CONF-9705121--

Proceedings of the

FIFTEENTH SYMPOSIUM ON ENERGY ENGINEERING SCIENCES

May 14-15, 1997

at

ARGONNE NATIONAL LABORATORY

Argonne, Illinois

Cosponsored by

Office of Basic Energy Sciences
U.S. DEPARTMENT OF ENERGY

and

Energy Technology Division
ARGONNE NATIONAL LABORATORY

Coordinated by

Argonne National Laboratory
9700 South Cass Avenue
Argonne, Illinois 60439

DISCLAIMER

FIFTEENTH SYMPOSIUM ON ENERGY ENGINEERING SCIENCES

FOREWORD

This Proceedings Volume includes the technical papers that were presented during the Fifteenth Symposium on Energy Engineering Sciences on May 14-15, 1997, at Argonne National Laboratory, Argonne, Illinois. The Symposium was organized into eight technical sessions, which included 32 individual presentations followed by discussion and interaction with the audience. A list of participants is appended to this volume.

The DOE Office of Basic Energy Sciences, of which Engineering Research is a component program, is responsible for the long-term mission-oriented research in the Department. It has the prime responsibility for establishing the basic scientific foundation upon which the Nation's future energy options will have to be identified, developed, and built. It is committed to the generation of new knowledge necessary for the solution of present and future problems of energy exploration, production, conversion, and utilization, consistent with respect for the environment.

Consistent with the DOE/BES mission, The Engineering Research Program is charged with the identification, initiation, and management of fundamental research on broad, generic topics addressing energy-related engineering problems. Its stated goals are: 1) to improve and extend the body of knowledge underlying current engineering practice so as to create new options for enhancing energy savings and production, for prolonging useful life of energy-related structures and equipment, and for developing advanced manufacturing technologies and materials processing with emphasis on reducing costs with improved industrial production and performance quality; and 2) to expand the store of fundamental concepts for solving anticipated and unforeseen engineering problems in the energy technologies.

In achieving these goals, the Engineering Research Program supports approximately 130 research projects covering a broad spectrum of topics cutting across traditional engineering disciplines with a focus on three areas: 1) mechanical sciences, 2) control systems and instrumentation, and 3) engineering data and analysis. The Fifteenth Symposium involved approximately one-fourth of the research projects currently sponsored by the DOE/BES Engineering Research Program.

The Fifteenth Symposium was held under the joint sponsorship of the DOE Office of Basic Energy Sciences and Argonne National Laboratory. Local arrangements were handled by Ms. Joan Brunsvold and Ms. Marianne Adair of the ANL Conference Services. Ms. Gloria Griparis of the ANL Office of Technical Communication Services was responsible for assembling these proceedings and attending to their publication.

I am grateful to all who contributed to the success of the program, particularly to the participants for their uniformly excellent presentations, their active involvement in discussions, and their infectious enthusiasm. The resulting interactions made this Symposium a most stimulating and enjoyable experience.

Robert Goulard, ER-15
Division of Engineering and Geosciences
Office of Basic Energy Sciences

FIFTEENTH SYMPOSIUM ON ENERGY ENGINEERING SCIENCES

May 14-15, 1997

Argonne National Laboratory

Argonne, IL

TABLE OF CONTENTS

Technical Session 3 - Mostly Optics

Technical Session 4 - Fluid Mechanics

Technical Session 5 - Nonlinear Fields

INTERFACIAL AREA TRANSPORT IN BUBBLY FLOW

M. Ishii, Q. Wu, S.T. Revankar, T. Hibiki, W.H. Leung, S. Hogsett, and A. Kashyap

Thermal Hydraulics and Reactor Safety Laboratory
Purdue University, West Lafayette, IN 47907 U.S.A.

ABSTRACT

In order to close the two-fluid model for two-phase flow analyses, the interfacial area concentration needs to be modeled as a constitutive relation. In this study, the focus was on the investigation of the interfacial area concentration transport phenomena, both theoretically and experimentally. The interfacial area concentration transport equation for air-water bubbly up-flow in a vertical pipe was developed, and the models for the source and sink terms were provided. The necessary parameters for the experimental studies were identified, including the local time-averaged void fraction, interfacial area concentration, bubble interfacial velocity, liquid velocity and turbulent intensity. Experiments were performed with air-water mixture at atmospheric pressure. Double-sensor conductivity probe and hot-film probe were employed to measure the identified parameters. With these experimental data, the preliminary model evaluation was carried out for the simplest form of the developed interfacial area transport equation, i.e., the one-dimensional transport equation.

I. INTRODUCTION

In multiphase flow modeling, the overwhelming difficulties encountered in a local instant formulation stem from the existence of deformable moving interfaces and the turbulent fluctuations of the field variables. Even for single phase flow without free interfaces, the exact solutions for the local instant fluctuations are beyond our present mathematical capability. Therefore, the common approaches to depict the macroscopic aspects of two-phase flow use averaging techniques (Vernier and Delhaye, 1965 [1]; M. Ishii, 1975 [2]; Boure, 1978 [3]) because the microscopic details of the interface behaviors and the fluid particle fluctuations are

rarely needed in engineering applications. In a certain sense the averaging techniques can be considered as a low-pass filtering process that excludes unwanted high frequency signals from local instant chaotic fluctuations. Subsequently, based on proper averaging operations, the well-known two-fluid model is established, which is widely accepted as the most sophisticated measure for two-phase flow analyses. In this model the averaged macroscopic properties of the flow field are considered separately in terms of two sets of conservation equations that govern the balance of mass, momentum and energy in each phase. However, the two averaged macroscopic fields are coupled to each other through certain phase interaction terms in the field equations to characterize the interfacial transfers. According to Ishii [2], the generalized interfacial interaction terms can be expressed in the following form:

$$\text{Interfacial Transfer Term} = a_i \times (\text{Driving Flux}), \tag{1}$$

where the Driving Flux depends on the transfer mechanisms and a_i is the interfacial area concentration, defined as the total interface area per unit mixture volume. In order to close the two-fluid model, these parameters need to be modeled individually. In this study, the focus is on the investigation of the interfacial area concentration transport phenomena, both theoretically and experimentally. The interfacial area concentration transport equation for air-water bubbly up-flow in a vertical pipe is developed, and the models for the source and sink terms are provided. The necessary parameters for the experimental studies are identified, including the local time-averaged void fraction, interfacial area concentration, bubble interfacial velocity, liquid velocity and turbulent intensity. Experiments are performed with an air-water mixture at atmospheric pressure. Double-sensor conductivity probe and hot-film probe are employed to measure the identified parameters. The experimental results provide a comprehensive data base for the model evaluation. At the present stage, the model evaluation is in progress with the detailed experimental data. The preliminary results are presented here for the simplest form of the developed interfacial area transport equation, i.e., the one-dimensional transport equation.

II. INTERFACIAL AREA TRANSPORT EQUATION

The interfacial area concentration is proportional to the fluid particle density, and thus its variation is primarily dominated by the fluid particle expansion, agglomeration and disintegration. Analogous to Boltzman's transport equation, Reyes (1989 [4]) proposed a Population Balance Approach (PBA) to develop a particle number density transport equation for chemically non-reacting, dispersed spherical fluid particles. For the interfacial area transport phenomena, Kocamustafaogullari and Ishii (1995 [5]) obtained the following form of the generalized particle number density transport equation:

$$\frac{\partial f(\mathbf{x}, \mathcal{V}, t)}{\partial t} + \nabla \cdot \left(f(\mathbf{x}, \mathcal{V}, t) \mathbf{v}_p \right) = \sum_j s_j(\mathbf{x}, \mathcal{V}, t) + s_{ph}(\mathbf{x}, \mathcal{V}, t), \tag{2}$$

where $f(\mathbf{x}, \mathcal{V}, t)$ is the local number density distribution function of the fluid particles with volumes between $\mathcal{V}$ and $\mathcal{V}+d\mathcal{V}$, while $\mathbf{v}_p(\mathbf{x}, \mathcal{V}, t)$ refers to the local time-averaged particle

velocity. The fluid particle sink or source rate due to phase change is denoted by $s_{ph}(\mathbf{x}, \mathcal{V}, t)$. Detailed treatment of this term may follow the approach suggested by Kocamustafaogullari and Ishii (1983 [6]), whereas the wall nucleation rate must be specified as a boundary condition. The other terms on the right-hand-side of Eq. (2), $s_j(\mathbf{x}, \mathcal{V}, t)$, represent the net rate of change in the fluid particle number density distribution function resulting from particle breakage and coalescence. Some phenomenological models for these source and sink terms were summarized by Prince and Blanch (1990 [7]) and Lafi and Reyes (1991 [8]). These models depend on the fluid particle volume and present the detailed insight of the driving mechanisms. However, the particle volume dependent transport equation described by Eq. (2) is too detailed for the averaged two-fluid model, in which the primary interest is in the average fluid particle behaviors. It is then advantageous to develop a transport equation averaged over all fluid particles. From the integration of Eq. (2) over the fluid particle volume, the overall local particle number density is given by:

$$\frac{\partial n(\mathbf{x}, t)}{\partial t} + \nabla \cdot \left(\bar{\bar{\mathbf{v}}}_p(\mathbf{x}, t) n(\mathbf{x}, t) \right) = \sum_j S_j(\mathbf{x}, t) + S_{ph}(\mathbf{x}, t), \tag{3}$$

where $n(\mathbf{x}, t)$ is the local particle number per unit mixture volume, and $\bar{\bar{\mathbf{v}}}_p(\mathbf{x}, t)$ is the average local particle velocity weighted by the particle number, which is identical to the local time-averaged particle velocity weighted by the void fraction if the sample size is sufficiently large and the internal circulation in each fluid particle is neglected.

In the two-fluid model, the parameter of interest is the interfacial area concentration rather than the bubble number density. To replace the bubble number density in Eq. (3) with the interfacial area concentration, the following geometric relation is applied:

$$n = \frac{\alpha}{\mathcal{V}_b} = \psi \left(\frac{a_i^3}{\alpha^2} \right) \tag{4}$$

where ψ is a factor depending on the shape of the bubbles, and $\mathcal{V}_b$ denotes the average bubble volume. For spherical bubbles ψ equals $1/(36\pi)$. Substituting Eq. (4) into Eq. (3) leads to:

$$\frac{\partial a_i}{\partial t} + \nabla \cdot \left(a_i \bar{\bar{\mathbf{v}}}_g \right) = \frac{1}{3\psi} \left(\frac{\alpha}{a_i} \right)^2 \left[\sum_j S_j + S_{ph} \right] + \left(\frac{2a}{3\alpha} \right) \left[\frac{\partial \alpha}{\partial t} + \nabla \cdot \left(\bar{\bar{\mathbf{v}}}_g \alpha \right) \right]. \tag{5}$$

The second term on the right-hand-side represents the effects of the variation in the mean bubble volume. Without phase change, this term is related to the gas phase continuity equation in the following manner:

$$\frac{\partial \alpha}{\partial t} + \nabla \cdot \left(\bar{\bar{\mathbf{v}}}_g \alpha \right) = -\frac{1}{\rho_g} \left[\frac{\partial \rho_g}{\partial t} + \left(\bar{\bar{\mathbf{v}}}_g \cdot \nabla \right) \rho_g \right]. \tag{6}$$

This term is important when the system pressure is low and the change in gas density is significant along the flow path, which is quite common especially in air-water experiments conducted under atmospheric pressure conditions. For steady-state bubbly flow, if the gas density decreases along the flow path, the bubble size grows resulting in larger interfacial area

concentration. By introducing Eq. (6) into Eq. (5), the interfacial area transport equation without considering phase change effects is readily reduced to:

$$\frac{\partial a_i}{\partial t}+\nabla\cdot\left(a_i\vec{\mathbf{v}}_g\right)=\frac{1}{3\psi}\left(\frac{\alpha}{a_i}\right)^2\left(\sum_j S_j\right)-\left(\frac{2a_i}{3\alpha}\right)\frac{1}{\rho_g}\frac{D_g\rho_g}{Dt}. \tag{7}$$

The remaining unknowns in the above equation are the source and sink terms due to bubble coalescence and breakage. According to the preliminary studies of Wu et al. (1997 [9]), three major mechanisms are responsible for bubble coalescence and disintegration in dispersed bubbly flow. These are the coalescence due to random collisions driven by turbulence, the coalescence due to wake entrainment, and the breakage upon the impact of turbulent eddies.

For the turbulence induced bubble coalescence, the net rate of change in the bubble number per unit mixture volume is given by [9]:

$$S_{RC}=-C_{RC}\left[\frac{\psi u_t a_i^4}{\pi\alpha^2}\right]\left[\frac{1}{\alpha_{\max}^{1/3}\left(\alpha_{\max}^{1/3}-\alpha^{1/3}\right)}\right]\left[1-\exp\left(-C\frac{\alpha_{\max}^{1/3}\alpha^{1/3}}{\alpha_{\max}^{1/3}-\alpha^{1/3}}\right)\right], \tag{8}$$

where, the subscript "*RC*" stands for Random Collision, C_{RC} is proportional to the coalescence efficiency of each random collision, and u_t is the root-mean-square of the turbulent velocity fluctuation caused by the eddies with a characteristic length comparable to the mean bubble size. It is postulated that smaller eddies do not provide considerable bulk motion to a bubble, while larger eddies transport groups of bubble without leading to significant relative motion for coalescence. The term in the first bracket is similar to the collision rate proposed by Coulaloglou and Tavlarides [10] in 1976 for a liquid-liquid droplet flow system, analogous to the particle collision model in an ideal gas. The term in the second bracket is due to the finite size of bubbles that cannot be neglected when compared to the mean free path. The maximum void fraction, α_{max}, is chosen to be 0.8, a value at the transition point between annular flow and slug flow (Wallis, 1969 [11]). In the final bracket, a modification factor is introduced to consider the effect of the mean bubble distance compared to the bubble traveling range. For the sparsely distributed bubbles, the collision probability decreases because the turbulent eddies that drive the bubbles toward each other cannot sustain such long range interactions. It should be emphasized that the coalescence efficiency in this study is assumed to be constant. The most popular model for the coalescence efficiency is the film thinning model developed by Oolman and Blanch (1986 [12]). However, their model suggests that the coalescence rate decreases exponentially with respect to the turbulent fluctuating velocity, which is much stronger than the linear dependence of the collision rate, resulting in an overall decreasing trend of the coalescence rate as the turbulent fluctuation increases. This may cause serious trouble when the model is applied to experimental data, especially at high liquid flow rates. Nevertheless, constant coalescence efficiency is only an approximation, further efforts are needed to model the efficiency mechanistically.

The bubble coalescence due to the so-called wake-entrainment mechanism may occur when bubbles enter the wake region of a leading bubble and then accelerate toward the preceding one. By considering the relative motion as the driving mechanism, the corresponding rate of change in bubble number density satisfies the following simple formula [9]:

$$S_{WE} = -\frac{\psi}{\pi} C_{WE} u_r \left(\frac{a_i^2}{\alpha} \right)^2 , \tag{9}$$

where u_r is the relative velocity between the gas bubble and the ambient liquid, given by:

$$u_r = \left(\frac{Dg}{3C_D} \frac{\Delta\rho}{\rho_f} \right)^{1/2} \tag{10}$$

with $$C_D = 24 \frac{\left(1 + 0.1 \mathrm{Re}_D^{0.75}\right)}{\mathrm{Re}_D}, \text{ and } \mathrm{Re}_D \equiv \frac{\rho_f u_r D}{\mu_f} (1-\alpha). \tag{11}$$

Again, the coefficient C_{WE} is proportional to the coalescence efficiency and is assumed to be constant here.

The final mechanism is the bubble disintegration due the impact of the turbulent eddies. From simple force balance on a gas bubble, the net rate of change in bubble number density is given by [9]:

$$S_{TI} = \frac{1}{18} C_{TI} u_t \left(\frac{a_i^2}{\alpha} \right) \left(1 - \frac{\mathrm{We}_{cr}}{\mathrm{We}} \right)^{1/2} \exp\left(-\frac{\mathrm{We}_{cr}}{\mathrm{We}} \right), \quad \mathrm{We} > \mathrm{We}_{cr}, \tag{12}$$

where, $$\mathrm{We} = \frac{6\rho_f u_t^2 \alpha}{\sigma a_i}. \tag{13}$$

The subscript "*TI*" refers to Turbulent Impact. The unique feature of this expression is that the breakup rate equals zero when the Weber number is less than a critical value, We_{cr}. At low liquid flow rates, the turbulent fluctuation is small and thus no breakup would be counted, which enables the fine-tuning of the adjustable parameters in the coalescence terms, independent of the bubble breakage at low liquid flow rate with small void fraction.

Eqs. (7), (8), (9) and (12) constitute the closure relations of the interfacial area concentration in two-phase vertical dispersed bubbly flow. The variables in these equations are coupled with the field equations in the two-fluid model. For the information of the local interfacial area concentration, the field equations should be solved together with the closure relations. On the other hand, these models also provide the detailed requirements to the experimental investigation for the interfacial area transport phenomena. In addition to the measurements of local time-averaged void fraction and interfacial area concentration, information about the bubble velocity, liquid turbulent intensity and bubble relative velocity should be obtained for the evaluation of the proposed models.

III. EXPERIMENTAL INVESTIGATION

For the purpose of studying the interfacial area transport process, a vertical cocurrent two-phase flow test facility was designed and constructed with three measurement ports at different axial positions. In order to obtain the parameters identified in the previous section for the characteristics of the interfacial area transport phenomena, double-sensor conductivity probe and hot-film probe were employed in the experimental investigations. Comprehensive measurements were carried out mainly in the bubbly flow regime. For every flow condition, the radial distributions of each parameter were obtained at three axial positions. These parameters include the local time-averaged liquid velocity, liquid turbulent intensity, gas velocity, void fraction, interfacial area concentration, and bubble Sauter mean diameter.

3.1. Test Facility and Instrumentation

The schematic of the test loop is depicted in Fig. 1. The test section is a 3.6 m long acrylic pipe (5.08 cm ID). Air and deionized water are used in the experiments. With a 1.5 hp water pump (SP15-150, Price Pump Company), the liquid superficial velocity in the test section can reach about 2.0 m/s, whereas the lab's stable 0.7 MPa dry air supply is sufficient for the experiment up to the churn-turbulent flow regime. The uniform two-phase mixture enters the test section from an air-water mixing chamber with a conical section connected to the bottom end of the test pipe. Inside the air-water mixer, an air-injector consists of 400 hypodermic needles held together by a resin base as shown in Fig. 1, providing bubbles with diameters of 3 mm to 5 mm. In the upper plenum, water is separated from air and flows back to the water tank, where a heat exchanger is installed to maintain a constant liquid temperature for the hot-film probe measurement. Along the flow path in the test section, three instrument ports are equipped at the axial positions of L/D equal to 2, 32 and 62, and a traversing mechanism is installed at each probe port. The traversing mechanism utilizes a micrometer that is accurate to 0.025 of a millimeter. Since the existence of the probe support can induce bubble breakage and coalescence and therefore affect the downstream measurements, the measurement at the four probe locations have to be performed individually. For the measurements of the time-averaged local void fraction, interfacial area concentration and bubble velocity, a double-sensor conductivity probe is developed, whereas the water velocity and turbulent intensity measurements are made using a hot-film anemometer system (FLOWPOINT, TSI Inc.). The local instantaneous data are acquired with a 486 PC, and the time-averaged variables are obtained through software developed in the TRSL of Purdue University.

The double-sensor probe is made of two thin electrodes as shown schematically in Fig. 2a. The tip of each electrode is exposed to the two-phase mixture and measures the impedance between the probe tip and the common ground. Due to the significant difference in the conductivities of the liquid phase and the gas phase, the impedance signal rises sharply when a bubble passes through the probe (Fig. 2b). From the front tip signal, the bubble residual time fraction is the local time-averaged void fraction. On the other hand, the time delay, Δt, of the two impedance signals can thus be utilized to characterize the time interval for the bubble surface

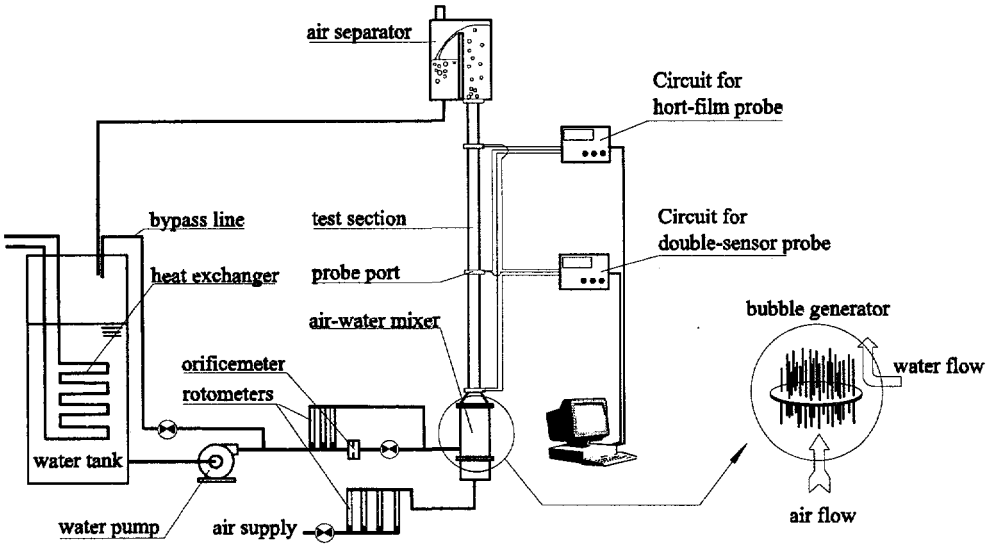


Fig. 1 Schematic diagram of the experimental loop

traveling from the front probe tip to the back tip. Since the separation of the two probe tips is known, defined as Δs, a measurable bubble axial velocity, Δs/Δt, is obtained. For the measurement of the local time-averaged interfacial area concentration, according to Wu and Ishii (1997 [13]), the following formula is suggested in considering the effects of the bubble lateral motion and the probe spacing:

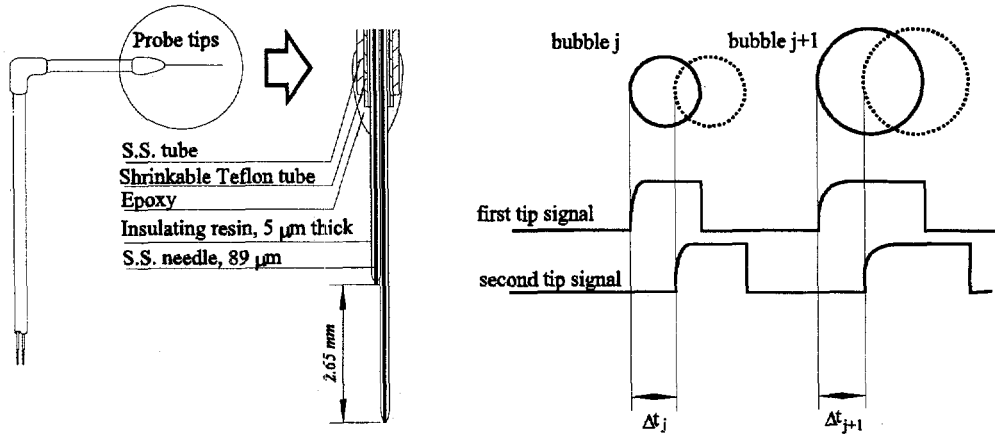
$$a_i = \left(\frac{2N_b}{\Delta T\left(N_b - N_{miss}\right)}\right)\left[2+\left(\frac{V_b'}{\overline{V}_b}\right)^{2.25}\right]\sum_j\left(\Delta t_j/\Delta s\right), \qquad D = 1.2\Delta s \sim 3\Delta s, \tag{14}$$

where the subscript j denotes the j-th measured bubble, N_b stands for the number of the bubbles that pass the first probe tip in the ΔT time interval for time-averaging, D refers to the Sauter mean diameter of the bubbles, and N_{miss} is the number of the missed bubbles. A bubble is defined as missed when the second probe tip either cannot touch the bubble or enters the bubble before the first tip. These cases may occur if the bubble has lateral velocity components [13]. If the Sauter mean diameter of the measured bubbles is in the range of 1.2 to 3 times the probe spacing and the sample size is sufficiently large, the error of this expression is within ±1.5%. The term $\left(V_b'/\overline{V}_b\right)$ in Eq. (14) represents the relative bubble velocity fluctuation, given by [13]:

$$\left(\frac{V_b'}{\overline{V}_b}\right) = 1.2\frac{\sigma_{\Delta t_j}}{\left(\Delta t_j\right)}, \qquad D = 1.2\Delta s \sim 3\Delta s, \tag{15}$$

where $\sigma_{\Delta tj}$ is the standard deviation of Δt_j. Moreover, in order to control the statistic fluctuations within ±5%, the sample for time-averaging should contain no less than 1000 bubbles [13]. With the measured local time-averaged void fraction and interfacial area concentration, if the bubbles are assumed to be spherical, the bubble Sauter mean diameter is given by the following relation:

$$D = \frac{6\alpha}{a_i}. \tag{16}$$



(a) Schematic of double-sensor probe (b) Signals form the probe

Fig. 2 Double-sensor conductivity probe

The hot-film probe used in the experiment for the measurements of the water velocity and turbulent intensity is a constant temperature probe with a conical tip (Model 1231W, TSI Inc.), which measures the one-dimensional liquid velocity with a maximum frequency response of 5 kHz. The calibration of the hot-film probe was performed with either an orifice flowmeter or a Pitot-static probe positioned at the test section centerline. In order to maintain constant liquid temperature, a cooling coil was installed in the water tank as shown in Fig. 1, which controls the temperature fluctuation within ±1°C. Typically, the loop is allowed to run for an hour or so until a steady-state temperature is reached. For small temperature fluctuations, a thermocouple is placed in the test section and the FLOWPOINT software adjusts the hot-film output signal accordingly.

The hot-film probe is designed for single phase water flow at an operating temperature of 66.7°C. When applied to air-water two-phase bubbly flow, however, the operation temperature should be reduced in order to achieve consistent probe sensitivity (Hogsett, 1996 [14]; Hibiki, 1997 [15]). Otherwise a frequent output voltage drift may occur due to the dramatic change in local convection heat transfer coefficients between the liquid phase and the gas phase. In the experiment, the operating temperature was chosen at 45°C and the probe calibration was carried out using the orifice flow meter installed in the liquid supply line as shown in Fig. 1. Even at low temperature conditions, significant drift of the calibration curve is observed. Hence, frequent recalibrations are conducted. In Fig. 3, the typical time history of a hot-film probe signal for a two-phase bubbly flow is presented. Besides the chaotic fluctuations caused by the turbulence in the liquid phase, some voltage spikes exist due to the bubbles passing through the probe tip. In order to obtain a liquid velocity representative of the flow, it is necessary to filter out these spikes. This has been done using a threshold scheme developed by Hogsett [14] and the filtered signal is also shown in Fig. 3. From the signal of the hot-film probe, average liquid velocity as well as the root-mean-square velocity fluctuations are obtained. Moreover, the filtered signal gives the local bubble residual time and provides an alternative measure to calibrate the void fraction measurement from the double-sensor conductivity probe.

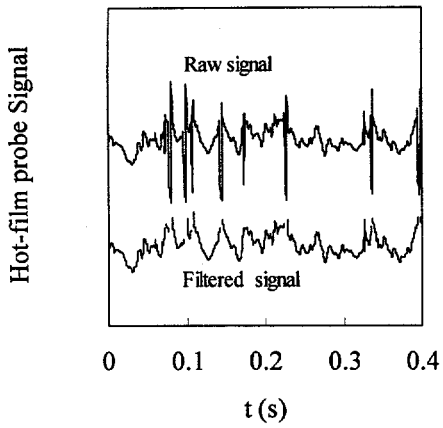


Fig. 3 Raw and filtered hot-film signals

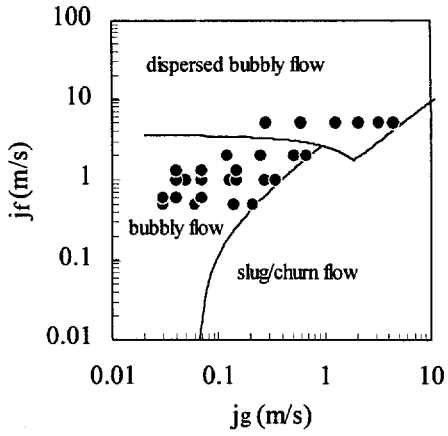


Fig. 4 Flow conditions for the tests

3.2 Typical Test Results

Comprehensive measurements were carried out mainly in the bubbly flow regime as shown in Fig. 4. For every flow condition, the radial distributions of each parameter were obtained at three axial positions. These parameters include the local time-averaged liquid velocity, liquid turbulent intensity, gas velocity, void fraction, interfacial area concentration, and the bubble Sauter mean diameter. The liquid velocity and turbulent intensity were measured with a hot-film anemometer, whereas the bubble velocity, void fraction, interfacial area concentration, and bubble Sauter mean diameter were acquired using a double-sensor conductivity probe.

Typical liquid velocity profiles at different axial locations are shown in Fig. 5. Generally, the liquid velocity profiles are flat, presenting a typical turbulent flow characteristics. Under the specific flow condition, a subtle wall peak is observed at the pipe entrance region with L/D equal to 2, which corresponds to a void fraction wall peak in Fig. 7. On the other hand, the turbulent intensity profile appears to depend on the relative difference between the bubble-induced turbulence and the turbulence generated by wall shear. Based on the data obtained by Hogsett [14] at low liquid flow rates, the turbulent intensity peaks in the core of the flow where bubbles tend to cluster, even for the lowest void fraction case with $\langle\alpha\rangle$=0.018. This evidence implies that the bubble-induced turbulence seems to be the dominant mechanism at low liquid flow rates. As the liquid flow rate increases, however, wall shear effects gradually overtake the bubble-induced turbulence, and the turbulent intensity profiles at different axial positions shown in Fig. 6 tend to collapse to a fixed line that resembles the behavior of a single phase turbulent flow.

All the void fraction profiles at different flow conditions exhibit similar trends, as shown in Fig. 7. The radial void fraction profile peaks near the wall. Along the flow path, the peak moves away from the wall and decreases considerably. The variation of the peak depends on the flow conditions. At low liquid flow rates the peak may decrease by 40% to 50%, while at high liquid flow rates the decrease is about 10%. In Fig. 8, the radial interfacial area concentration profile

exhibits the same trends as the void fraction profile. Hence, the bubble Sauter mean diameter, which is proportional to the ratio of the void fraction to the interfacial area concentration, remains approximately constant within the pipe cross-sectional area (Fig. 9). However, as the bubbles move upward, their average size becomes larger because of expansion and agglomeration. Moreover, the bubble velocity profiles measured by the double-sensor probe are presented in Fig. 10. Near the entrance, at L/D equal to 2, the bubble velocity is slightly higher in the wall region. Downstream, at L/D equal to 32 and 62, the bubble velocity reaches its terminal velocity and the velocity profiles resemble the liquid velocity profile.

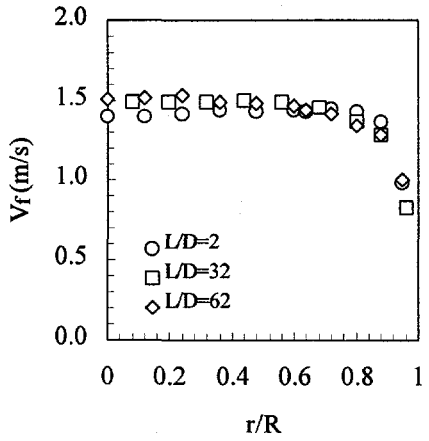


Fig. 5 Liquid velocity profile (jg=0.039m/s jf=1.3m/s)

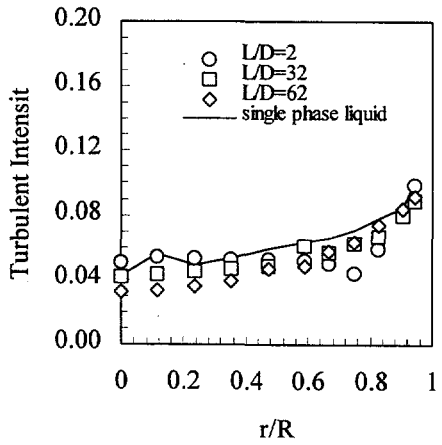


Fig. 6 Turbulent intensity profile (jg=0.039m/s jf=1.3m/s)

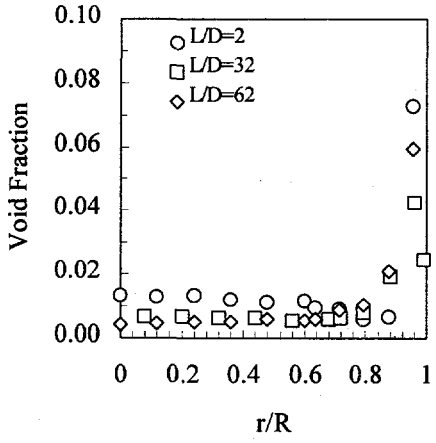


Fig. 7 Void fraction profile (jg=0.039m/s jf=1.3m/s)

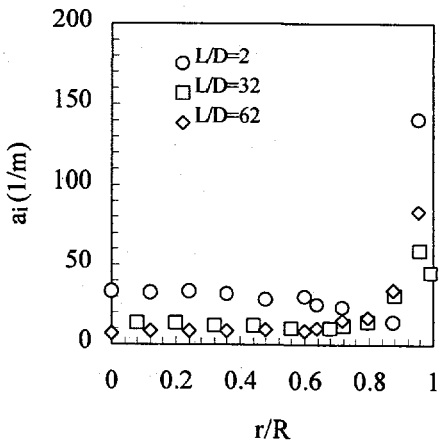


Fig. 8 Interfacial area concentration profile (jg=0.039m/s jf=1.3m/s)

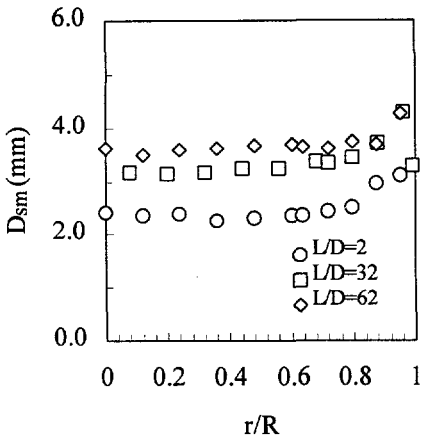


Fig. 9 Bubble Sauter mean diameter profile (jg=0.039m/s jf=1.3m/s)

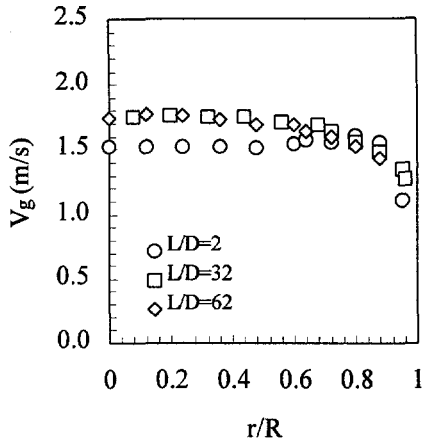


Fig. 10 Bubble velocity profile (jg=0.039m/s jf=1.3m/s)

IV. PRELIMINARY MODEL EVALUATIONS

At the present stage, the modeling evaluation is in progress with the detailed experimental data presented in the previous section. Here, the preliminary results are presented for the simplest form of the developed interfacial area transport equation, i.e., the steady-state, one-dimensional transport equation obtained by applying cross-sectional area averaging over Eq. (7):

$$\frac{\partial}{\partial z}\left(\langle a_i \rangle \langle \langle v_{gz} \rangle \rangle\right) = \langle S_{a,RC} \rangle + \langle S_{a,WE} \rangle + \langle S_{a,TI} \rangle, \tag{17}$$

where the bubble expansion effects are neglected and the gas phase is assumed to be incompressible. Due to the uniform bubble size assumption, the area-averaged bubble interface velocity weighted by interfacial area concentration is the same as the conventional area-averaged gas velocity weighted by the void fraction, if the internal circulation in bubbles is neglected. The exact mathematical expressions for the area-averaged source and sink terms would involve many covariances, which may further complicate the one-dimensional problem. However, since these local terms were originally obtained from a finite mixture element, the functional dependence of the area-averaged source and sink terms on the averaged parameters may be approximately the same if the hydraulic diameter of the flow path is considered as the length scale of the finite element. Therefore, Eqs. (8), (9), and (12) with the parameters averaged within the cross-sectional area are applicable for the area-averaged source and sink terms Eq. (17)

Following the procedure suggested by Wu et al. [9], seven cases of tests (Kashyap, 1994 [16]) with identical initial bubble injection mechanisms are employed for the preliminary model evaluation. With the measured interfacial area concentration at L/D=2 as the initial condition, Eq. (17) is integrated numerically to predict the axial distribution, and the adjustable parameters C_{RC}, C, C_{WE}, C_{TI} and We_{cr} are determined once the predictions at the other two locations match

the experimental data. The final results of these parameters are listed in Table 1. It should be mentioned that the coalescence efficiency is assumed to be constant in this evaluation. For very small bubbles, the chance for bubble coalescence upon each collision may be smaller, and thus the coalescence efficiency should be investigated.

Table 2 Adjustable parameters

C_{RC}	C	C_{WE}	C_{TI}	We_{cr}
0.0565	3	0.151	0.18	2.0

Compared to the experimental data, all 7 cases are shown in Fig. 11a and 11b. The model predicting interfacial area concentration is generally in good agreement with the measurements. The maximum relative difference is about 8% at very small void fraction with high liquid flow rate. Nevertheless, the conclusions are based on the simplified one-dimensional model with constant coalescence efficiency. Fine-tuning of these models and the adjustable parameters is needed.

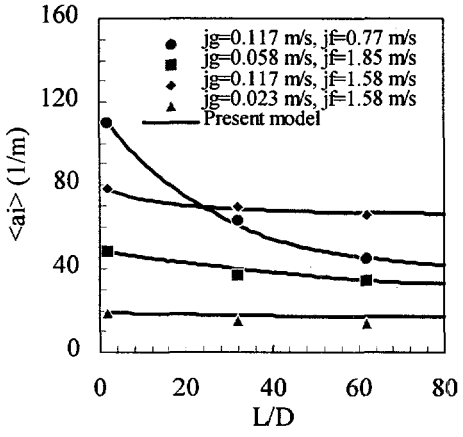


Fig. 11a a_i versus L/D

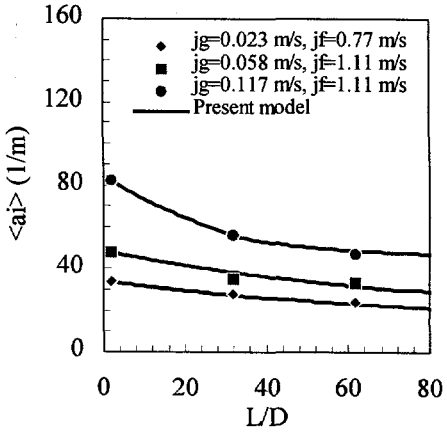


Fig. 11b a_i versus L/D

V. CONCLUSIONS

In this study, the efforts focused on the investigation of the interfacial area concentration transport phenomena, both theoretically and experimentally. Starting from the bubble number transport equation, the interfacial area concentration transport equation for the air-water bubbly up-flow in a vertical pipe was developed. The models for the source and sink terms due to bubble agglomeration and disintegration were provided. For bubble coalescence, two mechanisms are considered to be important in bubbly flow, i.e., the random collisions caused by turbulence and the wake-entrainment process. For bubble disintegration, the impact of turbulent eddies is included. From these mechanistic models, the necessary parameters for the experimental studies were identified, including the local time-averaged void fraction, interfacial

area concentration, bubble interfacial velocity, liquid velocity and turbulent intensity. Experiments were performed with an air-water mixture at atmospheric pressure. Double-sensor conductivity probe and hot-film probe were employed to measure the identified parameters. Detailed local time-averaged measurements were obtained at three axial positions and the experimental results provided a comprehensive data base for the model evaluation. At the present stage, the model evaluation is in progress with the detailed experimental data. Only the preliminary results were presented for the simplest form of the developed interfacial area transport equation, i.e., the one-dimensional transport equation. For application to three-dimensional problems, however, the adjustable parameters need to be evaluated. It should be emphasized again that the coalescence efficiency in the models was assumed to be constant. In reality, the liquid surface tension and bubble size have strong influences on the coalescence probability, and thus this assumption may not be universal. Further studies need to be performed to understand the bubble coalescence mechanism upon each collision. Moreover, the present transport equation is applicable only to bubbly flow. When cap/slug bubbles appear, another group transport equation should be constructed to depict the cap/slug bubble behaviors. In such a situation, some inter-group interactions would play important roles in the interfacial area transport process.

ACKNOWLEDGMENTS

This study was performed under the auspices of the U.S. Department of Energy. The authors would like to express their sincere appreciation to Dr. O.P. Manley of DOE/BES for his support of this project.

REFERENCES

1. P. Vernier and J.M. Delhaye, "General Two-Phase Flow Equations Applied to the Thermohydrodynamics of Boiling Nuclear Reactor," Engr. Primaire, Vol. No. 1, 5 (1968).
2. M. Ishii, *Thermo-Fluid Dynamic Theory of Two-Phase Flow*, Collection de la Direction des Etudes et Reserches d'Electricité de France, Eyrolles, Paris (1975).
3. J.A. Boure, "Mathematical Modeling of Two-Phase Flows," Proc. of CSNI Specialist Meeting, S. Banerjee and K.R. Weaver, Eds. A.E.C.L., Vol. 1, 85, Aug. 3-4, Toronto (1978).
4. J.N. Reyes, "Statistically Derived Conservation Equations for Fluid Particle Flows," 5th Proc. ANS-THD, P. 12, 1989 ANS Winter Meeting, San Francisco, CA, Nov. (1989).
5. G. Kocamustafaogullari and M. Ishii, "Foundation of the Interfacial Area Transport Equation and its Closure Relation," *Int. J. Heat and Mass Transfer*, Vol. 38, No. 3, 481 (1995).
6. G. Kocamustafaogullari and M. Ishii, "Interfacial Area and Nucleation Site Density in Boiling Systems," *Int. J. Heat and Mass Transfer*, Vol. 26, 1377 (1983).
7. M.J. Prince and H.W. Blanch, "Bubble Coalescence and Break-Up in Air-Sparged Bubble Columns," *AIChE Journal*, Vol. 36, No. 10, 1485 (1990).

8. A.Y. Lafi and J.N. Reyes, Jr., "Phenomenological Models for Fluid Particle Coalescence and Breakage," OSU-NE-9120, Report of the Department of Nuclear Engineering, Oregon State University, Corvallis, Oregon 97331, USA (1991).

9. Q. Wu, S. Kim, M. Ishii, and S.G. Beus, "One-Group Interfacial Area Transport in Vertical Air-Water Bubbly Flow," Accepted by the 1997 National Heat Transfer Conference to be held in Baltimore, Maryland, Aug. 10-12 (1997).

10. C.A. Coulaloglu and L.L. Tavlarrides, "Drop Size Distributions and Coalescence Frequencies of Liquid-Liquid Dispersion in Flow Vessels," *AIChE J.*, Vol. 22, 289 (1976).

11. G.B. Wallis, *One-dimensional Two-phase Flow*, McGraw-Hill Book Company, NY. 350 (1969).

12. T. Oolman and H.W. Blanch, "Bubble Coalescence in Stagnant Liquids," *Chem. Eng. Commun.*, Vol. 43, 237 (1986).

13. Q. Wu and M. Ishii, "Sensitivity Studies on Double-Sensor Conductivity Probe for the Measurement of Interfacial Area Concentration in Bubbly Flow," submitted to the *Int. J. Multiphase Flow*, (1997).

14. S. Hogsett, "The Axial Development of Liquid Turbulence and Interfacial Area in Bubbly Two-Phase Flow," M.S. Thesis, Department of Nuclear Engineering, Purdue University, (1996).

15. T. Hibiki, S. Hogsett, and M. Ishii, "Local Measurement of Interfacial Area, Interfacial Velocity and Liquid Turbulence in Two-Phase Flow," OECD/CSNI Specialist Meeting on Advanced Instrumentation and Measurement Techniques, Santa Barbara, March 17-20 (1997).

16. A. Kashyap, M. Ishii and S.T. Revankar, "An Experimental and Numerical Analysis of Structural Development of Two-Phase Flow," PU-NE-94/2, Report of the Nuclear Engineering Department, Purdue University, West Lafayette, IN 47907, USA (1994).

ASPECTS OF SUBCOOLED BOILING

S. G. Bankoff

Chemical Engineering Department
Northwestern University
Evanston, IL 60208

ABSTRACT

Subcooled boiling boiling refers to boiling from a solid surface where the bulk liquid temperature is below the saturation temperature (subcooled). Two classes are considered: 1) nucleate boiling, where, for large subcoolings, individual bubbles grow and collapse while remaining attached to the solid wall, and 2) film boiling, where a continuous vapor film separates the solid from the bulk liquid. One mechanism by which subcooled nucleate boiling results in very large surface heat transfer coefficient is thought to be latent heat transport within the bubble, resulting from simultaneous evaporation from a thin residual liquid layer at the bubble base, and condensation at the polar bubble cap. Another is the increased liquid microconvection around the oscillating bubble. Two related problems have been attacked. One is the rupture of a thin liquid film subject to attractive and repulsive dispersion forces, leading to the formation of mesoscopic drops, which then coalesce and evaporate. Another is the liquid motion in the vicinity of an oscillating contact line, where the bubble wall is idealized as a wedge of constant angle sliding on the solid wall. The subcooled film boiling problem has been attacked by deriving a general long-range nonlinear evolution equation for the local thickness of the vapor layer. Linear and weakly-nonlinear stability results have been obtained. A number of other related problems have been attacked.

INTRODUCTION

Subcooled boiling refers to boiling from a hot solid surface in the presence of subcooled bulk liquid. Two major regimes can be distinguished: a) subcooled nucleate boiling, in which myriads of small bubbles grow and collapse while remaining attached to the heating surface, and b) subcooled film boiling, in which a continuous film (except possibly for transient and limited contacts) separates the hot wall from the subcooled bulk liquid. The transition from nucleate boiling to film boiling occurs at some critical value of

the surface heat flux, and may lead to disastrous overheating of the solid wall. On the other hand, subcooled nucleate boiling is a highly efficient mode of heat transfer, with reported heat fluxes in the neighborhood of 3×10^3 w/cm^2, or more in some cases. On the other hand, film boiling is much less efficient, although with sufficient liquid subcooling the quasi-steady film thickness may be quite small. At a critical surface temperature, known as the minimum film boiling temperature, the vapor film collapses, and the transition to nucleate boiling occurs at constant heat flux.

The mechanism whereby such large heat fluxes are attained in subcooled nucleate boiling is naturally of great interest. Two major theories have been proposed: 1) increased turbulent mixing and transient liquid-wall contact owing to the presence of the growing and collapsing bubbles, 2) latent heat transport within the bubbles, in which evaporation occurs from a thin liquid residual layer left behind by the bubble growth, and simultaneous condensation at the colder polar region of the bubble in contact with the bulk liquid. Arguments have been advanced for and against these theories [1]. Better understanding of the basic mechanisms is needed. Some specialized problems have therefore been attacked in order to make progress on understanding the overall mechanism of subcooled nucleate boiling. The first examines the rupture of the residual microlayer at the base of the bubble while it is attached to the wall [2]. The formation of an array of isolated microdrops connected by extremely thin liquid film is predicted. These drops evaporate while coalescing, and eventually disappear. A second problem addresses the liquid motion in the vicinity of the oscillating contact line as the bubble grows and collapses periodically. The bubble wall is idealized as a wedge of constant contact angle moving with respect to the solid wall [3]. Approximate solutions are obtained, using double (Laplace and Mellin) transforms, showing attainable regions of parameter space, and flow reversals even at very small Strouhal numbers (dimensionless frequencies).

At the other end of the spectrum, the mechanism of destabilization of subcooled film boiling is poorly understood. There is a steady base-state thickness at which condensation exactly balances evaporation for a vapor film separating a plane heating surface from the bulk subcooled liquid [4]. To make progress on this problem, one can assume an average heat transfer coefficient at the vapor-liquid surface. Small perturbations from the base state can be studied by linear theory, but contact with the solid wall involves wave amplitudes which are of the same order as the base vapor film thickness. Lubrication theory, as in the treatment of the stability of thin liquid films, can then be employed prior to contact, since long-wave theory is appropriate. An interesting difference between the present problem and the instability of an evaporating liquid film is that vapor recoil effects, owing to the spatial variations of the reactive pressure on the evaporating liquid surface, are stabilizing for the vapor film, but destabilizing for the liquid film. A weakly nonlinear analysis reveals a supercritical branch of new equilibrium solutions corresponding to wave-like modulations of the uniform film. These disturbances satisfy the Newell-Whitehead-Segel (NWS) equation on an unbounded heating surface. Work on this problem is continuing.

SUBCOOLED NUCLEATE BOILING

Unsteady Stokes flow near an oscillating contact line [3].

Consider the flow in a wedge of incompressible fluid bounded by a solid surface and

a passive gas, as shown in Figure 1. We assume that the no-slip boundary condition is relaxed near the contact line, so that the stress singularity is removed by posing a slip velocity. The origin of the polar-coordinate system is placed at the contact line. The flow within the wedge is governed by the Navier-Stokes equations with continuity.

$$\nabla \cdot \mathbf{u} = 0 , \tag{1}$$

$$\rho(\mathbf{u}_t + \mathbf{u} \cdot \nabla\mathbf{u}) = -\nabla p + \upsilon\nabla^2\mathbf{u} , \tag{2}$$

with interfacial conditions of continuity of shear stress, the jump in normal stress balanced by surface tension σ times the curvature κ, and the kinematic boundary condition at $\theta = \alpha(r,t)$

$$[\mathbf{n} \cdot \mathbf{T} \cdot \mathbf{T} = 0 , \tag{3}$$

$$[\mathbf{n} \cdot \mathbf{T} \cdot \mathbf{t}] = \sigma\kappa , \tag{4}$$

$$\left[1 + (r\frac{\partial\alpha}{\partial r})^2\right]^{1/2} \alpha_t + u \cdot n = 0 , \tag{5}$$

where [f] is the jump in the quantity f across the interface. A tangential velocity profile on the solid boundary at $\theta = 0$ is prescribed

$$u(r,t) = U_o(r) \cos \omega t , \tag{6}$$

$$v(r,t) = 0 , \tag{7}$$

where (u,v) are the radial and azimuthal velocities respectively. The prescribed velocity profile $U_o(r)$ is assumed to be zero at $r = 0$, and approaches a general constant velocity amplitude U as $r \rightarrow \infty$. The transition scale between these velocities occurs on a prescribed slip length-scale L_s. This prescribed condition is an assumption on the plate velocity.

Lengths are scaled on this slip length scale L_s, velocities on the amplitude of the oscillating boundary U, time on the period of the oscillation $1/\omega$, along with a Stokes pressure scale $\mu U/L_s$ to arrive at the scaled system

$$\nabla \cdot \mathbf{u} = 0 , \tag{8}$$

$$S\ \mathbf{u}_t + Re\ \mathbf{u} \cdot \nabla\mathbf{u} = -\nabla p + \nabla^2\mathbf{u} , \tag{9}$$

where $S = L^2\omega/\upsilon$ is the Strouhal number, and $Re = UL/\upsilon$ is the Reynolds number. The interfacial conditions at $\theta = \alpha(r,t)$ become

$$[\mathbf{n} \cdot \mathbf{T} \cdot \mathbf{t}] = 0 \tag{10}$$

$$Ca[\mathbf{n} \cdot \mathbf{T} \cdot \mathbf{t}] = \sigma \tag{11}$$

$$\left[1 + (r\frac{\partial\alpha}{\partial r})^2\right]^{-1/2} \alpha_T + u \cdot n = 0 \tag{12}$$

where $Ca = \mu U/\sigma$ is the capillary number and

$$\kappa = \frac{1}{r}\frac{d}{dr}\left[r^2\frac{d\alpha}{dr}\right] + r^2\left[\frac{d\alpha}{dr}\right]^3\left\{1 + r\left[\frac{d\alpha}{dr}\right]\right\}^{-3/2} \tag{13}$$

The prescribed tangential velocity profile on the solid boundary at $\theta = 0$ becomes

$$u(r,t) = u_o(r) \cos t \tag{14}$$

$$v(r,t) = 0 , \tag{15}$$

with the prescribed velocity $u_o(r) \rightarrow 1$ as $r \rightarrow \infty$.

In studying local effects near a contact line that is strongly unsteady, it is assumed that inertial effects are small ($Re = 0$) but time variations are significant ($S = O(1)$). Capillary effects are also neglected ($Ca = 0$) and hence also the dynamics of the interface ($\alpha(r,t) = \alpha_0$). With these assumptions, a stream function formulation is used to consider the following problem:

$$S(\nabla^2\psi)_t = \nabla^4\psi \tag{16}$$

with, at $\theta = 0$,

$$\psi = 0 \tag{17}$$

$$\psi_{\theta\theta} = 0 . \tag{18}$$

By using Mellin and Laplace transforms in r and t, one obtains a stream function solution, whose poles provide the eigenvalues for each of the characteristic modes. One finds that nonuniformities are always present when there is both a steady forcing component and an arbitrarily small amplitude of an oscillatory component. One further finds that the drag force on the plate increases as the frequency of the oscillation increases.

NONLINEAR DYNAMICS OF ULTRA-THIN LIQUID FILMS

The breakdown and evaporation of the thin residual film left on the solid wall as the bubble grows is modeled as the spatio-temporal evolution of a two-dimensional thin volatile liquid layer, subjected to long-range molecular attractive forces and short-range repulsive forces. The evaporating film is described in dimensionless form by the nonlinear evolution equation

$$h_t + \frac{E}{h + K} + [(Ah^{-1} - BH^{-2})h_x]_x + S(h^3 h_{xxx})_x = 0, \tag{19}$$

provided that the disjoining/conjoining pressure is represented by the dimensionless potential

$$\Pi = \frac{A}{3}h^{-3} - \frac{B}{4}h^{-4}, \tag{20}$$

where t is time and x is the spatial coordinate along the solid substrate supporting the film. Here A,B are, respectively, dimensionless Hamaker constants for attractive long-range and repulsive short-range molecular forces, E is the dimensionless evaporation number and K is a thermal resistance at the interface owing to evaporation.

Linearizing the rescaled equations for a non-volatile liquid (E=0) yields

$$\omega = (\frac{1}{h_o} - \frac{\beta}{h_0^2})k^2 - h_0^3 k^4, \tag{21}$$

where k and ω are the wave number and growth rate of the interfacial disturbance, which is proportional to $\exp(\omega t + ik_x x + ik_y y)$, $k = \sqrt{k_x^2 + k_y^2}$. It follows from Eq. (21) that destabilization of the interface is due to attractive molecular forces, while its stabilization stems from both repulsive molecular and capillary forces.

Figure (2a,b,c,d) display the evolution of a volatile liquid film subject to a relatively slow evaporation rate, E = 0.5, and a finite interfacial resistance K = 1. The nucleation of the "dry spot" begins at t = 1.835 and the process of dewetting is accompanied by receding of the liquid, which is due to the squeeze effect of attractive molecular forces, Fig. (2a), formation of the rims and their disintegration into drops separated by new dry spots, Fig. (2b). Later, evaporation of the liquid and widening of the dry spots lead to the disappearance of smaller drops and growth of the larger ones, Fig (2c). Further, the larger drops coalesce due to evaporation and recession of the rims walls, Fig. (2d). Once a single large drop is formed, the dominant effect is evaporation and flattening of the interface. The last stage of the evolution begins with the last curve shown in Fig. (2d) which represents an almost flat interface. The film then disappears uniformly in space.

ACKNOWLEDGMENTS

This work was supported by Grant DE-FG02-86ER 13641, S.G. Bankoff and S.H. Davis, Co-Principal Investigators. Contributors were A. Oron, Visiting Scholar, B.S. Tilley, Applied Mathematics, New Jersey Institute of Technology, and C.A. Panzarella, graduate student.

REFERENCES

1. G. TSUNG-CHANG, and S.G. BANKOFF, "On the Mechanism of Forced-Convection Subcooled Nucleate Boiling," *J. Heat Transfer*, 112, 213-217 (1990).

*2. A. ORON and S.G. BANKOFF, "Nonlinear Dynamics of Ultra-Thin Liquid Films," manuscript in preparation (1997).

*3. B.S. TILLEY, S.H. DAVIS and S.G. BANKOFF, "Unsteady Stokes Flow Near an Oscillating Contact Line, submitted to *J. Fluid Mech.* (1997).

*4. C.H. PANZARELLA, S.G. BANKOFF and S.H. DAVIS, "A Long-Wave Analysis of Film Boiling," manuscript in preparation (1997).

*Work supported by grant in report period.

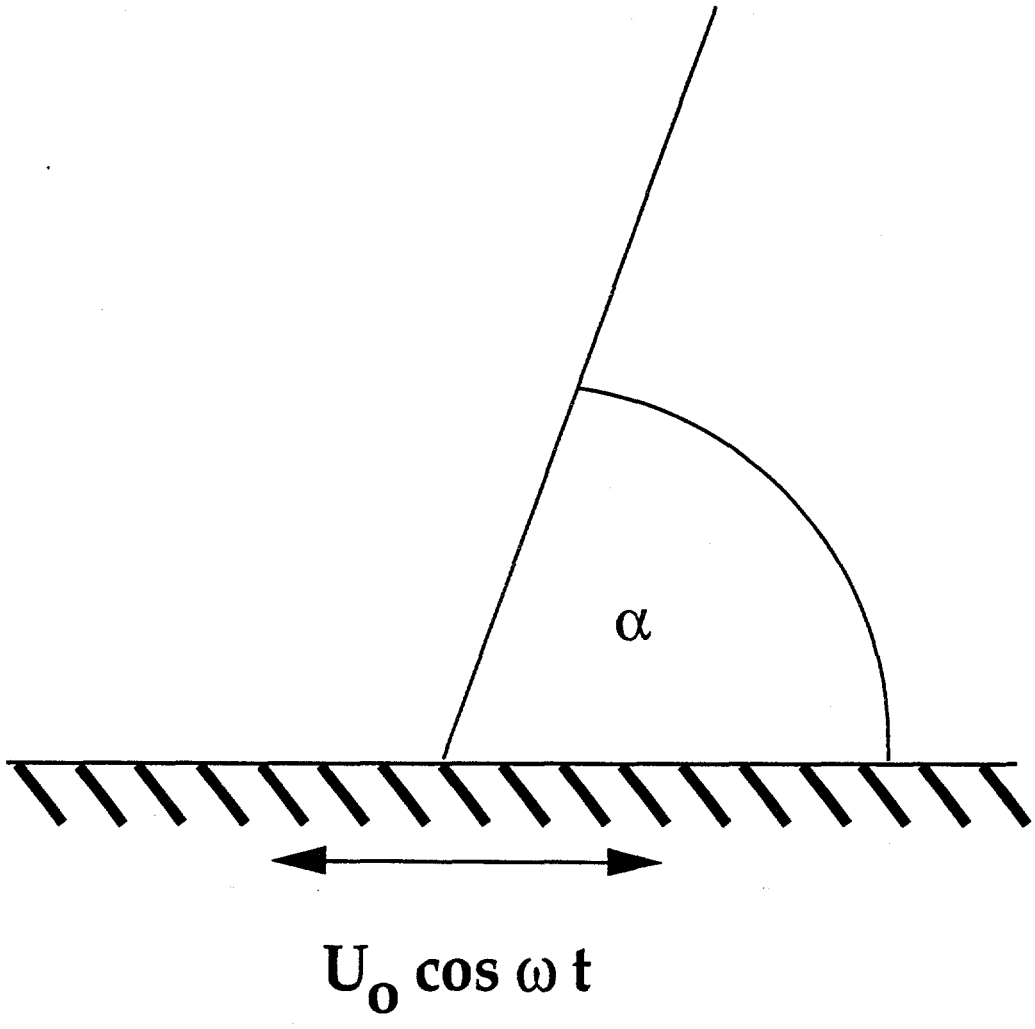


Fig. 1 Oscillating Wedge

Fig. 2a

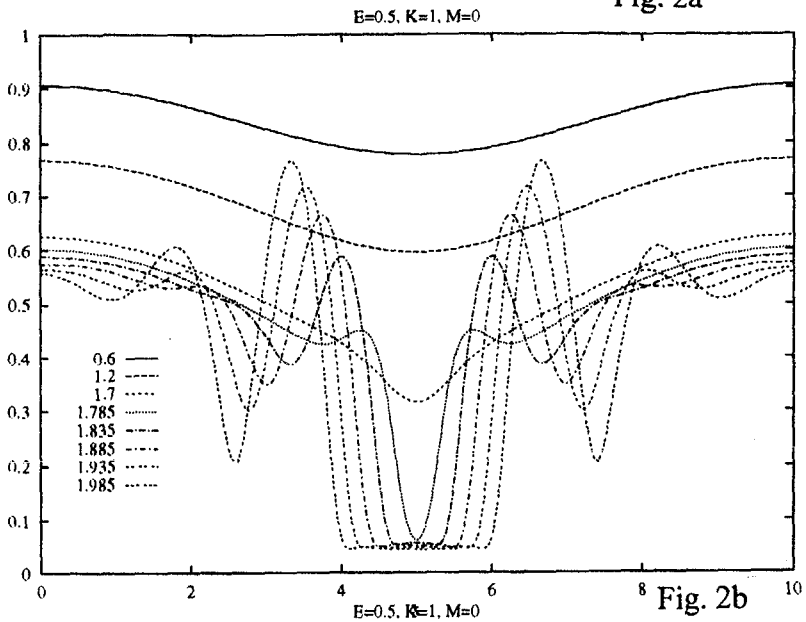


Fig. 2b

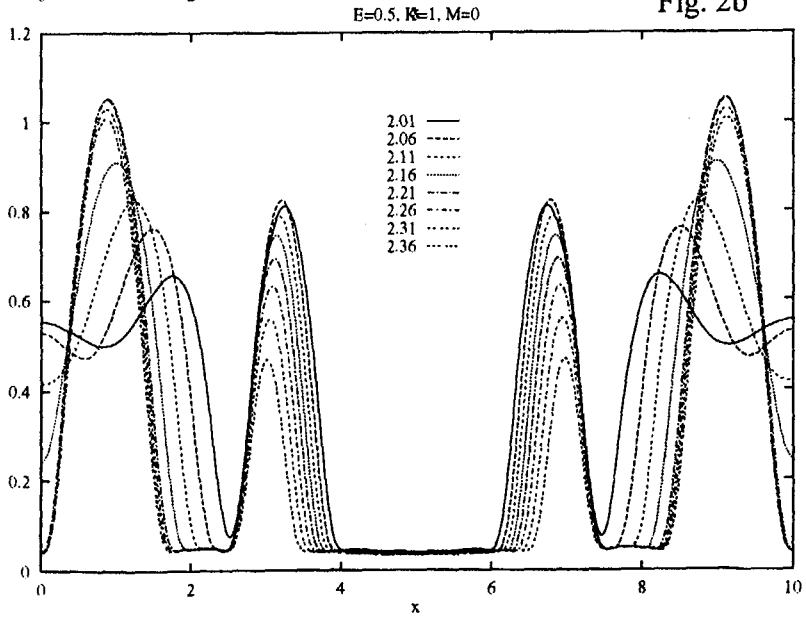


Fig. 2c

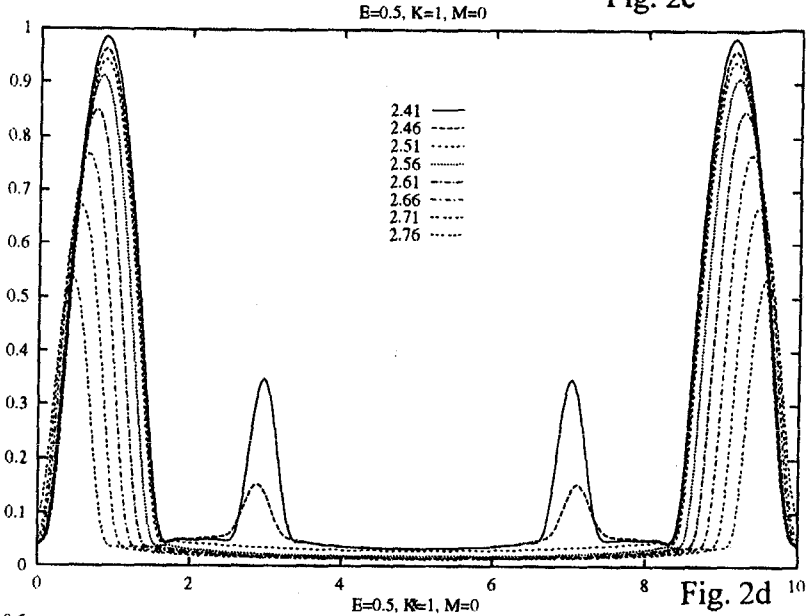
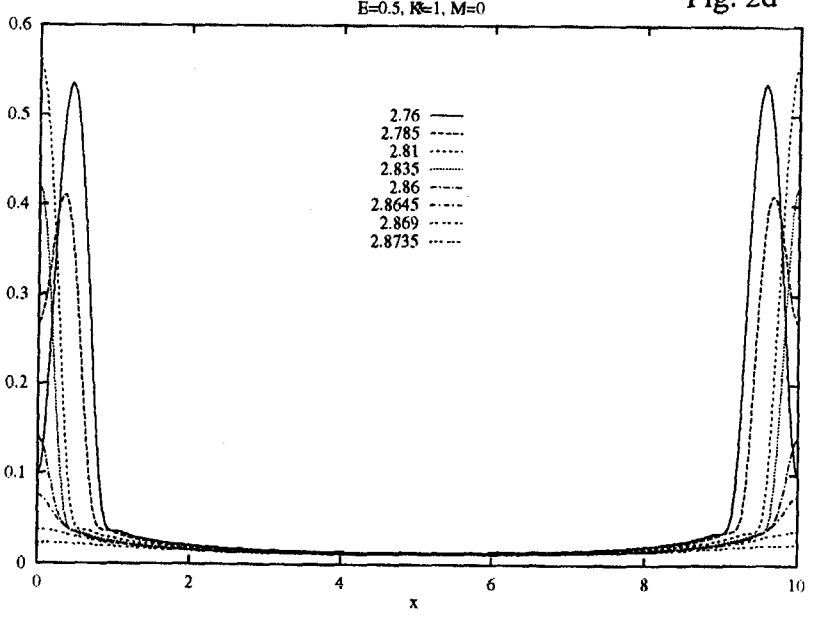


Fig. 2d



DESIGN OF A VAPOR-LIQUID-EQUILIBRIUM, SURFACE TENSION, AND DENSITY APPARATUS[1]

C.D. Holcomb and S.L. Outcalt

National Institute of Standards and Technology
Physical and Chemical Properties Division
325 Broadway
Boulder, Colorado 80303

ABSTRACT

The design and performance of a unique vapor-liquid equilibrium (VLE) apparatus with density and surface tension capabilities is presented. The apparatus operates at temperatures ranging from 218 to 423 K, at pressures to 17 MPa, at densities to 1100 kg/m^3, and at surface tensions ranging from 0.1 to 75 mN/m. Temperatures are measured with a precision of ±0.02 K, pressures with a precision of ±0.1% of full scale, densities with a precision of ±0.5 kg/m^3, surface tensions with a precision of ±0.2 mN/m, and compositions with a precision of ±0.005 mole fraction.

The apparatus is designed to be both accurate and versatile. Capabilities include: 1) the ability to operate the apparatus as a bubble point pressure or an isothermal pressure-volume-temperature (PVT) apparatus, 2) the ability to measure densities and surface tensions of the coexisting phases, and 3) the ability for either trapped or capillary sampling. We can validate our VLE and density data by measuring PVT or bubble point pressures in the apparatus. The use of the apparatus for measurements of VLE, densities, and surface tensions over wide ranges of temperature and pressure is important in equation of state and transport property model development. The use of different sampling procedures allows measurement of a wider variety of fluid mixtures. VLE measurements on the alternative refrigerant system R32/134a are presented and compared to literature results to verify the performance of the apparatus.

INTRODUCTION

Experimental properties of fluids in the two-phase region are very valuable in the characterization of the thermodynamic surface of fluids. The search for alternative refrigerants is

[1] Contribution of the National Institute of Standards and Technology, not subject to copyright in the U. S.

one of many research areas where high quality data on the vapor-liquid equilibria and densities of the coexisting phases are needed. These properties are essential in the design and optimization of many industrial processes including air-conditioning and refrigeration systems. Vapor-liquid equilibrium and coexisting density measurements are essential for developing models for the saturation boundaries of fluids. Surface tension measurements are important in calculating heat transfer coefficients for boiling and condensation. In an effort to increase our capabilities for making accurate vapor-liquid-equilibrium measurements over a large temperature range and to include the ability to measure coexisting densities and surface tensions, a new apparatus has been designed and constructed which covers the temperature range of 218 to 423 K at pressures up to 17 MPa.

APPARATUS DESIGN

Physical Description

The apparatus is designed for density and surface tension measurements in addition to the vapor-liquid-equilibrium measurements. The primary components of the apparatus are an equilibrium cell, two recirculation pumps, three platinum resistance thermometers (PRTs), two quartz crystal pressure transducers, two differential pressure transducers, a gas chromatograph (GC), two vibrating tube densimeters, a recirculating refrigerator, two heating systems, a temperature controlled enclosure, and a data acquisition system. Figure 1 shows a schematic of the apparatus. In order to obtain accurate measurements, the equilibrium cell, densimeters, and recirculating pumps must be maintained at the desired temperature and the gradients must be minimized between these components. A well-stirred liquid bath is used to achieve temperature control and minimize temperature gradients. A large glass Dewar is used to contain the silicone-oil bath and insulate it from the room. The five components are immersed in the oil. Other items located in the Dewar include the heating/cooling coils, a stirrer with three impellers, and three PRTs. The PRT located closest to the equilibrium cell is a standard-reference-quality four-wire PRT while the other two (located near the densimeters) are industrial quality four-wire PRTs. These three PRTs allow us to monitor the temperature gradients within the bath and adjust the stirrer speed to minimize the gradients.

A length of stainless steel tubing connects the system to a temperature controlled enclosure which houses the differential pressure transducers and valving for filling and evacuating the system. Located outside of the enclosure are the filling system (a liquid piston pump and a manual screw pump) and the pressure transducers. The enclosure is maintained at a temperature 10 K above the main apparatus temperature to ensure that only vapor is present in the enclosure. Liquid condensation in this part of the apparatus would adversely affect the pressure and surface tension measurements.

The vapor and liquid samples needed to determine the compositions are obtained using capillary sampling. Capillary lines are connected to the top and bottom of the main cell. The samples are withdrawn from the system and flow through two gas chromatograph injection valves. The valves inject small amounts of the samples into a helium carrier gas stream. This carrier stream flows directly into a gas chromatograph where the samples are analyzed.

Temperature Control

The uncertainties of the properties measured in this apparatus are close to state-of-the-art, especially considering the wide range of properties simultaneously measured on the same sample and the wide ranges of temperatures and pressures covered. The pressures are measured with a

precision of ±0.1% of full scale, the temperatures with a precision of ±0.02 K, the compositions with a precision of ±0.005 mole fraction, the densities with a precision of ±0.5 kg/m^3, and the surface tensions with a precision of ±0.2 mN/m. Accurate measurements are attainable because possible sources of systematic errors like large temperature gradients or a low degree of temperature stability are eliminated.

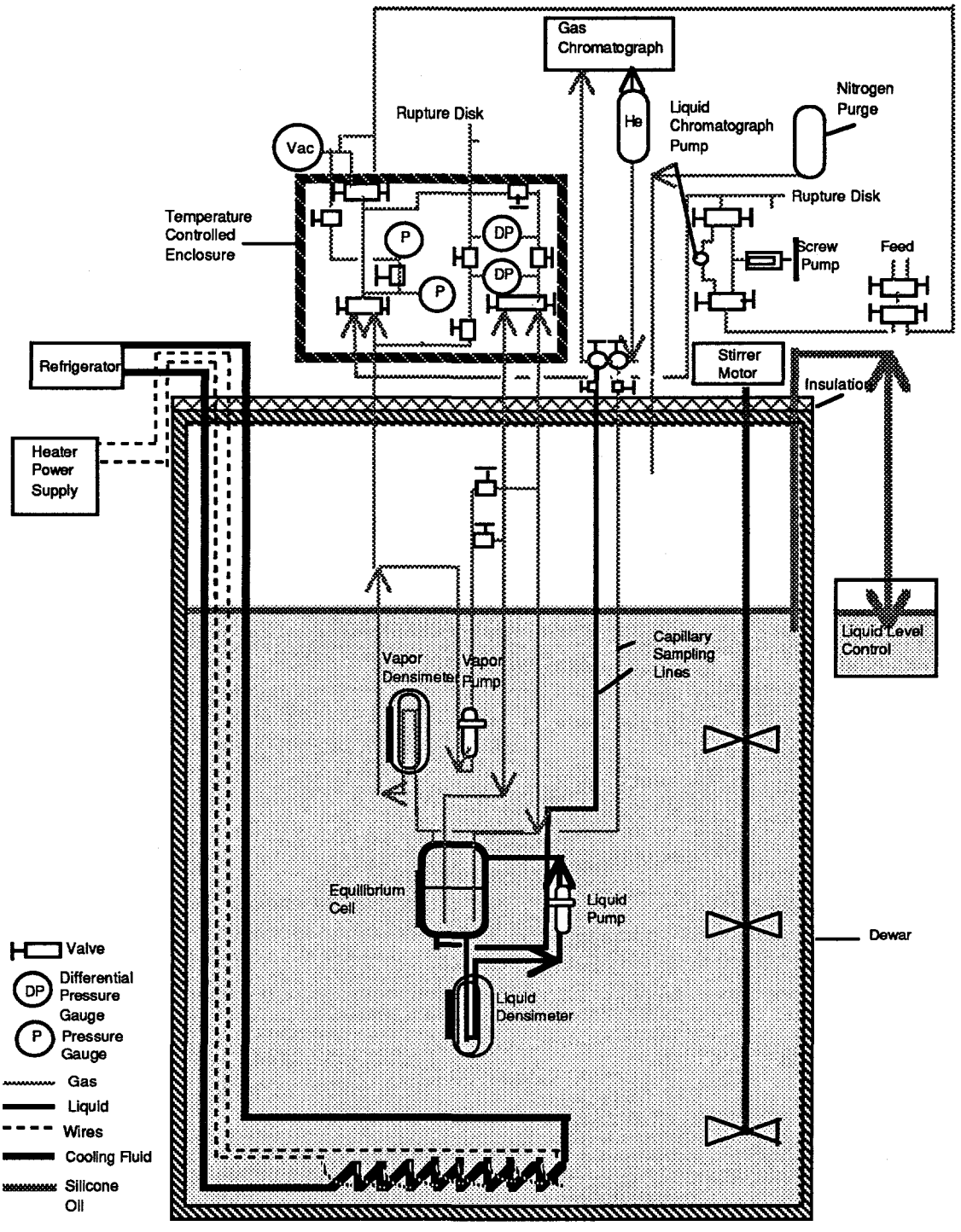


Figure 1. Schematic Diagram of the Coexisting Density and Vapor-Liquid-Equilibrium Apparatus with Surface Tension Measurement Capabilities

In designing this liquid-bath apparatus, particular care was taken to minimize thermal gradients within the Dewar. Silicone oil was selected as the bath fluid because of its excellent heat transfer properties and because it is one of the few commercially available heat transfer fluids that has a low viscosity and low vapor pressure over the entire temperature range of the apparatus. It is also non-flammable, non-toxic and relatively inexpensive.

Temperature control of the silicone oil bath is achieved with two heating systems and a recirculating refrigerator. The fine heating system is a low-power, computer-controlled system used for fine temperature control. It is capable of maintaining the bath temperature within ± 0.02 K of the set-point temperature for several hours. The coarse heating system consists of a high-power system with a programmable controller. It is the major source of heat when increasing the temperature of the system.

A large recirculating refrigerator cools the silicone oil bath. The set-point temperature of the recirculating fluid exiting the refrigerator is ten degrees colder than the set-point temperature of the apparatus. This temperature difference allows the refrigerator to remove heat from the system without introducing a cold spot. A metering valve is used to adjust the flow rate through the cooling coils. The valve is fully open when cooling the bath to a new set-point temperature and almost closed when providing the necessary cooling to maintain a given set-point temperature.

Vapor-Liquid Equilibrium Measurements

The apparatus is evacuated before initiating measurements on a new mixture. After the system reaches the desired temperature, the fluids of interest are loaded into the system. The components are added to the system using either the screw pump or the piston pump. Then, the recirculation pumps are turned on and the liquid and vapor phases are recirculated to ensure thorough mixing. The vapor is returned to the bottom of the cell and the liquid to the top. Next, the pumps are turned off and the pressure and temperature of the system are monitored. Once equilibrium is achieved, the pressure and temperature are recorded and the compositions of the two phases are determined. To advance to the next point on the isotherm, an additional amount of one component is added to change the composition. The phases are mixed and equilibrium is attained before the next set of measurements is obtained.

Sampling Systems

The compositions of the coexisting phases are determined by withdrawing vapor and liquid samples and analyzing them with a GC. The two basic methods for withdrawing the samples are known as capillary sampling and trapped sampling. A review and comparison of these two methods is reported by Noles[1]. The apparatus can be configured to use either type of sampling system, but is currently arranged for capillary sampling. Capillary sampling is used with systems that are sensitive to temperature gradients. These samples must be withdrawn directly from the equilibrium cell. This method relies on plug flow in the capillary lines to allow the mixture to flash to an all-vapor phase without stratification. A small quantity of the flashed vapor is injected into a helium carrier stream using a rotary GC valve. The sample is carried to the GC for analysis.

Trapped sampling is used when studying mixtures that are toxic, are available only in small quantities, or have a bubble-point pressure less than 0.3 MPa. Trapped sampling for this system involves mounting high pressure GC valves inside the lid of the dewar. The valves are piped into the recirculation loops, and the quantity of material is trapped in the rotor and then injected into a helium carrier stream. The sample is carried to the GC for analysis.

Density and Surface Tension Measurements

Coexisting densities and surface tension measurements can also be made with this apparatus. Two commercially available vibrating tube densimeters are used to measure the liquid and vapor densities. The densimeters are mounted so that they are self-draining, one densimeter is located above and one below the equilibrium cell. They are located in the recirculation lines to ensure that they are entirely filled with the equilibrium liquid or vapor. The temperature differences for the densimeters, pumps, and the equilibrium cell are within ±0.05 K. This ensures that the densities measured are representative of the bulk liquid and vapor.

The surface tension is measured using the differential bubble pressure method[2]. The vapor recirculation pump is used to supply vapor to two dip tubes mounted in the equilibrium cell. The equilibrium vapor is bubbled into the liquid through one dip tube at a time, and the maximum pressure that the bubble reaches during the bubble growth is measured with a differential pressure transducer. The surface tension is a function of the difference in the two maximum pressures, the radii of the two dip tubes, and the density difference between the vapor and liquid phases.

Apparatus Versatility

The system can be operated as an isothermal single-phase PVT or a two-phase bubble-point pressure apparatus as well as a vapor-liquid-equilibrium apparatus. The system can be filled completely with a gravimetrically prepared standard mixture and the densimeters used to measure the density as a function of pressure at a given temperature for that composition. This provides isothermal PVT measurements which can be used to supplement or validate existing PVT data.

The system can be nearly filled with a standard mixture and the bubble-point pressure can be measured as a function of temperature. The bubble-point pressure measurements eliminate the uncertainty introduced by sampling and analysis to determine the liquid composition. These measurements represent an independent verification of the bubble-point measurements from the vapor-liquid equilibrium measurements, and can be used to check the validity of the GC calibration.

PERFORMANCE VERIFICATION

Alternative refrigerant mixtures are excellent candidates for testing the performance of a vapor-liquid-equilibrium apparatus. These systems have small heats of vaporization and are especially sensitive to temperature gradients. Comparisons of the data measured for the alternative refrigerant mixture R32/134a with independent measurements demonstrate the ability of the apparatus for accurate vapor-liquid-equilibrium measurements.

The data for the R32/134a system cover 5 isotherms from 279 to 340 K. A total of 48 vapor-liquid equilibrium measurements are presented in Table 1. The data are compared to four other data sets available in the literature. We compare the data sets with the predictions of the Lemmon-Jacobsen model [3] in REFPROP 6.0 [4]. The deviations between the data and the values predicted from the Lemmon-Jacobsen model are presented in Figures 2 and 3. In Figure 2 the deviations between the measured bubble-point pressures and the predicted values are shown.

The scatter in the data might appear large compared to the experimental uncertainties quoted for the apparatus. However, a distinction must be made between the experimental uncertainty of a specific parameter measurement (such as temperature, pressure, or composition)

and the state-point uncertainty. The state point is system dependent and is defined as the equilibrium bubble-point pressure and vapor composition at a given liquid composition and temperature. Each of the parameters (P,T,x,y) has an experimental uncertainty associated with the measurement of that parameter. However, the state-point uncertainty is a function of the uncertainties of each of these four parameters as well as the dependencies between them. The state-point uncertainties are different for each system and vary with temperature and composition. The bubble-point state point uncertainty is calculated by multiplying the uncertainty in the composition measurement by dP/dx at that state point, adding the uncertainty of the pressure measurement, and adding the temperature measurement uncertainty multiplied by dP/dT at that state point. The bubble-point pressure state-point uncertainty for the R32/134a system at temperatures from 280 to 340 K ranges from ±0.22 to ±0.30% (±2σ)

Table 1

Vapor-Liquid Equilibrium Data for R32/134a Mixtures from 280 to 340 K [5].

x(R32) Liquid Mole Fraction	Temperature K	Pressure MPa	y(R32) Vapor Mole Fraction	x(R32) Liquid Mole Fraction	Temperature K	Pressure MPa	y(R32) Vapor Mole Fraction
0.000	280.39	0.379	0.000	0.619	299.96	1.357	0.761
0.299	279.98	0.555	0.489	0.262	303.87	1.076	0.403
0.478	280.05	0.674	0.676	0.609	303.07	1.453	0.756
0.604	279.98	0.739	0.757	0.000	309.31	0.913	0.000
0.642	279.98	0.772	0.797	0.181	309.98	1.167	0.292
0.642	280.09	0.774	0.793	0.225	310.00	1.225	0.370
0.650	279.99	0.778	0.805	0.228	309.97	1.215	0.363
0.650	280.19	0.778	0.796	0.360	309.98	1.392	0.515
0.737	280.13	0.852	0.869	0.681	309.99	1.826	0.793
1.000	280.00	1.006	1.000	1.000	309.99	2.294	1.000
0.302	283.02	0.609	0.483	0.000	324.50	1.367	0.000
0.632	283.06	0.840	0.782	0.162	325.00	1.643	0.256
0.640	283.01	0.846	0.795	0.294	324.99	1.889	0.422
0.640	283.08	0.847	0.795	0.595	325.00	2.423	0.710
0.283	289.91	0.746	0.466	0.733	324.99	2.698	0.809
0.637	289.79	1.030	0.782	0.783	324.99	2.805	0.860
0.282	292.93	0.811	0.457	1.000	324.99	3.279	1.000
0.634	292.92	1.124	0.773	0.000	340.16	1.986	0.000
0.000	294.69	0.596	0.000	0.466	339.87	3.036	0.569
0.275	294.95	0.856	0.448	0.577	339.93	3.315	0.669
0.452	294.99	1.018	0.626	0.763	339.99	3.807	0.820
0.529	295.13	1.091	0.696	1.000	339.98	4.561	1.000
0.584	294.94	1.149	0.748				
0.632	294.96	1.187	0.773				
0.720	294.94	1.271	0.838				
1.000	295.48	1.575	1.000				

The deviations in figure 2 compare the data to the predictions of the model. There is a slight composition dependent systematic deviation between the model and the data that appears as a spread in the deviations at a given temperature. The data taken with this apparatus agree with

Nagel and Bier [5] within ±1.0 %, and most of the data of Higashi [6] and of Widiatmo et al. [7] within ±2 % except for a few points. The data of Fujiwara et al. [8] at 273 K are consistently lower than any of the other data sets.

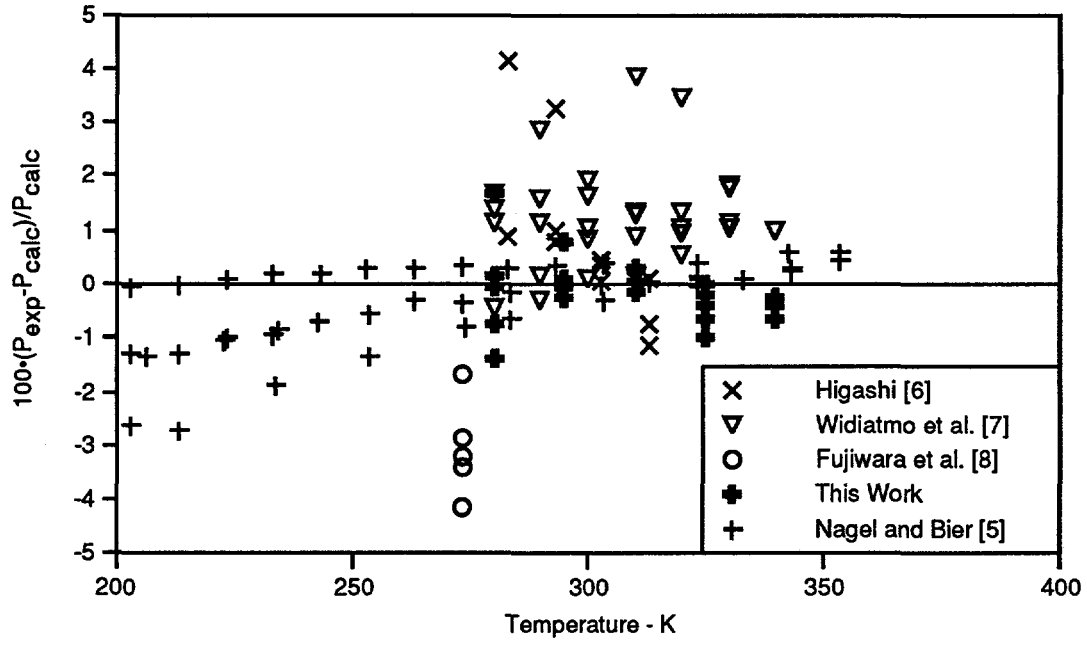


Figure 2. Comparison of Bubble-Point Pressures for the R32/134a System; the baseline is from the Lemmon-Jacobsen Model in REFPROP 6.0.

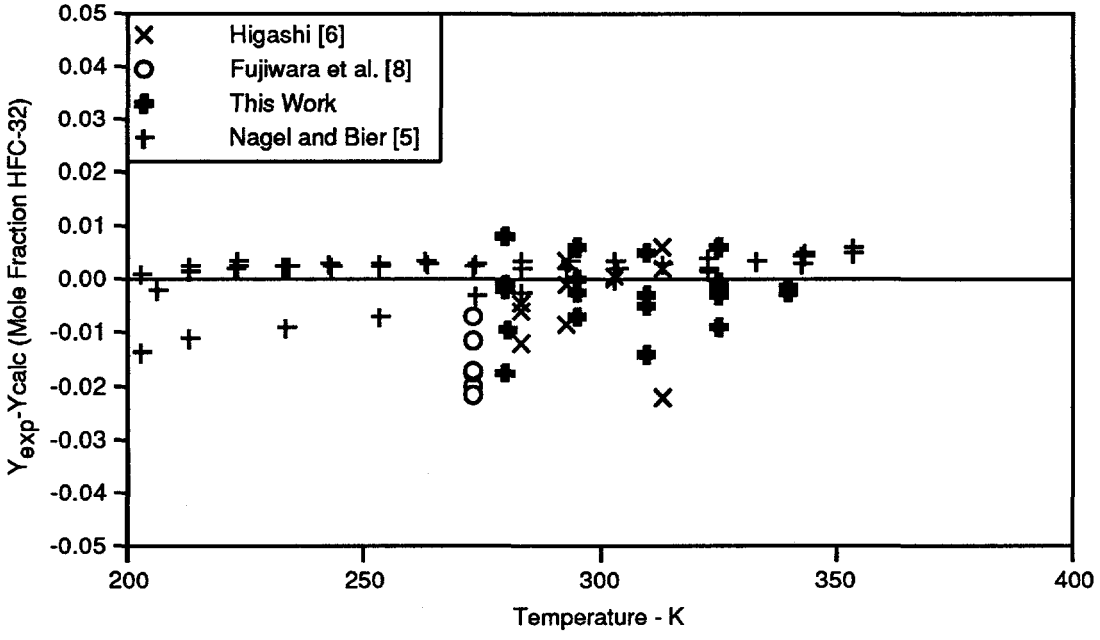


Figure 3. Comparison of Vapor Compositions for the R32/134a System; the baseline is from the Lemmon-Jacobsen Model in REFPROP 6.0.

In Figure 3 the vapor compositions of the data sets are compared. In general, the four data sets agree within ±0.015 mole or mass fraction HFC-32. The data of Fujiwara et al. [8] at 273 K are systematically 0.015 mole fraction HFC-32 lower than the data of Nagel and Bier [5].

The vapor-composition state-point uncertainty ranges from ±0.006 to ±0.013 mole fraction HFC-32 for the R32/134a system at temperatures from 280 to 340 K.

R32/134a is one of the many refrigerant mixtures that have been studied with this apparatus. Some of the systems included in the study were R32/125, R125/134a, R32/125/134a, R32/143a, R143a/125, R143a/134a, R32/290, R125/290, R134a/290, and R41/744 [9]. The apparatus is currently being used to study mixtures containing CF_3I and hydrocarbon mixtures.

ACKNOWLEDGEMENT

The financial assistance of the United States Department of Energy is gratefully acknowledged. The apparatus was constructed under a grant from the Division of Engineering and Geosciences in the Office of Basic Energy Sciences and the measurements on R32/134a mixtures were funded under a grant from the Building Equipment Division in the Office of Building Technologies.

REFERENCES

1. J. R. NOLES, "Vapor-Liquid Equilibria of Solvating Binary Mixtures," Ph. D. Thesis, Department of Chemical Engineering, Cornell University, Ithaca, NY. (1991).

2. C. D. HOLCOMB and J. A. ZOLLWEG, "An Improved Differential Bubble Pressure Surface Tensiometer," *Journal of Colloid and Interface Science* 154, 51 (1992).

3. E. W. LEMMON, "A Generalized Model for the Prediction of the Thermodynamic Properties of Mixtures Including Vapor-Liquid Equilibrium," Ph. D. Thesis, Department of Mechanical Engineering, University of Idaho, Moscow, ID. (1996).

4. M. O. MCLINDEN and S. A. KLEIN, "A Next Generation of Refrigerant properties Database," Proceedings of the 6th International Refrigeration Conference at Purdue, West Lafayette, IN, July 23-26, 409 (1996).

5. M. NAGEL and K. BIER, "Vapour-Liquid Equilibrium of Ternary Mixtures of the Refrigerants R32, R125 and R134a," *International Journal of Refrigeration* 18, 534 (1995).

6. Y. HIGASHI,"Vapor-Liquid Equilibrium, Coexistence Curve, and Critical Locus for Binary HFC-32/HFC-134a Mixture," *International Journal of Thermophysics* 16 ,1175 (1995).

7. J. V. WIDIATMO, H. SATO, and K. WATANABE, "Measurements of the Liquid Densities of the Binary HFC-32 + HFC-134a System," *Fluid Phase Equilibria* 99, 199 (1994).

8. K. FUJIWARA, H. MOMOTA, and M. NOGUCHI, "Vapor-Liquid Equilibria of HFC-32 Mixtures," 13th Japanese Symposium Thermophysical Properties A116, 61 (1992).

9. W. M. HAYNES, "Thermophysical Properties of HCFC Alternatives", Final Report, Air-Conditioning and Refrigeration Technology Institute", November 1996.

INTERFACIAL WAVE BEHAVIOR IN OIL-WATER CHANNEL FLOWS: PROSPECTS FOR A GENERAL UNDERSTANDING

M. J. McCready, D. D. Uphold, K. A. Gifford
Department of Chemical Engineering
University of Notre Dame
Notre Dame, Indiana 46556 USA

ABSTRACT

Oil-water pressure driven channel flow is examined as a model for general two-layer flows where interfacial disturbances are important. The goal is to develop sufficient understanding of this system so that the utility and limitations of linear and nonlinear theories can be known *a priori*. Experiments show that sometimes linear stabilty is useful at predicting the steady or dominant evolving waves. However in other situations there is no agreement between the linearly fastest growing wave and the spectral peak. An interesting preliminary result is that the bifurcation to interfacial waves is supercritical for all conditions that were studied for an oil-water channel flow, gas-liquid channel flow and two-liquid Couette flow. However, three different mechanisms are dominant for each of these three situations.

INTRODUCTION

All facets of energy production and distribution involve processes that contain multi-fluid flows. Hydrocarbon production and transportation, energy exchange devices such as condensers are common examples. Further, there is extensive use of multiphase reactors for chemical production in which hydrodynamic instabilities can alter the chemical product distribution.

Unfortunately, even for the simplest class of two-fluid flows, stratified, it is not possible to predict the nature of the instabilities that occur and the eventual downstream behavior. While it is in principle possible to perform a linear stability analysis of the base flow (the complication being the possible presence of turbulence), linear theory tells information only about the initial instability. This is sufficient to predict if any disturbances will grow and their initial nature. If the waves grow and then saturate at small amplitude without changing wavelength, then only the saturation process involves nonlinear effects. However, often, the initial waves form subharmonics[1,2], overtones [1] or even interact with wavelengths much longer than the fundamental [1].

Because of this (for general flows), the next level of theory, nonlinear analysis, cannot be done without looking at what happens in the experimental system. For example, nonlinear analysis is often based on the long wave assumptions [3,4]. For falling films the waves often evolve to long wavelengths, (even if the initial instability is short) so that this approach works. However, for two-layer systems where gravity is stabilizing, the positive linear growth rate region is often bounded away from 0 wavenumber [5,6]. Further, waves may remain of moderate wavenumber even though the linear growth rate is positive down to 0 wavenumber[7]. Thus long wave theory is not of general applicability for two-layer systems.

Blennerhassett [8] Renardy and Renardy [9] have formulated the interfacial problem as a weakly nonlinear expansion and derived equations for the amplitude of the one or two dominant modes. Sangalli et al. [5] showed that for a two-layer Couette flow that experiments agree quite well with theory. Thus the eigenfunction expansion and center manifold projection approach

works for this system. However, a single Stuart Landau equation [8,9] will be valid only for situations when a single short-wave mode is dominant. Formation of subharmonic or other modes cannot be predicted and this single equation does not tell when a different one is required. (At least when the bifurcation is supercritical which it is for all interfacial systems that we have done calculations for.) Thus at the present time the weakly nonlinear problem is not well understood.

A serious need exists for a well defined system to study to sort out as many of this issues as possible. The gas-liquid channel flow does not conveniently exhibit a useful range of behavior for two-layer laminar flows. We have found that an oil-water system has several features that make it appropriate for study. First is a significant range of laminar flow. Second is that there are two distinct modes that lead to interfacial waves and which travel at different speed. Third, these two modes can be simultaneously neutrally stable. Further there is another condition where the long wave region and the short wave region of the same mode can be simultaneously neutrally stable. Thus we can study conditions where two distinct wave modes can be generated and then interact and other conditions where the same mode, but at different wavelengths, interact. Therefor we have an experimental system that allows a broad range of different nonlinear interactions.

In this paper, preliminary results from this study are given. We have found the linear stability theory predicts the onset well. Further, both waves lead to observable interfacial disturbances. (which is not the case for gas-liquid flows). Nonlinear theory suggests that even though the bifurcation seems to always be supercritical, the mechanism responsible for stabilization is different in different ranges.

EXPERIMENTAL SYSTEM

Figure 1 shows a schematic of the oil-water channel that is used for the experiments. Data are obtained from visual and video observations and from conductance probes. We are currently working on an optical technique for measuring the interface tracings to replace the conductance probes. The fluids are water, with Sodium Silicate added to improve its ability to wet the Plexiglas® channel and a light hydrocarbon oil with a density of 0.88 g/cm^3 and a viscosity of 17.8 cP. More details about the flow system and its construction are included in a thesis by McKee[10].

THEORY

Theoretical analysis for this system includes (temporal) linear stability analysis for a two-layer laminar flow that has been completely formulated in papers by Yih[11]or Blennerhassett [8]. The problem is solved numerically with Chebyshev-Tau spectral technique [12] using a scheme devised by Gardner et al.[13]. It is difficult to get accurate results in the long wave region so the results are compared to the analytical long wave solution of Blennerhassett [8]. Strictly speaking, the channel flow geometry is a convective situation and should require a spatial stability formulation. However, this is much more complicated to program and lengthy to compute and we have found that satisfactory results can be obtained from the temporal stability problem with a Gaster transformation if necessary.

Nonlinear analysis for the two-layer problem has been formulated with a multiple scales technique by Blennerhassett [8] and an eigenfunction, center manifold approach by Renardy and Renardy [9]. We use the basic approach of Renardy and Renardy [9] except that the individual contributions to the coefficient β, in the Stuart-Landau equation

$$\frac{\partial A}{\partial t} = L(\lambda)\, A + \beta\, |A|^2 A \qquad [1]$$

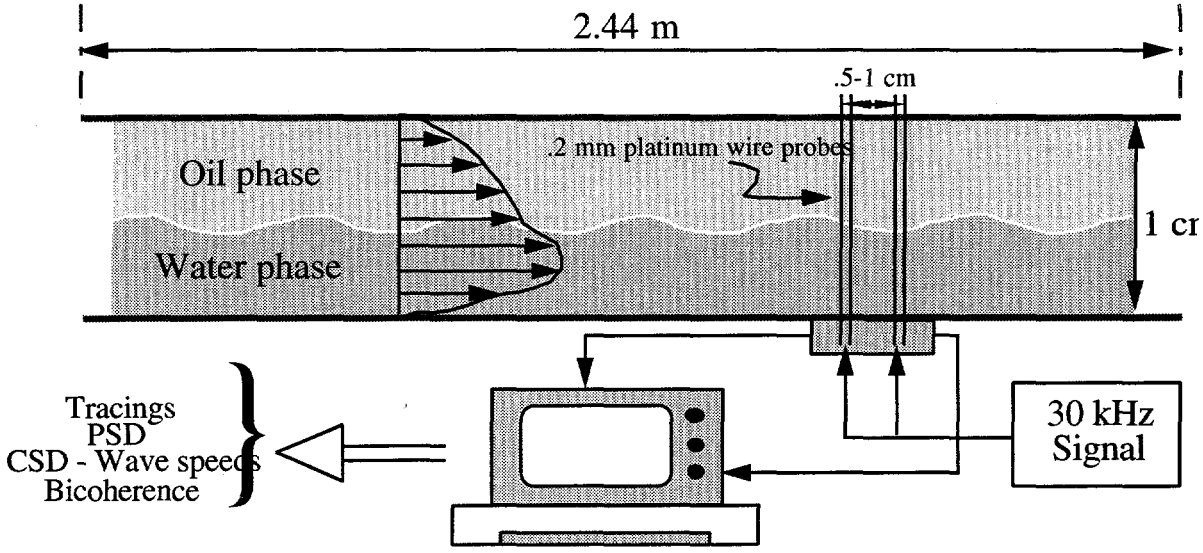


FIGURE 1. OIL-WATER CHANNEL FOR STUDYING INTERFACIAL WAVES

are separated into a quadratic contribution, a cubic self interaction and a cubic interaction with the mean flow (see [6]) for details. This equation tells the nature (i.e., super or subcritical) of the initial bifurcation from a steady base state if the neutral wave number occurs away from 0 wavenumber. If the flow is unstable to long waves, then the neutral wavenumber is 0 and this analysis is not valid. A further difficulty is that even when the initial bifurcation is supercritical, at sufficient forcing, the region of unstable modes and the growth rates get large enough for other modes to occur -- even though the amplitude/wavelength ratio of any mode remains small.

RESULTS

The linear stability diagram should predict the initial transition behavior. Figure 2 shows a plot in Re_{oil} versus Re_{water} of the long wave and short wave stability boundaries. At two points, the different wave types are simultaneously neutrally stable. Thus there are regions where two distinct modes are simultaneously linearly unstable and (by change of the inclination angle of the flow) can have a variable growth rate ratio. It is important to examine the nature of these modes. Figure 3 shows the growth and speed plots for conditions close to the crossing point at low Re_o. The three most unstable modes are plotted. Because an individual mode would be expected to vary smoothly with wavenumber (particularly on the speed plot) it is clear that the same mode is not most unstable always. The mode that is unstable at long wavenumber is distinct from the mode that is unstable at intermediate wavenumbers.

The disturbance eigen function is plotted in figure 4 as a contour plot (for one complete wave) to show what the different modes look like. It is seen that the low k disturbance has its water part centered in the water (lower) phase. However, there is a second unconnected (weaker) part that starts in the oil but is centered at the interface. The high k disturbance has the oil part of the

disturbance centered at the interface connected with a skewed part in the water phase. As the water phase travels much faster than the oil, this could be the reason that the high k disturbance has a

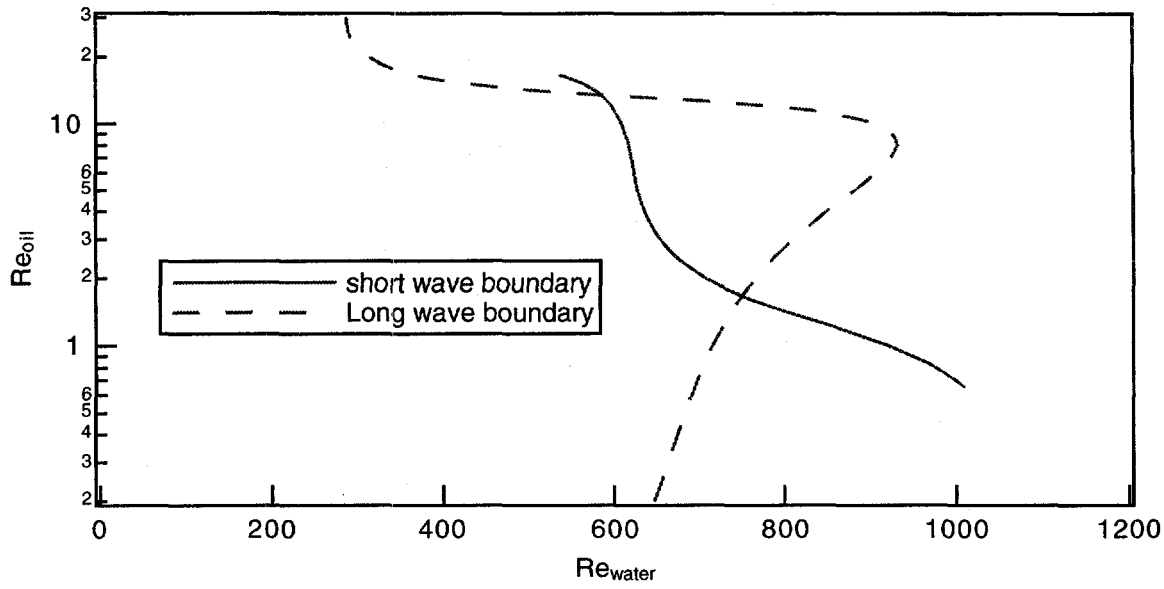


Figure 2. Plot of neutral stability boundaries for long and short waves.

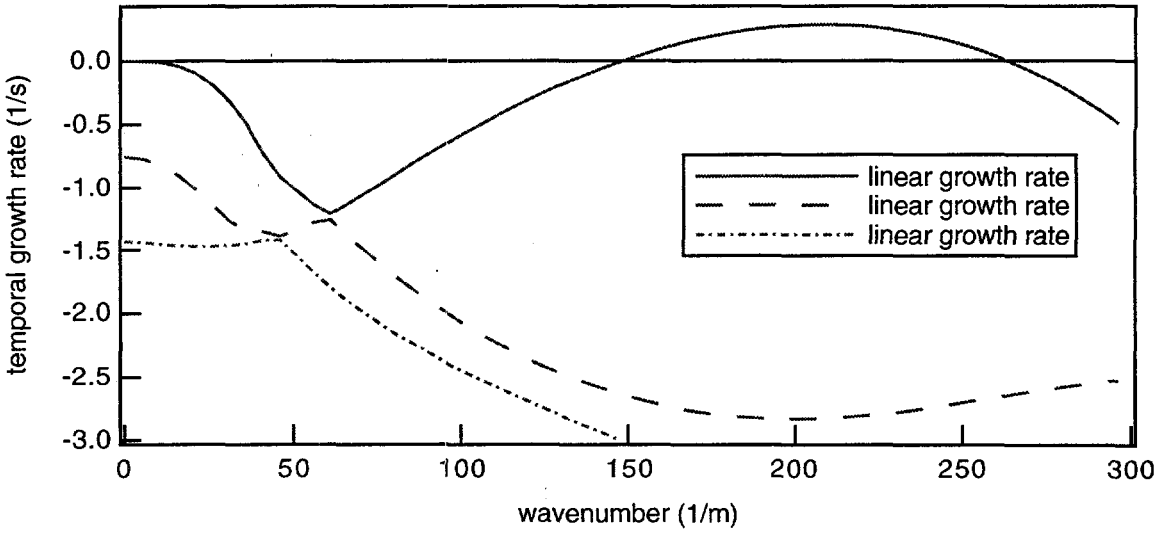


Figure 3a. Plot of growth curves for first three modes, $Re_o = 1.8$, $Re_w = 770$

much higher phase velocity. It should be noted that neither of these disturbances has the "pure" form of interfacial and internal modes as described by Yiantsios and Higgins[14]. Experiments show that both of them lead to interfacial waves. This is in contrast to situations for say, gas-liquid flows where a gas phase internal mode can be unstable, and perhaps be associated with turbulence, but not cause interface waves.

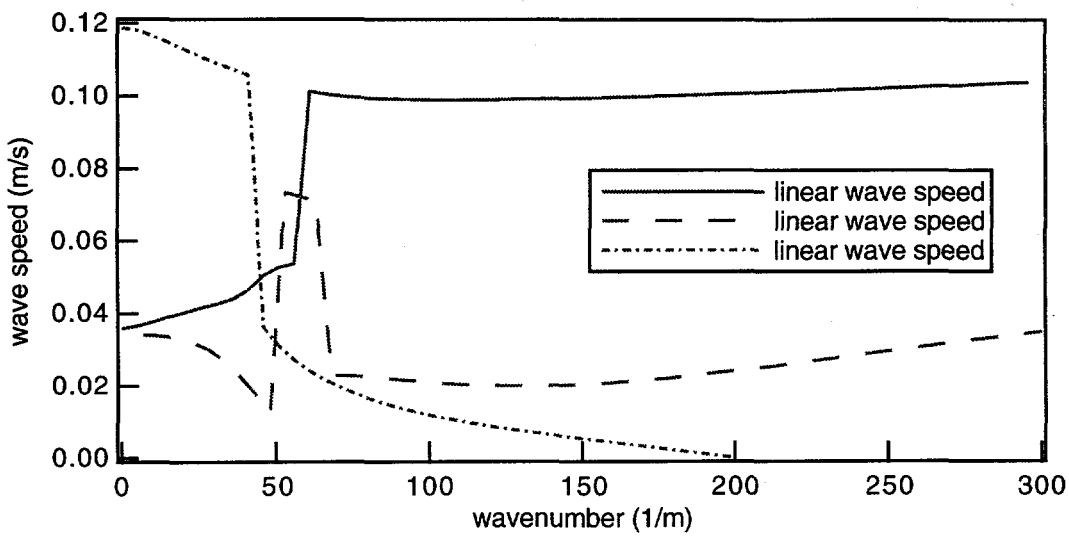


Figure 3b. Plot of wave speeds for first three modes, $Re_o = 1.8$, $Re_w = 770$

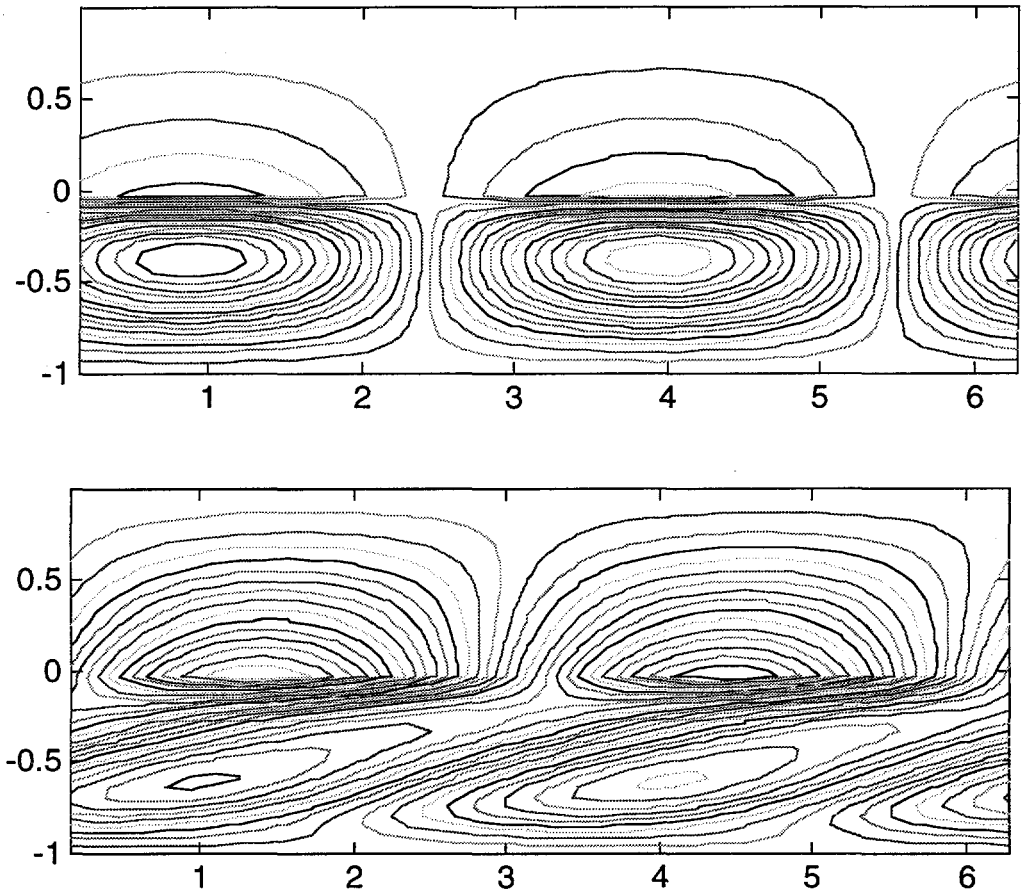


Figure 4. Disturbance eigen function at $Re_o = 1.8$, $Re_w = 770$. a. top: k=1/m
b. bottom: k=211/m.

The situation is different at the long-short crossing point at higher Re_o. Plots of the disturbance function would show that both the long and short waves are the same qualitative shape, (similar to figure 3b) and that the speeds of the long and short waves are comparable.

The importance of developing an better understanding of the nonlinear processes that affect wave behavior is shown by reference to figure 5. In figure 5a, the linear growth curve has a peak close to 3 Hz and the wave data have a corresponding spectral peak. Thus the wavelength and speed of steady waves are close to the linear stability values. However, if the water Reynolds number is increased the situation is very different. Now the spectrum of figure 5b has no relation to the peak in the linear growth region but occurs at much smaller frequency. While the low frequency region is linearly unstable, there is no indication that the low mode will dominate completely with almost no evidence of short waves. Further, video images of this condition show no evidence that the short wave instability is reforming even though the growth rate is large enough that short waves would be expected to reform between the large waves.

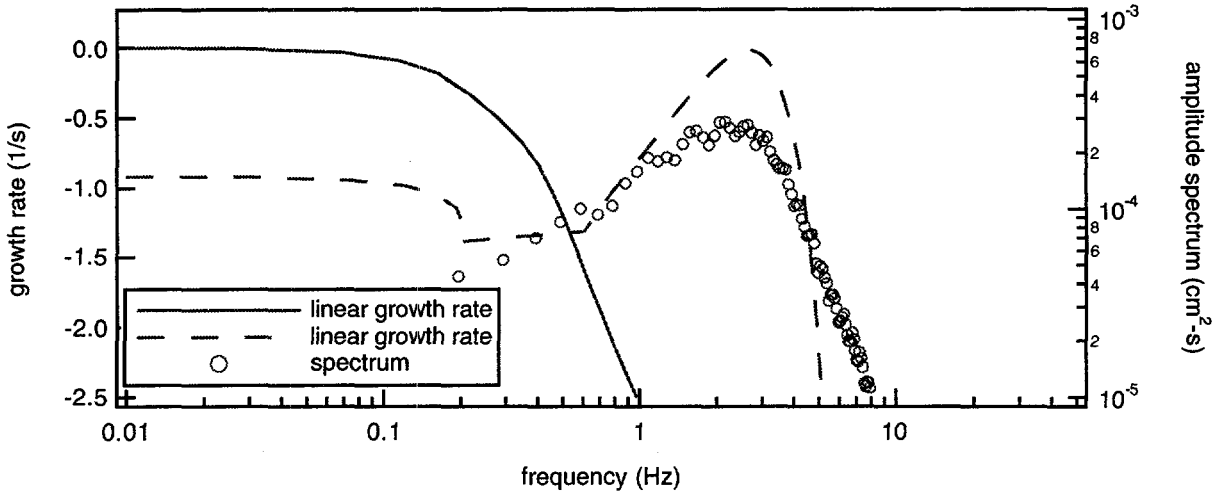


Figure 5a. Growth rate of two modes and measured spectrum at Re_o=3, Re_w = 700

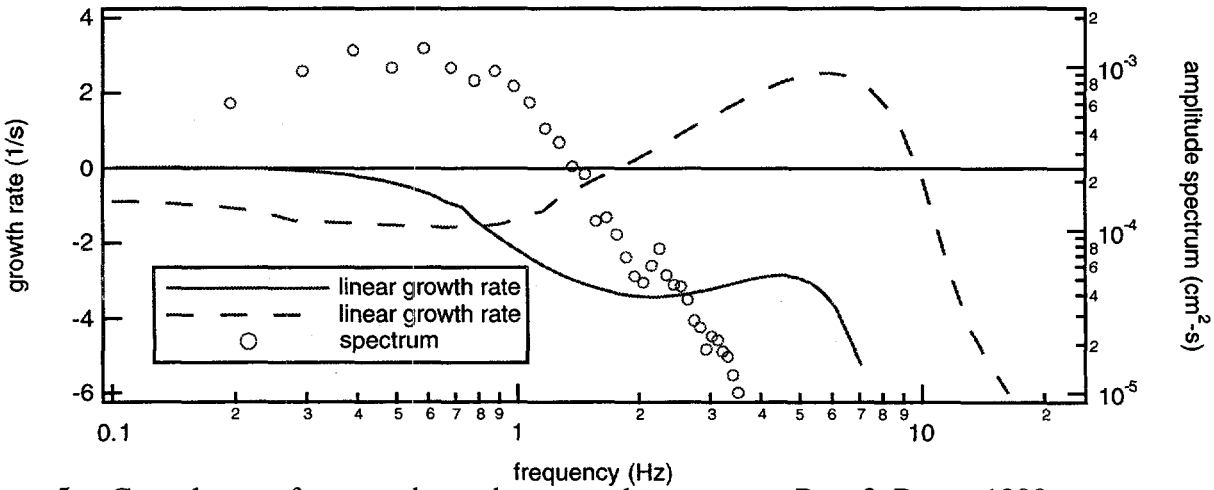


Figure 5a. Growth rate of two modes and measured spectrum at Re_o=3, Re_w = 1200

DISCUSSION

While this project is just beginning, there are several issues that have emerged which were not expected based on numerous previous studies of gas-liquid flows.

First the relative interaction rates of different wavelength modes with different speeds should be successfully determined by experiments in the vicinity of the two-mode crossing points. Preliminary indications are that close to the low mode crossing point there is no interaction between the modes. We are not yet sure about the region around the other crossing point. In either case, the two-mode weakly nonlinear theory is expected to be adequate to describe the process close to the neutral point.

An interesting issue emerges when the oil-water channel is compared to two other two-layer systems that we have studied. Sangalli et al.[6]studied a gas-liquid cocurrent flow close to neutral stability and found that weakly-nonlinear theory predicted the qualitative behavior of the system quite well. The results of most interest here are that the bifurcation of the short wave mode was always supercritical (for all cases calculated) and that the origin of the nonlinear stabilization was the cubic terms that originate in the boundary conditions except when the speeds of the fundamental and overtone were close to the same. In this case the system was nearly-resonant and efficient transfer of energy from the unstable fundamental to the stable overtone stabilized the system at amplitudes smaller than expected for cubic stabilization. A third mechanism [6] cubic interaction between the fundamental and the base flow was found to be unimportant. In contrast, for a two-layer Couette flow [7], we have calculated results for quite a few cases and found that the bifurcation is again supercritical and the overtone interaction is dominant even if the fundamental and overtone are not close to resonant. (In most cases the cubic interaction is destabilizing and the base state interaction is not significant.) Thus it is quite interesting that the oil-water channel flow should again exhibit a supercritical bifurcation for all tested conditions but, the dominant interaction is usually the cubic interaction between the fundamental and the mean flow. For the three different experiments available to us, there are three different mechanisms of stabilization. However, in all cases, cubic order stabilization is observed and the bifurcation is supercritical.

In light of these calculations a check of the calculation procedure as applied to two-layer flows is warranted (note that we had already reproduced the single phase subcritical behavior). For a gas - liquid flow it is possible to get a gas-phase internal mode to be neutrally stable before any interfacial modes occur by raising the gas Reynolds number above 6000-7000, while keeping the liquid one low. For $Re_G = 8000$, $Re_L = 2$, $\mu_L = 5$ cP, it was found that the gas mode appeared and was, as expected, subcritical. Thus the observed supercritical nature for two-layer flows is for the interfacial mode when the instability is bounded away from 0 wavenumber.

Of course it is not known if the bifurcation is supercritical of *all* conditions. It would be interesting to determine if this is the case. A proof seems impossible given the complexity of the system although perhaps worth looking for some restricted parameter ranges. Certainly we will continue to look for subcritical regions. If can find any, experiments will be done in this range to see what differences in behavior exist. Because this issue is so fascinating and we have not been able to find a subcritical region for the fundamental, calculations have been done to find how wavenumbers other than the peak would bifurcate if they were present. Interestingly, for the gas-liquid systems, most modes at wavenumbers less than the fundamental are subcritical. A way to simulate the behavior of these modes is to excite waves artificially with a paddle. For one set of flow conditions, paddling was done over the range of 1-10 Hz. In all cases no amplification of the paddled modes were observed and sufficiently far downstream the spectrum was identical to cases where no paddling was done. This result can not be considered conclusive, however the experiments do not seem consistent with the calculations and further work is needed.

One last issue that will require more study and probably a new formulation is how to deal with conditions far above criticality. For the conditions of figure 5b, the initial bifurcation was supercritical and the wave amplitude are never real large. However, once the growth curve has a sufficiently wide region of instability, it is clear that one or two amplitude equations will not be sufficient. Perhaps the weakly-nonlinear formulation of the Navier-Stokes equations can be solved by integration in time and space to produce results that predict the experiments.

ACKNOWLEDGMENT

This work has been supported by the U. S. Department of Energy, Office of Basic Energy Sciences

REFERENCES

1. L. A. Jurman, S. E. Deutsch, S. E. and M. J. McCready, "Interfacial mode interactions in horizontal gas-liquid flows," *J. Fluid Mech.* **238**, 187-219 (1992).
2. Z. Fan, F. Lusseyran, F. and T. J. Hanratty, "Initiation of slugs in horizontal gas-liquid flows, "*AIChE J.* **39** 1741-53 (1993).
3. B. S. Tilley, S. H. Davis and S. G. Bankoff, "Long-wave nonlinear stability of superposed fluids in an inclined channel," Phys. Fluids **6**, 3906 (1994).
4. F. Charru and J. Fabre, *Phys. Fluids* , **6** , 1223 (1994).
5. M. Sangalli, C. T. Gallagher, D. T. Leighton, H. -C. Chang and M.J. McCready,"Finite amplitude wave evolution at the interface between two viscous fluids", *Phys. Rev. Let.* **75**, pp 77-80, (1995).
6. M. Sangalli, M. J. McCready and H. -C. Chang, "Stabilization mechanisms of short waves in gas-liquid flow", *Phys. Fluids*, **9,** pp 919-939, (1997).
7. C. T. Gallagher, M. J. McCready and D. T. Leighton, "Experimental investigation of a two-layer shearing instability in a cylindrical Couette cell", *Phys. Fluids*, **8**, pp 2385-2392, (1996).
8. P. J. Blennerhassett, "On the Generation of waves by wind," *Proc. R. Soc. Lond.* A **298**, 451-494 (1980).
9. M. Renardy and Y. Renardy, "Derivation of amplitude equations and analysis of sideband instabilities in two-layer flows", *Phys. Fluids*, **A 5**, 2738-2762 (1993).
10. William McKee, "An experimental study of interfacial waves in cocurrent oil-water flows" -- M.S. Thesis, University of Notre Dame, (1995).
11. C. S. Yih, "Instability due to viscosity stratification", *J. Fluid Mech.* **27**, 337-352 (1967).
12. Y. Y. Su and B. Khomami, B., "Numerical Solution of Eigenvalue Problems Using Spectral Techniques", *Journal of Computational Physics* **100**, 297-305, (1992) .
13. D. R. Gardner, S.A. Trogdon, and R.W. Douglass, "A modified tau spectral method that eliminates spurious eigenvalues", *J. Comp. Phys.* **80**, 137 (1989).
14. S. G. Yiantsios and B. G. Higgins, "Linear stability of plane Poiseuille flow of two superposed fluids", *Phys. Fluids* **31**, 3225-3238 (1988).

SELF LUBRICATION OF BITUMEN FROTH IN PIPELINES

Daniel D. Joseph

University of Minnesota
Minneapolis, MN 55455
May, 1997

ABSTRACT

In this paper I will review the main properties of water lubricated pipelines and explain some new features which have emerged from studies of self-lubrication of Syncrudes' bitumen froth. When heavy oils are lubricated with water, the water and oil are continuously injected into a pipeline and the water is stable when in a lubricating sheath around the oil core. In the case of bitumen froth obtained from the Alberta tar sands, the water is dispersed in the bitumen and it is liberated at the wall under shear; water injection is not necessary because the froth is self-lubricating.

INTRODUCTION

There is a strong tendency for two immiscible fluids to arrange themselves so that the low-viscosity constituent is in the region of high shear. We can imagine that it may be possible to introduce a beneficial effect in any flow of a very viscous liquid by introducing small amounts of a lubricating fluid. Nature's gift is evidently such that the lubricating fluid will migrate to the right places so as to do the desired job. This gives rise to a kind of gift of nature in which the lubricated flows are stable, and it opens up very interesting possibilities for technological applications in which one fluid is used to lubricate another.

Water-lubricated transport of heavy viscous oils is a technology based on a gift of nature in which the water migrates into the region of high shear at the wall of the pipe where it lubricates the flow. Since the pumping pressures are balanced by wall shear stresses in the water, the lubricated flows require pressures comparable to pumping water alone at the same throughput, independent of the viscosity of the oil (if it is large enough). Hence savings of the order of the oil to water viscosity ratio can be achieved in lubricated flows. Lubricated flow in an oil core is called core annular fow, CAF for short.

Typically, waves appear on the surface of the oil core and they appear to be necessary for levitation of the core off the wall when the densities are different and for centering the core when the densities are matched. We call these flows wavy core annular flow (WCAF). Perfectly centered

core flows (PCAF) of density matched fluids in horizontal pipes and, generally in veritcal pipes, are possible but are rarely stable.

The science behind the technology of CAF has given rise to a large literature which has been reviewed by Oliemans & Ooms [1], by Joseph & Renardy [2], and by Joseph, Bai, Chen and Renardy [3], hereafter called [JBCR]. The story of self-lubrication appears in the open literature first here; apart from the internal Syncrude study of Neiman [4], there is no prior literature.

INDUSTRIAL EXPERIENCE

It is best to start this review with industrial experience since the potential of lubricated lines for energy efficient transport of heavy oil gives this interesting subject an even greater urgency. Heavy crudes are very viscous and usually are somewhat lighter than water, though crudes heavier than water are not unusual. Typical crudes might have a viscosity of 1000 poise and a density of 0.99 g/cm^3 at 25°C. Light oils with viscosities less than 5 poise do not give rise to stable lubricated flows unless they are processed into water/oil emulsions and stiffened.

Oil companies have had an intermittent interest in the technology of water-lubricated transport of heavy oil since 1904. Isaacs & Speed [5] in U.S. Patent #759374 were the first to discuss water lubrication of lighter oils which they proposed to stabilize by centripetal acceleration created by rifling the pipe. For stratified flow, Looman [6] patented a method of conveying oils by passing them over an array of water traps at the bottom of the pipe. An extended history of patents is presented in Joseph & Renardy [2]. The patent history of the subject as it is presently understood starts with the application of Clark & Shapiro [7] of Socony Vacuum Oil company who used additives to reduce the density differences between the oil and water and anionic surfactants to reduce emulsification of water into oil. Clifton & Handley [8] of Shell development proposed to prevent the emulsification of oil at pumps by removing the water before and inserting the oil after the pumps. In fact, water-in-oil emulsions an be pumped in a sheath of water despite the fact that the viscosity of the emulsion can be orders of magnitude larger than the oil alone. In general, lubricated flows are more effective when the oil is more viscous; the water/oil emulsion is an "effective" thickened oil whose density is closer to water. Kiel [9] of Exxon patented a CAF process for pumping heavy oils and water in oil emulsions, surrounded by water, for fracturing subterranean formations to increase oil and gas production. Ho & Li [10] of Exxon produced a *concentrated* water in oil emulsion with 7 to 11 times more water than oil, which they successfully transported in CAF.

Lubricated transport of concentrated oil-in-water emulsions is also an issue. The viscosity of such emulsions can be much smaller than the viscosity of the oil and may be independent of the oil viscosity for large viscosities. This has motivated the consideration of pumping heavy crudes through pipelines as concentrated oil-in-water emulsions. Lamb & Simpson [11] reports a commercial line in Indonsia which carries 40,000 barrels/day of 70% oil/water emulsion in a 20-inch diameter line, 238 kilometers long. Another commerical lubricated transport of Orimulsion®, a coal substitute fuel of 70% oil-in-water produced in Venezuela and marketed by Bitor, can be accomplished naturally since the water for lubrication is already there and will stick to the wall if the surfactant used to stabilize the emulsion and the material of wall construction is suitable (Núñez et al [12]).

Probably the most important industrial pipeline to date was the 6-inch (15.2 cm) diameter, 24-mile (38.6 km) long Shell line from the North Midway Sunset Reservoir near Bakersfield, California, to the central facilities at Ten Section. The line was run under the supervision of Veet Kruka for 12 years from 1970 until the Ten Section facility was closed. When lubricated by water at a volume

flow rate of 30% of the total, the pressure drop varied between 900 psi and 1,100 psi at a flow rate of 24,000 barrels per day with the larger pressure at a threshold of unacceptability which called for pigging. In the sixth year of operation the fresh water was replaced with water produced at the well site which contained various natural chemicals leached from the reservoir, including sodium metasilicate in minute 0.6 wt.% amounts. After that the pressure drop never varied much from the acceptable 900 psi value; the CAF was stable as long as the flow velocity was at least 3 ft/s. Industrial experience suggests that inertia is necessary for successful CAF.

Syncrude Canada Ltd. has undertaken studies of lubricated transport of a bitumen froth which is obtained from processing of oilsands of Alberta for upgrading to synthetic crude. The oil (bitumen) is extracted from mined oilsands rather than pumped directly from the reservoir. A hot-water extraction process is used to separate bitumen as froth from sand and the average composition of the froth is 60, 30 and 10 weight % bitumen, water and solids, respectively. Internal studies led by Neiman et al [4] and recent studies at the University of Minnesota showed that the produced bitumen froth will self lubricate in a pipe flow. The Minnesota studies showed also that once established, lubricating flows were stable and ran continuously for 4 days and nights without any pressure drop buildup or loss of mass flow. Further testing of self-lubrication in 24″ by 1 km pilot pipeline near Fort McMurray confirmed the Minnesota results at a commercial scale. Encouraged by these results, Syncrudes' management has decided to build a 35km line from the new Aurora mine to the upgrading plant at Mildred Lake. Further research impacting on the design of this self-lubricated pipeline of bitumen froth is being done by Syncrude and the University of Minnesota.

LUBRICATED PIPELINES (WITH WATER INJECTION)

Even though lubricated flows are hydrodynamically stable, oil can foul the wall. This is an adhesion rather than a hydrodynamic effect and is not taken into account in the equations used to study stability. The hydrodynamic stability of lubricated flow is very robust even when oil wets the wall. A water annulus can lubricate an oil core even in a pipe whose walls are spotted with oil. Sometimes, however, the fouling builds up, leading to rapidly increasing pressure drops even blocking the flow. Oil fouls some places more than others, near pumping stations where the pressure is highest and the holdup and core wave structure are developing and around line irregularities such as unions, bends, flanges and curves. Another major problem is an unexpected shut-down in the line; the oil and water stratify, causing the oil to stick to the pipe wall, making it harder to restart the line.

It is desirable to lubricate the oil core with as little water as possible because a small water input alleviates the problem of dewatering. On the other hand, oil is more likely to foul the pipe wall when a small amount of water is used, so it is desirable to suppress fouling for this as well as other reasons.

Remedial strategies to prevent fouling naturally alter the adhesive properties of the wall which depend on the solid surface and the oil used. The different strategies that have been employed are discussed in [JBCR]. Clean-up procedures for fouled pipes are important for a viable lubricated pipelines technology.

Obviously, the restart of a fouled pipe will be easier if the oil does not strongly stick to the pipe wall. The restart is also easier if there is an open channel through which water may flow. Such a channel can be opened by stratification under gravity in a large diameter horizontal line. The flowing water will produce a propagating solitary wave near the pump, which tends to partially block the flow of water in such a way that the high local pressure fingers water bewteen the oil and

pipe wall in an unzipping motion which restores core flow as the wave moves forward. The open channel may be closed at places where the pipe goes over a hill since the lighter oil will fill the pipe at high places and make restart more difficult. In small pipes, in which capillarity may dominate gravity, the oil will stratify in slugs separted by water lenses in which water is trapped.

Two-phase flows, liquid-liquid, gas-liquid and liquid-solid flows, can be arranged by dynamics into different configurations called flow types. Various properties of the solutions, like the pressure gradient needed to drive the flow, the area averaged velocity of each phase and the holdup ratio depend on the flow type. The holdup ratio of oil to water is the ratio:

$$h = \frac{Q_0}{A_0} \Big/ \frac{Q_w}{A_w} = \frac{\overline{U}_0}{\overline{U}_w}$$

of the area averaged oil velocity $\overline{U}_0 = Q_0/A_0$ and water velocity $\overline{U}_w = Q_w/A_w$ where Q is a volume flow rate. In the case of an emulsion of small water droplets uniformly dispersed in oil, we could assume that the water moves with the oil so that $\overline{U}_w = \overline{U}_0$, hence $h = 1$. This does not mean that the flow rates or superficial velocities are equal because, say, only a relatively small amount of water could disperse into small drops. In the case of PCAF, with only a very small water input so that the layer of water outside is thin and a uniform velocity for very viscous oil inside, we get

$$\overline{U}_w = \frac{1}{2}\overline{U}_0, \tag{1}$$

hence $h = 2$. In vertical flow with a pressure gradient producing flow of water down against a stationary bouyant slug of oil held in place by gravity, we have $h = 0$ because $Q_0 = 0$ even though the oil fills nearly all of the pipe.

Bai et al [19] found some remarkable holdup results for vertical core flows. For bamboo waves in up-flow, $h \approx 1.39$ independent of the flow rates. Though h varies between 0.8 and 1.4 with flow rates of oil and water, $h \to 1.4$ for fast flows $U_0/U_w > 5$. Thus, $h = 1.4$ is a good value for up-flow and fast down-flows in which buoyancy is not important. In the absence of better knowledge, reasers may assume $h \approx 1.4$.

The lubricated arrangements of oil in water flow are PCAF, WCAF, slugs of oil in water and drops of oil in water. Well-dispersed drops of oil in water are sometimes called oil-in-water (o/w) emulsions. Very concentrated o/w emulsions can be stabilized by surfactants.

Water-in-oil (w/o) emulsions are an oil continuous phase of an "effective" oil which can be lubricated with water. Effective oils of high viscosity can be formed from water in low viscosity oil emulsions (see Ho & Li [10]). When the water content of such w/o emulsions is high, the density of the effective oil will be close to water. It is desirable to transport as much oil as possible at a fixed pressure gradient or to minimize the pressure gradients for fixed oil flow. Minimum pressure gradients are found for flow types near to PCAF; wavy flows are energy efficient and the input ratio can be controlled to give maximum efficiency [JBCR].

Arney et al [14] plotted the friction factor λ vs. Reynolds number $\mathbb{R}$ data for core flows from 12 authors using special definitions which depend on the holup ratio [JBCR]. Most of the data points are in the region of turbulent flow where the Blasius formula $\lambda = 0.316/\mathbb{R}^{1/4}$ applies. There is a considerable scatter of the data which may be due to fouling. Huang et al [15] did turbulence modeling using a K-E model assuming that there were no waves on the core and the core moves forward as a rigid body. The agreement between model predictions, which have not adjustable parameters, and experimental and field data from all sources, was excellent. This result suggests that the major source of drag is turbulence in the water.

Lubrication may fail to fouling, to stratification under gravity when the density difference is large (and core speeds too low to levitate the core off the wall) and to inversion to w/o emulsions

(which can be relubricated as "effective" oils). Failure to stratification at low speeds is characteristic in large diameter pipes in which capillarity is not important. In small pipes, slugs of oil separated by water will stratify in the pipe.

STEEP WAVES & LEVITATION OF CORE FLOWS

A surprising property of core flow is that the flow in a horizontal line will lubricate with the core levitated off the wall even if the core is lighter or heavier than the lubricating water. This levitation could not take place without a hydrodynamic lifting action due to waves sculpted on the core surface. In the case of very viscous liquids, the waves are basically standing waves which are convected with the core as it moves downstream. This picture suggests a lubrication mechanism for the levitation of the core analogous to mechanisms which levitate loaded slider bearings at low Reynolds numbers. Ooms et al [16] and Oliemans and Ooms [1] gave a semi-empirical model of this type and showed that it generated buoyant forces proportional to the first power of the velocity to balance gravity. In this theory, the shape of the wave must be given as empirical input.

Consider water-lubricated pipelining of crude oil. The oil rises up against the pipe wall because it is lighter than the water. It continues to flow because it it lubricated by waves. However, the conventional mechanisms of lubrication cannot work. The saw tooth waves shown in Figure 1 are like an array of slipper bearings and the stationary oil core is pushed off the top wall by lubrication forces. If c were reversed, the core would be sucked into the wall, so the slipper bearing picture is obligatory if you want levitation.

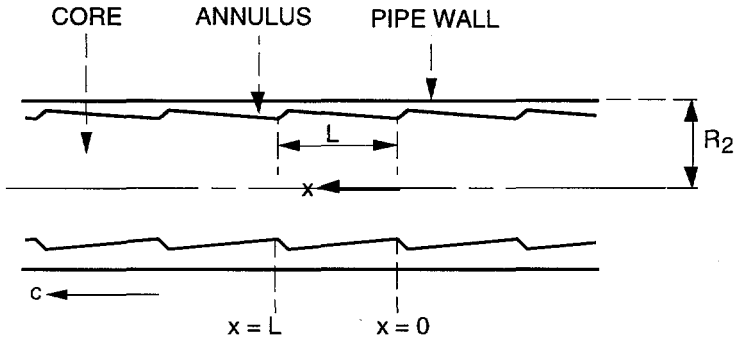


Figure 1: The core is at rest and the pipe wall moves to the left.

Obviously the saw tooth waves are unstable since the pressure is highest just where the gap is smallest, so the wave must steepen where it was gentle, and smooth where it was sharp. This leads us to the cartoon in Figure 2. To get a lift from this kind of wave it appears that we need inertia, as in flying. Liu's [17] formula for capsule lift-off in a pipeline in which the critical lift-off velocity is proportional to the square root of gravity times the density difference is an inertial criterion. Industrial experience also suggests an inertial criterion, since CAF in the Shell line could be maintained only when the velocity was greater than 3 ft/s; at lower velocities the drag was much greater.

Bai et al [19] did a direct simulation of steady axissymetric, axially periodic CAF, assuming that the core viscosity was so large that secondary motions could be neglected in the core. They found

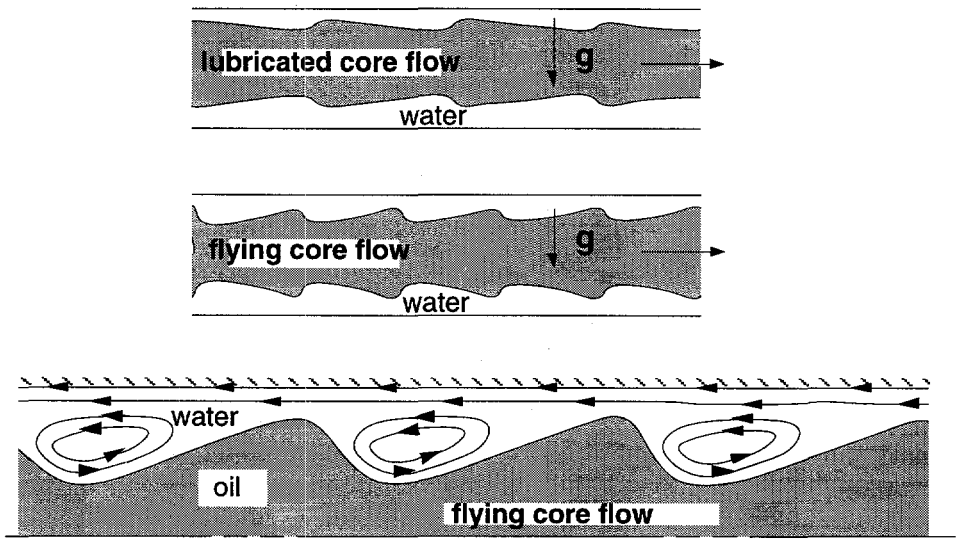


Figure 2: (After Feng et al 1995) (*top*) The interface resembles a slipper bearing with the gentle slope propagating into the water; the shape of these waves is unstable. (*middle*) The high pressure at the front of the wave crest steepens the interface and the low pressure at the back makes the interface less steep. (*bottom*) The pressure distribution in the trough drives one eddy in each trough.

that wave shapes with steep fronts like those shown in Figure 2 always rise from the simulation. The wave front steepens as the speed increases.

A new and important feature revealed by the simulation is that long waves do not arise when the gap size tends to zero as is usually assumed in long wave theories. As the gap size decreases, $\eta \to 1$, the wavelength $\overline{L}(\eta)$ decreases linearly with η. This means that the wave shape hardly changes and a steep wave will stay steep in this limit. The aforementioned properties of waves are in excellent agreement with the observed shape of waves in experiments [JBCR].

SELF LUBRICATION OF BITUMEN FROTH

The first studies of self-lubrication of bitumen froth were carried out in 1985 in unpublished studies by O. Neiman [4] for Syncrude Canada. Tests were conducted on the pipelining characteristics of froth using a research 50mm hydraulic test facility with froth produced on the External Extraction Circuit. They found that

> "The pipeline flow of froth was found to be governed by a special type of multi-phase behavior involving separated water. A concentration of water was observed to occur near the pipe wall, which effectively formed a lubricating water layer around the central plug of froth... Pipelining of froth in the syncrude expansion case may be considered technically and economically feasible."

Syncrude's business is to upgrade the bitumen found in the tar sands into high quality synthetic crude for which there is an active world market. After 25 years of operation, they ran into the

problem that untapped oilsands are located at increasing distances from the upgrading facilities at Mildred Lake. The next ore body they want to mine is at the Aurora mine, 35km north of the upgrading facility, hence the requirements for a bitumen froth pipeline.

The froth is very viscous; to pipeline froth you would have to heat and dilute it, or to use a core flow technology. Neiman's study, industrial experience and the literature on core flow favor the core flow approach but froth is a special material; Neiman did not do tests of long duration, the effects of water addition, pipeline fouling and restart after stopping were not considered. Syncrude identified the DOE supported lubricated pipeline research at the University of Minnesota as most relevant to their needs and they gave us a contract to look at the pipelining of bitumen froth. Our objective was to investigate restart and fouling and to establish an operating envelope and the mechanism for core flow to derive scaling parameters.

The Minnesota studies, in results to be published, established that

- Froth could be pumped in a core flow mode for up to 100 hours.
- External water addition was not required.
- There was no indication of a pressure drop increase even after four days of continuous running.
- An operating envelope from 0.7 m/sec to 2.5 m/sec was established. Core flow was lost below 0.7 m/sec but 2.5 m/sec was an equipment limited value. Possibly unlimited velocities in core flow are possible.
- Shutdowns less than three minutes were re-started without water injection at peak pressure less than 3 times operating pressure.
- Longer shutdown required water addition equivalent 20 wt% lubricating water.

Core annular flow has always referred to a pipelining process requiring a dedicated water addition to form the annulus. One of the insights gained in Minnesota is that froth is self-lubricating, shear liberates some of the water tied up in the froth emulsion. Free water migrates to the point of highest shear– at the pipe wall. Therefore, self-lubricating flow or natural froth lubricity, is probably a better description of the technology.

Second, fouling appears to be inhibited by the liberated water. Bench scale tests show that even carbon steel fouled by the Venezuelans' Zuata crude can be cleaned by liberated water from Syncrude froth.

Third a minimum velocity of 0.7 m/s is required. Below this velocity, the froth core would begin to separate into slugs of lubricated froth between sections of water. This core separation would continue slowly over time until enough water left the annulus that a plug of froth would become unlubricated. At this point pipeline pressures would increase dramatically.

PROPOSED MECHANISM OF SELF-LUBRICATION OF BITUMEN FROTH

We know that froth is highly viscous and offers resistance to *low rates* of shearing. Our proposed mechanism for self lubrication depends on the fact that froth is unstable to *higher shearing rates*, resulting in the coalescence of water droplets to form a lubricating layer of free water. The unique property of bitumen froth that promotes the coalescence of water droplets arises from the dispersion of clay particles in the water.

The clay present in the water inhibits the coalescence of bitumen droplets and promotes the coalescence of the water droplets through a mechanism we refer to as "powdering the dough". Anyone who has baked bread will appreciate this analogy.

Bread dough is sticky, but when a layer of flour is sprinkled on top, the dough will no longer stick to itself or other surfaces. The fine solids surrounding a droplet tend to act as a barrier protecting the oil droplets from coalescing with one another.

Given the encouraging results at Minnesota, Syncrude decided to continue with self lubricating flow as a base case for further development.

The intent for the commercial test was to run the equivalent of 110 000 BPD of bitumen as froth through a 24″ by 1 km pipeline to duplicate the commercial conditions at full scale with respect to line diameter. They were less concerned about line length.

Their first objective for the commercial program was to understand whether or not they could establish self lubricating flow given the shear regime that would exist in a large diameter pipe. It wasn't obvious how to scale-up the re-start results from a 1 inch line diameter to a 24″ line diameter. They also wanted to gain some confidence that the technology would work by seeing it at full scale.

A hurdle to overcome for commercial tests was to understand whether or not they could use a centrifugal pump. Centrifugal pumps are inexpensive and have high capacity but will not operate with fluid viscosity's higher than about 3000 cP. (The viscosity of froth at 50 C is closer to 10000 cP). Positive displacement pumps can operate at much higher viscosities, but require three pumps operating in parallel for the commercial tests, and each pump would have to be among the largest in the world.

However, if our insights about self lubrication were correct then a centrifugal pump could also work in this service. Some quick tests at Minnesota with a centrifugal pump verified that it could indeed work.

Results from the 24″ × 1km pilot pipeline showed that

- Lubricated froth can be established in large diameter pipelines and a centrifugal can be used to drive the flow
- Pipeline shutdowns from 1 minute to 6 hours were tested and restarted without high pressure water injection.

The present decision of management at Syncrude is to build a 35 km lubricated froth pipeline. This pipeline would possibly cost 200 million dollars. The design if this pipeline is presently being worked out.

The scientific details of the lubricated froth studies of Joseph, Bai, Sury, Grant & Mata will be reported in an Archival journal in a future publication.

REFERENCES

1. R.V.A. OLIEMANS and G. OOMS, "Core-annular flow of oil and water through a pipeline. Multiphase Science and Technology," ed G.F. Hewitt, J.M. Delhay, N. Zuber, Vol. 2, Hemisphere Publishing Corporation (1986).
2. D.D. JOSEPH and Y.Y. RENARDY, *Fundamentals of Two-Fluid Dynamics,* Springer-Verlag, New York (1993).

3. D.D. JOSEPH, R. BAI, K.P. CHEN and Y.Y. RENARDY, "Core-Annular Flows," *Annual Review of Fluid Mechanics* 13, 739 (1997).

4. O. NEIMAN, "Froth pipelining tests," *Syncrude Canada Research and Development Progress Report*, 15(1):373-407 (1986).

5. J.D. ISAACS, J.B. SPEED, U.S. Patent No. 759,374.

6. M.D. LOOMAN, U.S. Patent No. 1,192,43, (1916).

7. A.F. CLARK and A. SHAPIRO, U.S. Patent No. 2,533,878 (1949).

8. E.G. CLIFTON, L.R. HANDLEY, U.S. Patent No. 2,821,205 (1958).

9. O.M. KIEL, U.S. Patent 3,378,047, (1968).

10. W.S. HO AND N.N. LI, "Core-annular flow of liquid membrane emulsion," *AIChE Journal* 40, (12): 1961-68 (1994).

11. M.J. LAMB and W.C. SIMPSON, *Proc. World Petroleum Cong., 6th, Frankfurt,* 23-33 (1973).

12. G.A. NÚÑEZ, M. BRICEÑO, C. MATA, H. RIVAS, D.D. JOSEPH, "Flow characteristics of concentrated emulsions of very viscous oil in water," *J. Rheol.*, In press.

13. R. BAI, K. CHEN and D.D. JOSEPH, "Lubricated pipelining: Stability of core-annular flow. Part 5: Experiments and comparison with theory," *J. Fluid Mech.* 240, 97-142. (1992)

14. M.S. ARNEY, R. BAI, E. GUEVARA, D.D. JOSEPH, K. LIU, Friction factor and holdup studies for lubricated pipelining. *Int. J. Mult. Flow*, 19(6): 1061- 1076.

15. A. HUANG, C. CHRISTODOULOU, "Friction factor and holdup studies for lubricated pipelining. Part II: Laminar and $k-\varepsilon$ models of eccentric core flow." *Intl. J. of Mult. Flow.* 20(3): 481-91.

16. G. OOMS, A. SEGAL, A.J. VAN DER WEES, R. MEERHOFF, R.V.A. OLIEMANS, "A theoretical model for core-annular flow of a very viscous oil core and a water annulus through a horizontal pipe," *Int. J. Mult. Flow.* 10: 41-60 (1984).

17. H. LIU. "A theory of capsule lift-off in pipeline," *J. Pipelines*, 2:23–33 (1982).

18. J. FENG, P.Y. HUANG, D.D. JOSEPH, "Dynamic simulation of the motion of capsules in pipelines," *J. Fluid Mech*, 286:201-227 (1995).

19. R. BAI, K. KELKAR and D.D. JOSEPH, "Direct simulation of interfacial waves in a high viscosity ratio and axisymmetric core annular flow," Supercomp. Inst. Res. Rep. #UMSI 95/107, University of Minnesota (1995).

HOMOGENEOUS AND NON-LOCAL HETEROGENEOUS TRANSPORT PHENOMENA WITH VAT APPLICATION ANALYSIS

I.Catton and V.S.Travkin

University of California, Los Angeles
Mechanical and Aerospace Engineering Department
Los Angeles, Ca 90095-1597

ABSTRACT

Some fundamental questions about the mathematical description of transport phenomena in heterogeneous media are addressed to highlight the importance of the smallest scale, of using the correct equations and of choosing a consistent set of scaling parameters. Micro-pore flow is treated using a viscosity correlation that properly treats the near wall region in a capillary. Equations resulting from use of volume averaging theory (VAT) are compared with exact solutions and it shown that certain terms normally ignored must be retained. A consistent set of scaling parameters are derived and shown to yield consistent results over a wide range of different media morphologies.

1. INTRODUCTION

There are many disagreements about the applicability of models based on conventional diffusivity type models of transport phenomena in porous media to problems with the following features: multi-scaled media, media with non-linear physical characteristics, polydisperse morphologies, materials with phase anisotropy, media with non-constant or field dependent phase properties, transient problems, and many others. The most practical way has been to seek a solution by doing numerical experiments over the more or less the exact morphology of interest. This leads to heavy use of large computers to solve large algebraic statements. As in each experiment, the treatment and analysis of the results of such "Direct Numerical Modeling" (DNM) is an unappealing and difficult problem.

Mathematical simulation of physical processes in a highly non-homogeneous media, in general, calls for obtaining averaged characteristics of the medium and, consequently, the averaged equations. The averaging of processes in a randomly organized media can be performed in many different ways. If a physical model has several interdependent structurally organized levels of processes underway, it is expedient to employ one of the hierarchical methods of simulation (see, for example, Kheifets and Neimark [1], Gushman [2] among others). The hierarchical principle of simulation consists of successively studying the processes at a number of structural levels.

One first deals with the smallest scale element, for example a small smooth capillary or globular media. Next, various types of capillary wall morphology are incorporated. This is followed by studies of a range of diameters, first smooth then rough, and then networking. Regular variations of the parameters are treated first, followed by random. This is done at each level. This approach is underway for capillary morphologies as well as granular morphologies. The process leads one to find ways to deal with the large number of closure expressions that result from the averaging used to obtain the original set of equations. Although of a common form, the resulting usable form depends on the media morphology and the local boundary conditions at the each level of the hierarchy. A particular closure expression will be different for energy, mass or momentum transfer between the fluid and the solid matrix.

Three separate topics have been chosen to demonstrate the need for care at every level of the model development process. First, flow in capillaries is treated and it is found that the viscosity is different near the wall than in the bulk fluid. One can postulate why but here we will rely on data and show that there is an impact. It is not large when considering ground water flow, but may well be important in the field of MEMS. The second topic deals with the equations resulting from the use of volume averaging theory (VAT). The equations have strange additional terms that are not usually seen. One needs to ask whether or not these new terms are small enough to ignore. In what follows, it will be shown that they are not. In fact, they are of the same order of magnitude as the one term that is normally kept. The third topic is scaling. There are many analysis and experimental results in the literature. Many use different parameters to scale their data for analysis or in developing an empirical correlation. Many different equations appear in the literature. Based on a simplified VAT model, a consistent set of scaling parameters is developed.

2. FINE PORE NONLINEAR RESISTANCE

The Darcy flow assumption breaks down when the very fine passages found in a porous medium at the smallest scale must be dealt with. Physical phenomena at this scale, less than a micrometer, are often neglected when the scale of the problem is meters or tens of meters. It should be emphasized that in fine pores, the structure of liquids is different from that of the bulk liquid. Indeed, it has been shown by the many convincing experiments of Churaev et al. [3], Derijagin et al. [4], Bondarenko [5,6], and Low [7] that polar liquid (for example, water) boundary layers are characterized by elevated viscosity near boundaries. The modified structure of the liquid affects the filtration process and leads to inapplicability of the Darcy law even at very low Reynolds numbers. Calculations of the flow rates must take into account the dependence of viscosity on distance from the pore wall.

To amplify, we start with the following equation governing the stationary flow of a viscous fluid in a pipe,

$$\frac{1}{r}\frac{d}{dr}\left(r\mu(r)\frac{dV}{dr}\right) = \nabla P,$$

with $\mu(r)$ being the local viscosity near the wall given by Low [7] for a water in the form

$$\mu = \mu_p / (1 - a \exp [-b(R-r)])$$

where μ_p is the bulk viscosity of the fluid far from a wall and the parameters a and b characterize the viscosity on the wall and the corresponding effective thickness of the boundary layer. Low used data from his measurements of water flow velocity in capillaries with radii ranging from $4.5.10^{-6}$ to 10^{-4} cm to determine the Newton viscosity of water close to a solid hydrophobic surface. The water data yielded $a = 0.47$ and $b = 6.35\times10^5$ [cm^{-1}]. Using these results, the mean capillary velocity is found to be

$$V/V_p = 1 - 8a\,[1/2bR - 3/2(bR)^2 + 3/(bR)^3 - 3(1-e^{-bR})/(bR)^4],$$

where $V_p = R^2|\nabla P|/8\mu_p$ is the mean velocity in a capillary when $\mu = \mu_p = \text{const}$. Results for water, using the a and b found by Low are shown on Fig. 1 along with the step-like layered viscosity function near the pore wall.

Overall pressure loss models in porous media with the presence of very fine pores need to take into account this phenomenon when polar liquids transport considered. That means that at least two scales of description have to be combined in the model. That VAT based hierarchical modeling naturally treats such problems was demonstrated by Travkin et al. [8], Travkin and Catton [9].

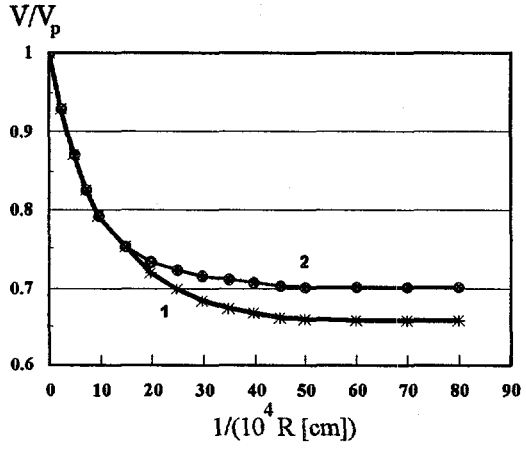


Figure 1. Nondimensional Velocity Against Inverse Radius of Pore for Exponential (1) and Discrete (2) Distribution of Viscosity

3. THEORETICAL VERIFICATION OF VAT BASED MODEL EQUATIONS

Many forms of the energy equation are used in the analysis of transport phenomena in porous media. The primary difference between such equations and those resulting from a more rigorous development based on VAT are certain additional terms. The best way to evaluate the need for these additional more complex terms is to obtain an exact mathematical solution and compare the results with calculations using the VAT equations. This will clearly display the need for using the more complex VAT mathematical statements.

Consider a two-phase heterogeneous medium consisting of an isotropic continuous (solid or fluid) matrix and an isotropic discontinuous phase (spherical particles or pores). The volume fraction of matrix, or f-phase, is $m_f = \Delta\Omega_f / \Delta\Omega$, the volume fraction of filler, or s-phase, is $m_s = 1 - m_f = \Delta\Omega_s / \Delta\Omega$, where $\Delta\Omega = \Delta\Omega_f + \Delta\Omega_s$ is the volume of representative elementary volume (REV). The constant properties (phase conductivities, λ_f and λ_s), stationary (time-independent) heat conduction differential equations for T_s and T_f, the local phase temperatures,

$$-\nabla\cdot\mathbf{q}_f = \lambda_f\nabla^2 T_f = 0 \quad \text{and} \quad -\nabla\cdot\mathbf{q}_s = \lambda_s\nabla^2 T_s = 0,$$

with the interfacial conditions representing perfect thermal contact

$$T_s = T_f \quad \text{and} \quad \mathbf{ds}_1\cdot\mathbf{q}_f = \mathbf{ds}_1\cdot\mathbf{q}_s \quad on\ S_\omega\ ,$$

describe the problem of interest. Here $\mathbf{q}_f = -\lambda_f\nabla T_f$ and $\mathbf{q}_s = -\lambda_s\nabla T_s$ are the local heat flux vectors, S_ω is the interfacial surface and $\mathbf{ds}_1$ is the unit vector outward to the s-phase. No internal heat sources are present inside the composite sample so that the temperature field is determined by the boundary conditions at the external surface of the sample. After correct formulation of these conditions, the problem is completely stated and has a unique solution.

Two ways to realize a solution to this problem are compared below (Travkin and Kushch [10]). The first is the conventional way to replace the actual composite by an equivalent homogeneous medium with an effective thermal conductivity coefficient, $\lambda = \lambda_{eff}(m_s, \lambda_f, \lambda_s)$, assuming you know what it is. To obtain the exact effective thermal coefficient was determined using DNM based on the mathematical theory of globular morphology multiphase fields developed by Kushch [11,12,13]. Averaging the heat flux $\langle\mathbf{q}\rangle$ and temperature $\langle T\rangle$ over the REV yields $\langle\mathbf{q}\rangle = \lambda_{eff}\nabla\langle T\rangle$, and for the stationary case there results

$$\nabla \cdot \left(\lambda_{eff} \nabla \langle T \rangle \right) = 0. \tag{1}$$

The boundary conditions for this equation are formulated in the same manner as for a homogeneous medium. The second way is to solve the problem using the VAT two equation, three term integro-differential equations (see, for example, Travkin and Catton [8, 9]). The solid phase equation to be solved is

$$\frac{\partial}{\partial z}\left(m_s \frac{\partial}{\partial z}\{T_s\}_s \right) + \frac{\partial}{\partial z}\left[\frac{1}{\Delta\Omega} \int_{\delta S_w} \hat{T}_s \vec{ds}_1 \right] + \left[\frac{1}{\Delta\Omega} \int_{\delta S_w} \frac{\partial T_s}{\partial x_i} \vec{ds}_1 \right] = 0, \tag{2}$$

where $\{T\}_s$ is the intrinsic phase temperature averaged over particles, and the second and third terms on the right hand side are the so called morpho-diffusive terms, MD_2 and MD_3 respectively. There is a corresponding equation for the "fluid" (matrix) phase. To evaluate and compare solutions to these equations with the DNM results one needs to know the local solution characteristics, the averaged characteristics over the both phases in each cell and in this case, the additional morpho-diffusive terms.

An infinite homogeneous isotropic medium containing a three-dimensional (3D) array of spherical particles is chosen for analysis. The particles are arranged so that their centers lie at the nodes of a simple cubic lattice with period a. The temperature field in this heterogeneous medium is caused by a constant heat flux $\mathbf{Q}_z$ prescribed at the sample boundaries, which, due to the absence of heat sources, leads to $\langle \mathbf{q} \rangle = \mathbf{Q}_z$. When all the particles have the same radii, the result is the triple periodic structure used widely, beginning from Rayleigh's [14] work, to evaluate the effective conductivity of particle reinforced composites (Fig. 2).

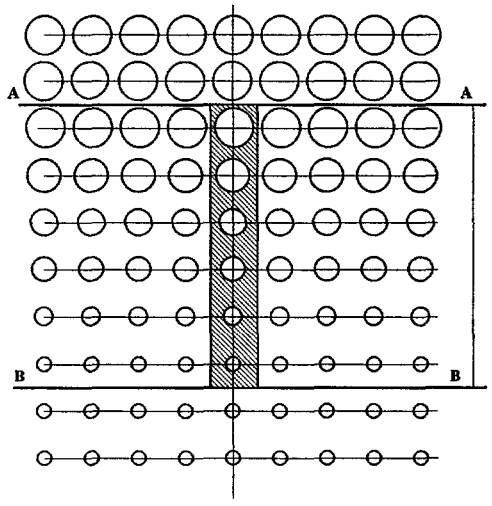


Figure 2. Model of Two-Phase Medium With Variable Volume Fraction of Disperse Phase.

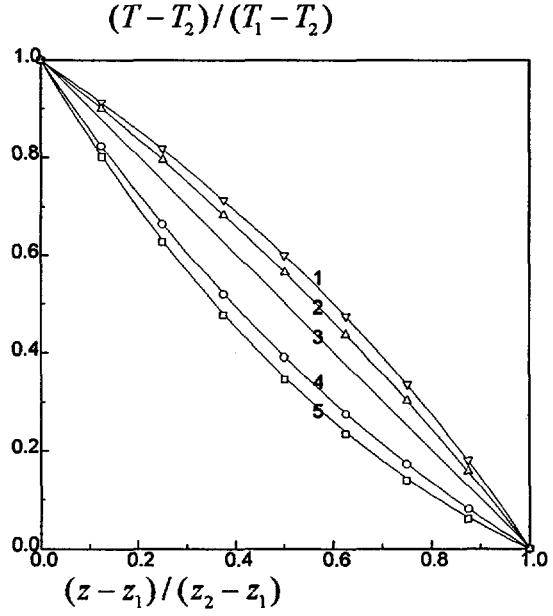


Figure 3. Comparison of VAT Three Term Equation Particle Temperature, Symbols, With the Exact Analytical Based on the Effective Conductance Coefficient Obtained by Exact DNM, Solid Lines.

The model composite medium consists of the three regions shown in Fig. 2. The half-space lying above the A-A plane has a volume content of the disperse phase $m_S = m_A$ and for half-space below the B-B plane $m_S = m_B$. To define the problem, let $m_A > m_B$. The third part is the composite layer between the plane boundaries A-A and B-B containing N double periodic lattices of spheres (screens). For the one-dimensional case, equation (1) becomes

$$\frac{\partial}{\partial z}\left(\lambda_{eff}\frac{\partial}{\partial z}T\right)=0,\ (z_1\le z\le z_2)\,,\ m_s=m_s(z)\,, \tag{3}$$

where $\lambda_{eff}(m_s)$ is the effective conductivity coefficient calculated for the perfect lattice structure composite with $\nabla m_s=0$.

The normalized solution of the both models (VAT and DNM) for the case of linearly changing porosity $m_s=m_s^{(1)}+z(m_s^{(2)}-m_s^{(1)})$, where $m_s(z_1)=m_s^{(1)}$, $m_s(z_2)=m_s^{(2)}$, $z_1=0$, $z_2=1$ are presented in Fig. 3.

Lines 1-5 represent solutions of the one-term equation with effective conductivity coefficients of $\lambda_{eff}=0$, 0.2, 1, 10, and 10000, respectively whereas the points (circles, triangles, etc.) represent the solutions of the VAT equations with accurate DNM closure of the micro model VAT integro-differential terms MD_2 and MD_3 for the composite with varying volume content of disperse phase. Here the number of screens is 9 corresponding to relatively small particle phase concentration gradient. The coincidence of the results of the exact calculation of two equation three term energy transport VAT model with the exact DNM of the one-temperature effective coefficient model for heterogeneous media with non-constant spatial morphology clearly demonstrates the need for using all the terms in the VAT equations. The need for the morpho-diffusive terms in the energy equation are further demonstrated by noting that their magnitudes are all of the same order.

4. SCALING IN EXPERIMENTS AND DATA ANALYSIS

Ergun [15] suggested two types of friction factors. One of them, the so-called kinetic energy friction factor, f_{ker}, is similar to the Fanning friction factor, f_f. When written with the same assumptions,

$$f_f=\frac{f_{ker}}{3}=\frac{d_h}{2\,\varrho_f\,\widetilde{\overline{u}}^2}\left(\frac{\Delta P}{L}\right),$$

where d_h is the hydraulic diameter and $\widetilde{\overline{u}}$ the intrinsic averaged velocity. The problem is what to choose for the hydraulic diameter for a given porous media that properly represents its morphology. Bird et al. [16] used the ratio of the "volume available for flow" to the "cross section available for flow" in their derivation of a hydraulic radius r_{hb}. This led them to

$$r_{hb}=\frac{\langle m\rangle\, d_p}{6(1-\langle m\rangle)},$$

where $\langle m\rangle$ is the porosity. A basis for a consistant hydraulic diameter for any systems is

$$d_h=\frac{4\,\langle m\rangle}{S_w}=\frac{4\,\langle m\rangle}{a_v(1-\langle m\rangle)}=\frac{2\,\langle m\rangle}{3(1-\langle m\rangle)}d_p=4r_{hb},$$

where a_v is the "particle specific surface area" which is equal to the total particle surface area divided by the volume of the particles and specific surface $S_w=a_v(1-\langle m\rangle)$. This expression is justified when an equal or mean particle diameter is $d_p=6/a_v$, which is exact for spherical particles and often used as substitute for granular media particles. The value of hydraulic radius given by Bird et al. [16] r_{hb} was chosen by Chhabra [17] and used in his determination of a specific friction factor for capillary models. The expression for friction factor given by Chhabra is

$$f_{cb}=\frac{d_p}{\langle m\rangle^2\,\varrho_f\,\widetilde{\overline{u}}^2}\frac{\Delta p}{L}.$$

This friction factor, the friction factor, f_b, given by Bird et al. [16] for a packed bed, see equation (6.4-1), the Fanning friction factor, f_f, and to Ergun's kinetic energy friction factor, f_{ker}, related by the following

$$f_{cb}=2f_b=\left(\frac{1-\langle m\rangle}{\langle m\rangle^3}\right)f_{ker}=3\left(\frac{1-\langle m\rangle}{\langle m\rangle^3}\right)f_f\,.$$

These models all use different length scales. It is not clear from evaluation of experimental data which is the most appropriate. To address this, the momentum equation for turbulent flow of an incompressible fluid in a porous

media based on a K-theory (Primak et al. [18], Travkin and Catton [19]) will be used to develop a consistant set of morphological properties and a characteristic length.

The one dimensional form of the momentum equation is

$$-\frac{d\widetilde{\overline{p}}}{dx} = \left(c_f + c_{dp}\frac{S_{wp}}{S_w}\right)\left(\frac{S_w}{\langle m\rangle}\right)\frac{\rho_f\,\widetilde{\overline{u}}^2}{2} = c_d\left(\frac{S_w}{\langle m\rangle}\right)\frac{\rho_f\,\widetilde{\overline{u}}^2}{2},$$

where c_f is the friction factor and c_{dp} the form drag, S_{wp} the cross flow form specific surface. The drag terms are combined into a single total drag coefficient

$$c_d = \left(c_f + c_{dp}\frac{S_{wp}}{S_w}\right),$$

to model the bulk flow resistence terms in the general simplifed VAT momentum equation. The drag resistance can be evaluated for a homogeneous porous media from measurements of pressure drop,

$$-\frac{d\widetilde{\overline{p}}}{dx} = f_f\left(\frac{S_w}{\langle m\rangle}\right)\frac{\rho_f\,\widetilde{\overline{u}}^2}{2}.$$

It was shown by Travkin and Catton [8, 19] that a good approximation for such media is

$$c_d \cong f_f,$$

where a bulk value of pressure loss coefficient f_f can be obtained from experimental correlations. There are few reasons that these quantities are not identical. One of them results because the media inflow and outflow pressure losses are usually incorporated into the correlations for f_f . A second reason is the loss of geometric characteristics represented by the ratio of S_{wp}/S_w. Some details of a media made up of regular globular morphologies can be described in terms of S_w, $\langle m\rangle$, d_p. For a spherical particle medium

$$S_w = \frac{6(1-\langle m\rangle)}{d_p},\quad d_h = \frac{2}{3}\frac{\langle m\rangle}{(1-\langle m\rangle)}d_p.$$

This leads to a relationship for d_h that takes the form

$$d_h = \frac{4\langle m\rangle}{S_w}.$$

This form is consistant with two major morphologies, both capillary and globular, and incorporates two properties of the media, both void fraction and specific surface area. It has a solid theoretical basis for at least for two types of canonical porous media morphology (straight capillary parallel pore morphology - SPPM and one diameter sphere globular morphology) and was arrived at with different theoretical reasonings by others in previous years like Kays and London [20]. There are a number of analysis and experimental studies of different porous media morphology leading to the use of a porous media Reynolds number with this scale. Although various porous media pressure resistance models are described by others (see, Bird et al. [16], Fand et al. [21], and Chhabra [17]), the above description for d_h allows the transformation and comparison of correlation equations and experimental results obtained for diverse morphology media and the use of various scaling. Also it allows experimentally determined characteristics of the media to be related to the closure relationship derived from the VAT analysis. It was shown by Travkin and Catton [8, 19] that a straight equal diameter tube morphology model yields the morphology functions $S_w/ < m >$ and d_h is obtained from the above experession. A c_d for this morphology was found from its corresponding skin friction coefficient expressed by the Darcy pressure loss coefficient by Travkin and Catton [19] $c_d = \widetilde{c}_d \cong f_D/4$.

With this, any experimental correlation can be reformulated for use in the simplified VAT bulk 1D momentum equation,

$$\frac{\Delta P}{L} = f_f\left(\frac{S_w}{\langle m\rangle}\right)\frac{\rho_f\,\widetilde{\overline{u}}^2}{2}.$$

where f_f is an experimentally determined function of $Re_{por} = \widetilde{\overline{u}}d_h/\nu$. With this expression, Fanning friction factor correlations can be easily compared and analyzed after reformulating corresponding expressions in terms of Re_{por} as shown in Fig. 4. The correlations presented in Fig. 4 are for experiments carried out by Gortyshov et al. [22, 23], and Beavers and Sparrow [24] among others. Gortyshov et al. used the Reynolds-Forchheimer momentum equation and a Darcy friction factor. Beavers and Sparrow used a friction factor approach but based their length scale on the measured permeability.

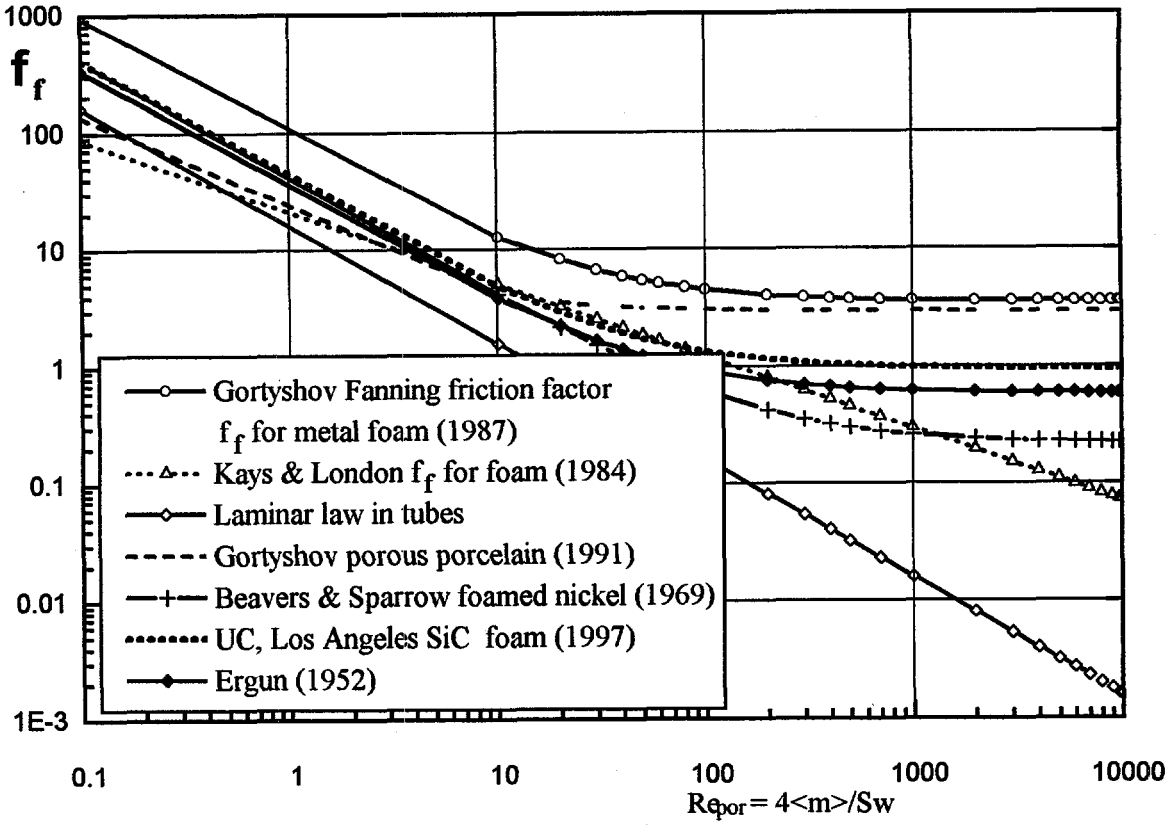


Figure 4. Comparison of Porous Media Pressure Loss Models Using Fanning Friction Factor f_f

5. CONCLUSIONS

Several conclusions are reached as a result of this work. First, one cannot neglect the small scale when dealing with a large problem. When the media includes fine pores, a hierarchical approach is needed that includes at least two levels. Calculations of the flow rates in porous media with a large diversity of morphological elements, e.g. capillaries, must take into account the dependence of viscosity on distance from the pore wall. Care must be taken when the problem statement is formulated that the equations contain all the necessary closure terms resulting from the averaging process. It was shown by comparison of a DNM solution with a VAT based solution that several terms usually ignored are of the same order as those usually kept. There are many terms that appear as a result of flow that are ignored and what their importance is yet to be determined. Analysis of many different experimental and analytical results combined with use of the VAT equations led to the conclusion that pore Reynolds number based on a hydraulic diameter that is a function of void fraction and media specific surface area make up a consistent, although not necessarily complete, set of parameters for correlation of friction factors.

6. ACKNOWLEDGMENT

The authors wish to acknowledge the help and collaboration of M.Pesenson and V.I.Kushch. This work was sponsored by the U.S. Department of Energy, Office of Basic Energy Sciences under the Grant DE-FG03-89ER14033 A002.

7. REFERENCES

1. L.I.Kheifets and A.V.Neimark, *Multiphase Processes in Porous Media*, Nadra, Moscow (1982).
2. J.H.Cushman, "Hierarchial Problems: Some Conceptual Difficulties in the Development of Transport Equations," *Inter. Seminar on Porous Media,* Dubrovnik, (1991).
3. N.Churaev, V.Sobolev, and Z.Zorin, "Thin Liquid Films and Boundary Layers," in *Special Discussion of the Faraday Society*, Academic Press, London, **1**, 213 (1970).
4. B.V.Derijagin, N.Churaev, and V.Muller, *Surface Forces,* Consultant Bureau, NY (1987).
5. N.F. Bondarenko, "On the Nature of Fluid Filtration Anomaly," *DAN SSSR*, **177**, 383 (1967).
6. N.F. Bondarenko, 1968, "Intermolecular Hydrogen Bonds Influence on the Character of Fluid Flow in the Capillaries," *Zh. Fiz. Khim.*, **42**, 225 (1968).
7. P.F.Low, "Viscosity of Interlayer Water in Montmorillonite," *Soil Sci. Amer. Journ.*, **40**, 500 (1976).
8. V.S.Travkin and I.Catton, "Turbulent Transport of Momentum, Heat and Mass in a Two Level Highly Porous Media," *Heat Transfer 1994, Proc. Tenth Intern. Heat Transfer Conf.*, Bighton, **5**, 399 (1994).
9. V.S.Travkin, L.Gratton, and I.Catton, "Modeling Technique for Closure of Transport Equations in Nonregular Globular and Capillary Porous Medium Morphology," *Procs. A.S.M.E./J.S.M.E. Thermal Engineering Joint Conference*, Hawaii, **3**, 319 (1995).
10. V.S.Travkin and V.I.Kushch, "Averaging Theorem Theoretical Closure and Verification," (to be submitted to the *J. Appl. Phys.*) (1997).
11. V.I.Kushch, "Heat Conduction in a Regular Composite With Transversely Isotropic Matrix," *Doklady AN Ukr. SSR*, No.1, 23 (1991), (in Russian).
12. V.I.Kushch, "Thermal Conductivity of Composite Material Reinforced by Periodically Distributed Spheroidal Particles," *Engng.-Phys. Journal*, **66**, 497 (1994), (in Russian).
13. V.I.Kushch, "Elastic Equilibrium of a Medium Containing Finite Number of Aligned Spheroidal Inclusions," *Int. J. Solids Structures*, **33**, 1175 (1996).
14. R.S.Rayleigh, "On the Influence of Obstacles Arranged in Rectangular Order Upon the Properties of a Medium," *Phil. Mag.*, **34**, 481 (1892).
15. S.Ergun,(1952), "Fluid Flow Through Packed Columns," *Chemical Engineering Progress*, **48**, 89 (1952).
16. R.B.Bird, W.E.Stewart, and E.N.Lightfoot, *Transport Phenomena*, Wiley, New York (1960).
17. R.P.Chhabra, *Bubbles, Drops, and Particles in Non-Newtonian Fluids*, CRC Press, Boca Raton (1993).
18. A.V.Primak, A.N.Shcherban, and V.S.Travkin, "Turbulent Transfer in Urban Agglomerations on the Basis of Experimental Statistical Models of Roughness Layer Morphological Properties," *Trans. World Meteor. Organization Conf. on Air Pollution Model. & Application*, Geneva, **2**, 259 (1986).
19. V.S.Travkin and I.Catton, "A Two-Temperature Model for Turbulent Flow and Heat Transfer in a Porous Layer," *J. Fluids Eng.*, **117**, 181 (1995).
20. W.M.Kays and A.L.London, *Compact Heat Exchangers*, 3rd ed., McGraw-Hill (1984).
21. R.M.Fand, B.Y.K.Kim, A.C.C. Lam, and R.T.Phan, "Resistance to the Flow of Fluids Through Simple and Complex Porous Media Whose Matrices are Composed of Randomly Packed Spheres," *J. Fluids Eng.*, **109**, 268 (1987).
22. Yu.F.Gortyshov, G.B.Murav'ev, and I.N.Nadyrov, "Experimental Study of Flow and Heat Exchange in Highly Porous Structures ," *Engng.-Phys. Journal*, **53**, 357 (1987), (in Russian).
23. Yu.F. Gortyshov, I.N.Nadyrov, S.R.Ashikhmin, and A.P.Kunevich, "Heat Transfer in the Flow of a Single-Phase and Boiling Coolant in a Channel with a Porous Insert, " *Engng.-Phys. Journal*, **60**, 252 (1991), (in Russian).
24. G.S.Beavers and E.M.Sparrow, "Non-Darcy Flow Through Fibrous Porous Media," J. Applied Mechanics, **36**, 711 (1969).

INTERFACIAL AREA, VELOCITY AND VOID FRACTION IN TWO-PHASE SLUG FLOW

Gunol Kojasoy

Department of Mechanical Engineering
University of Wisconsin-Milwaukee
Milwaukee, WI 53201

Jovica R. Riznic

Atomic Energy Control Board
Safety Evaluation Division
Ottawa, Canada K1 P5S9

ABSTRACT

The internal flow structure of air-water plug/slug flow in a 50.3 mm dia transparent pipeline has been experimentally investigated by using a four-sensor resistivity probe. Liquid and gas volumetric superficial velocities ranged from 0.55 to 2.20 m/s and 0.27 to 2.20 m/s, respectively, and area-averaged void fractions ranged from about 10 to 70%. The local distributions of void fractions, interfacial area concentration and interface velocity were measured. Contributions from small spherical bubbles and large elongated slug bubbles toward the total void fraction and interfacial area concentration were differentiated. It was observed that the small bubble void contribution to the overall void fraction was small indicating that the large slug bubble void fraction was a dominant factor in determining the total void fraction. However, the small bubble interfacial area contribution was significant in the lower and upper portions of the pipe cross sections.

1. INTRODUCTION

The intermittent plug/slug two-phase flow-pattern exists over a wide range of gas and liquid flow rates in a horizontal two-phase flow configuration. This flow-pattern is described as a gas slug in the form of a large elongated gas bubble in the upper part of the pipe followed by a liquid slug occupying the entire cross section. Depending on the flow rates of the gas and liquid, small bubbles may break off the large gas slug bubbles and either reside in the liquid slug and/or in the liquid underneath the gas slug or coalesce with front of the oncoming gas slugs. In order to advance the study of such a two-phase flow structure, it is essential to experimentally obtain detailed local values of fundamental parameters characterizing the internal structure of this flow-pattern.

The void fraction and interfacial area are two fundamental geometric parameters characterizing the internal structure of two-phase flow. The void fraction represents the volumetric fraction of gas phase or the probability density of the gas to exist at a given time and point and is a required parameter for hydrodynamic and thermal calculations in various industrial processes. On the other hand, the interfacial area describes the

available interfacial area for the interfacial transport of mass, momentum and energy in steady and transient two-phase flows and is a required parameter for two-fluid formulation of a two-phase flow field. However, little information is currently available on these parameters, and it is limited to vertical two-phase flow configurations. Particularly, there exists very little knowledge on the local interfacial area concentration in spite of its importance in multi-dimensional two-fluid model analysis, Boure [1] and Ishii and Kocamustafaogullari [2]. Accurate information about these parameters describing the internal structure of the plug/slug flow-pattern and generalized relationships among them are necessary to understand transport phenomena associated with this flow-pattern.

At present, several methods are available to measure interfacial area concentration in gas-liquid two-phase flows. Measurement techniques can be broadly classified into two categories: (a) chemical absorption methods, and (b) physical methods. Chemical absorption methods provide a global measurements of interfacial area concentration, and thus, do not provide the local information of interest in closure models. Briefly, this method involves measurement of a liquid film controlled gas absorption accompanied by rapid depletion of the solute through reaction in the liquid film. The physical methods involves techniques such as photography, light attenuation, ultrasonic attenuation and various intrusive probe techniques, i.e., resistivity or impedance probes, and fiber-optic probes, etc. Detailed review of all these methods have been provided by Landau et al. [3], Veteau [4], Ishii and Mishima [5], Chang et al. [6,7] and Wang and Kocamustafaogullari [8]. However, all these methods are limited to dispersed two-phase flow patterns, i.e., bubbly and droplet two-phase flows, and with the exception of Revankar and Ishii [9], no established local measurements of the interfacial area concentration exists for flow patterns other than bubbly and droplet flows. Recently, Revankar and Ishii [9] reported a theoretical formulation for the local measurements of interfacial area concentration in cap bubbly flow configuration by using four-sensor resistivity probe. According to the authors knowledge there have been no experimental data or fundamental studies available on the local distribution of interfacial area concentration in horizontal plug/slug two-phase flow-pattern.

In view of this brief discussion, it is evident that much experimental work is still necessary to attain a thorough physical understanding of the internal structure of an intermittent horizontal two-phase slug flow-pattern. In this context, an experimental investigation has been underway at the University of Wisconsin–Milwaukee to investigate the internal structure of this flow-pattern. In these experiments

- the double-sensor resistivity probe method was developed to measure local void fractions, interfacial area concentration, interfacial velocity, local bubble chord-length, size and frequency distributions of bubbles in a horizontal bubbly two-phase flow [10, 11].
- the four-sensor resistivity probe method was developed to measure the local void fraction and interfacial area concentration in a horizontal slug flow [12], and
- the hot-film anemometry technique was used to measure the time-averaged local void fractions due to small and large slug bubbles, as well as, the local axial velocity and turbulence in the liquid phase of an air-water intermittent flow in a horizontal channel [13, 14, 15].

In the present work, the local void fraction and interfacial area concentration in a horizontal plug/slug two-phase flow were studied experimentally by using four-sensor resistivity probe method. Local void fraction, interfacial area concentration, interface velocity have been measured, and the results are documented and discussed here. Furthermore, the dependence of these local parameters on the other flow variables are also presented. Besides the four-sensor resistivity probe method, the hot-film anemometer technique for liquid velocity measurements and two-sensor resistivity method for small bubble interfacial area concentration measurements were also used to supplement the results.

2. FOUR-SENSOR RESISTIVITY PROBE METHOD

2.1 Theoretical Foundation

The theoretical foundation for the local measurement of interfacial area concentration in bubbly and cap bubbly flow using four-sensor electrical resistivity probe was recently provided by Revankar and Ishii [9]. By starting the defining equation of the time-averaged local interfacial area concentration a_i at (x_o, y_o, z_o, t) as

$$a_i\ (x_o, y_o, z_o) = \frac{1}{\Omega} \sum_j \left(|\nabla f_j| / | \frac{\partial f_j}{\partial t} | \right) \tag{1}$$

they showed that the local time-averaged interfacial area concentration can be measured by three interfacial velocity components and the known geometric configuration of the four sensor probe. When the front and three rear sensors of the four-sensor probe are arranged such that they make an orthogonal system with the front sensor located at the apex and the three rear sensors at the rear plane perpendicular to the flow direction, the above relation is simplified to

$$a_i = \frac{1}{\Omega} \sum_j \left[\left(\frac{1}{v_{s1j}} \right)^2 + \left(\frac{1}{v_{s2j}} \right)^2 + \left(\frac{1}{v_{s3j}} \right)^2 \right]^{1/2}$$

In these equations, f_j (x, y, z, t) function represents jth interface in three-dimensional space, Ω is the total sampling time, the subscript j identifies the jth interface and v_{skj} (k = 1,2,3) is the passing velocity components in the three orthogonal directions with respect to rear sensor k.

Schematic of the four-sensor probe arrangement is shown on Figure 1. By considering the jth interface passing the front and any rear sensor k, with the time interval, Δt_{kj} , the passing velocity component in the k'th direction, v_{skj}, is given by

$$v_{skj} = \Delta s_{kj} / \Delta t_{kj}, \quad k = 1, 2, 3 \tag{3}$$

where Δs_{kj} is the spacing between the front and the k'th rear sensor tips. Referring to Figure 1, it is evident that combination of each rear sensor and front sensor constitutes a double-sensor to measure the interfacial velocity component in the respective orthogonal direction. The front sensor is designed to be the common sensor. Thus there are three double-sensor probes enabling the measurement of interfacial velocity components in three orthogonal directions. According to Eq. (2), measurement of the three velocity components v_{s1j} , v_{s2j} and v_{s3j} will be sufficient for determining the local time-averaged interfacial area concentration.

The time-averaged local void fraction, α, at any location, r can be obtained by either front or one of the rear sensors. Due to the finite size of the probe, only the front sensor was used for the void fraction measurements. The time-averaged void fraction at any location r is then determined as follows:

$$\alpha(r) = \frac{1}{\Omega} \left[\sum_{j=1}^{N_{lb}} \left(t_{0fj} - t_{0rj} \right) + \sum_{i=1}^{N_{sb}} \left(t_{0fi} - t_{0ri} \right) \right] \tag{4}$$

where the subscript 0 identifies the front sensor, whereas j and i identify large elongated bubbles, and small spherical bubbles respectively. t_{orj} is the time the front sensor tip enters into the large j'th bubble, and t_{ofj} is the time the front sensor enters in to the liquid phase. t_{ori} and t_{ofi} denote corresponding times for the small spherical bubbles. N_{lb}, and N_{sb}, respectively, are the number of large and small bubbles passing the front sensor tip in

the total sampling time of Ω. The first term appearing in Eq. (4) gives the large bubble void fraction, whereas the second term denotes the small bubble void fraction contribution to the overall void fraction in a slug flow.

2.2 Measurement Principle

In principle, the electrical resistivity probe consists of the instantaneous measurement of local electrical resistivity in the two-phase mixture by means of a sensor electrode. In an air-water flow the air can be considered as electrically insulating, whereas water is electrically conducting. When the sensor is in contact with the liquid, the circuit is closed. On the other hand, when it is in contact with a bubble, the circuit is open. Since the circuit is open or closed depending on whether the sensor is in contact with gas or liquid, the voltage drop across a sensor fluctuates between a V_{min} and V_{max}.

In the case of four-sensor probe method, each sensor and the return electrodes are connected to their own measuring circuits and, therefore, each sensor is used independently as a phase identifying device. Furthermore, from the timing of the shift in the voltage between V_{min} and V_{max}, the time when the gas-liquid interface passes the sensor can be recorded. Therefore, two pieces of parallel and independent information related to the phase identification and the transit time of the gas-liquid interface is obtained.

A schematic diagram indicating a typical time history record of signals from a four-sensor electrical resistivity probe in a slug flow is illustrated in Figure 2. As seen from the figure, the signals, even for the large slug bubbles, deviate from the ideal two-state square-wave signals. This deviation is largely due to the finite size of the sensor causing flow disruption and the possible deformation of the interface before the sensor enters from one phase to the other. The trailing edges are generally steeper than the leading edges. This difference is probably due to the wetting of the sensor by the residual liquid when the sensor enters into the gas phase. When the sensor tips encounter small gas bubbles in the liquid, the residence time in a small bubble is too small for the probe to react to changes fast enough because of drying period of wet sensor tips. As a result, signals for small bubbles do not vary between V_{min} and V_{max} as observed in the case of large bubbles.

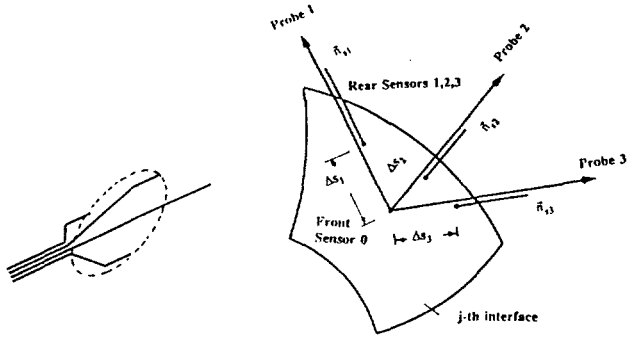


Figure 1. Schematic of four-sensor probe arrangement.

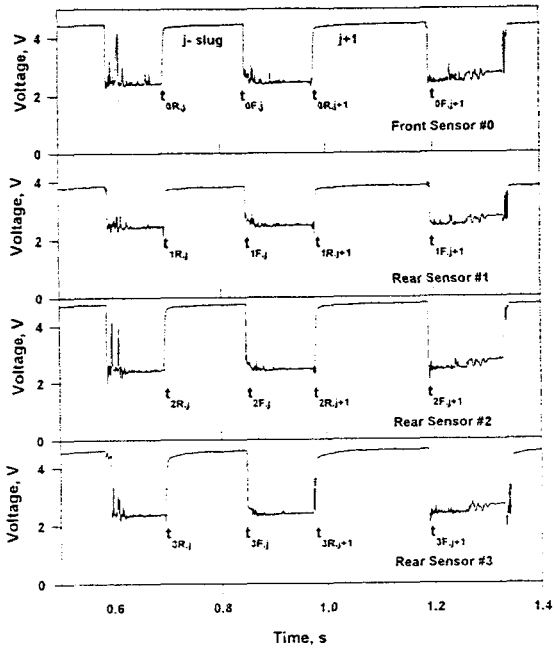


Figure 2. Typical output of four-sensor probe in a slug flow (j_f=2.20 m/s, j_g=1.1 m/s).

2.3 Four-Sensor Probe Design and Signal Processing

A typical four-sensor resistivity probe is shown in Figure 3. It consists of four identical platinum / 13% rhodium or chrome alloy wire sensors of 0.127 mm in diameter. They are completely insulated from the environment except at their exposed tips. The three rear sensor locations are positioned such that the tip of these sensors were on one plane and they were placed axisymetrically with respect to the central front sensor. Typical vertical distance between the front and rear sensor was 2 to 2.5 mm. The locations of the four sensors were arranged such that the tips of the sensors make an orthogonal coordinate system with the front sensor in the apex and the three rear sensors at the rear plane perpendicular to the flow direction. The complete assembly fits into a probe holder from which wires run to the electronic circuit. The electronic circuit uses a 5 V d.c. power supply. Variable resistors were used to enable adjustment of the maximum and minimum voltage signals.

It was found that the proper distance between the front and rear sensors was critical for analyzing the experimental data. Preliminary experiments were conducted to determine a proper distance between the front and rear sensor tips. The distance was dictated by possible bubble size and bubble interface velocity. It was found that 2 - 2.5 mm in lateral direction and 3 - 4 mm in longitudinal direction were the appropriate separation distance for the experimental conditions that were selected. It is to be noted that a very small distance results in inaccuracies in time duration measurements, since it requires very high sampling frequencies or very small bubble velocities. On the other hand, if the distance is too large, there is a strong possibility of misinterpretation of signals since multibubble contact occur between two signals originating from the same bubble.

For the case of the bubble velocity and void fraction measurements the right selection of two closely corresponding signals from each probe is important. This is so, because two sequential signals detected by the front and rear sensors do not always correspond to the same interface, and the residence time intervals of the gas or liquid phase at the sensors are not exactly the same. The signal validation was made by judging whether the following conditions were satisfied:

(1) By assuming the forward motion of the bubbles, the front sensor signal rises or falls before the rear sensor signals do. Therefore, referring to Figure 2, the following conditions should be satisfied

$$t_{kfj} > t_{krj},\ j=1,...N,\ k=0,1,2,3 \quad \text{and} \quad t_{0rj} < t_{krj},\ t_{0fj} > t_{krj},\ j=1,...N,\ k=2,3,4 \qquad \textbf{(5)}$$

where the second indices *f* and *r*, respectively denote the times of the fall and rise in each sensor, N is the number of bubbles passing through a given sensor in the total sampling time Ω.

(2) The residence time of a bubble for the front and rear sensors should be comparable to ensure that both the front and rear sensors detect the same bubble. Hence, the following condition should be also satisfied:

$$t_{0fj} - t_{0rj} \simeq t_{kfj} - t_{krj},\ j=1,...N,\ k = 1, 2, 3 \qquad \textbf{(6)}$$

(3) A passing interface should be properly identified. When there are different sized bubbles as in the slug flow the large slug bubbles and small bubbles should be identified separately. For this the bubble residence time on the front sensor was used to distinguish the elongated, large slug bubbles from the small spherical bubbles. The following condition on the bubble residence time was used to discriminate large and small bubbles:

$$c_1 \le t_{0fj} - t_{0rj} \le c_2$$

where c_1 and c_2 are time constants determined by the calibration process where the slug bubble size is estimated with flow visualization.

3. EXPERIMENTAL SETUP AND PROCEDURE

The experimental flow loop consists of a horizontal line of 50.3 mm ID Pyrex glass tubings with pressure tabs installed between them. The entire flow loop is about 15.4 m in length and is entirely transparent to facilitate flow visualization, high-speed photography and cinematography. Details of the flow loop are given elsewhere [10,11].

The air and distilled water are used as coupling fluids. The air to the test section is supplied from the university central air system. It is regulated through 0.95 m^3 capacity high-pressure storage tank, and metered by a series of turbine flowmeters. The water is pumped from 1.9 m^3 capacity storage tank by a stainless steel centrifugal pump end regulated by a transistor inverter. The flow rate is measured by a series of paddlewheel flow meters assembled in a parallel configuration. The air enters the mixing chamber from a vertical leg and is injected into the water flow through a cylindrical porous media of 100 μm porosity to achieve a uniform mixing. The two-phase mixture from the test section is directed to an air-water separator, where the air is vented to the atmosphere, and the water is returned to the water storage tank.

A Vernier, with graduations to an accuracy of 0.01 mm was used to traverse the probe in a direction perpendicular to the axis of the tube. The probe was traced through the vertical axis of the pipe, stopping at 17 positions to take measurements. The local, instantaneous interface velocity, void fraction and interfacial area concentration were measured at each location. When the experimental studies were intended to investigate the flow development and local interfacial parameters over the cross sectional plane of the pipe, the test section was affixed to the flow channel by rotary seals. The seal provided a water tight joint, yet allowed the test section to rotate freely.

The data from the probe was collected by the data acquisition system and stored in a PC. Due to the limitations of the computer, the sampling rate was set up to a maximum value of 10 kHz which allowed a statistically meaningful sampling time. At this sampling rate the measurement error of the velocity would be less than 5% for the sensor separation distance of 2 mm. Once the data were taken and stored in the memory of the computer, a FORTRAN program was used to process the data, separating the phases and calculating the interfacial velocity components.

4. EXPERIMENTAL RESULTS AND DISCUSSION

4.1 Local Void Fraction

Experimental studies were performed on plug/slug flow patterns, and data were acquired at the axial location of L/D=253 from the mixing chamber which represents a region of semi-fully developed two-phase flow. Four different liquid flow rates in combination with five different gas injection rates were studied. The gas volumetric fluxes were j_g=0.27, 0.55, 1.10, 1.65, and 2.20 m/s.

Equation (4) is used to calculate the local time-averaged void fraction. Typical slug bubble and total void fraction distributions are illustrated on Figure 4, where the total void fraction includes the small and slug bubble contributions. It is interesting to note that the shape of the total void fraction profile is similar to that of the slug bubble void fraction profile, and that the contribution from the slug bubble to the total void fraction is much larger than the small bubble contribution. It is also interesting to note that same findings were observed by Revankar and Ishii [9] for the case of vertical two-phase cap bubble flow. It is also evident from Figure 4 that the relative effect of the small bubbles on the total void distribution is much larger at the lower side of the pipe. This may be attributed to the generation of small bubbles from unstable lower surface of slug bubbles and to the vortex motion generated in front of liquid slugs where the liquid flow field is accelerated to fill the vacuum generated by the faster moving gas slugs. Toward the top of the pipe the small bubble void

fraction contribution stays almost constant although there is a slight increase toward the top of the pipe. This region is characterized by the liquid slug where the small bubbles are homogeneously distributed. However, the small bubbles tend to migrate toward the upper wall under the dominating influence of buoyancy force. Thus, the small bubble void fraction shows a slight increase near the top wall.

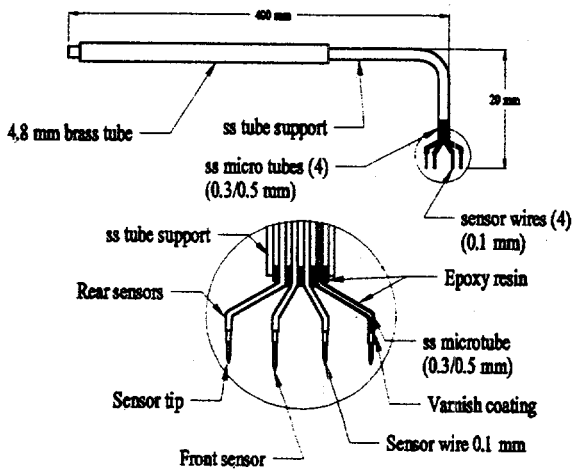


Figure 3. Four-sensor resistivity probe design.

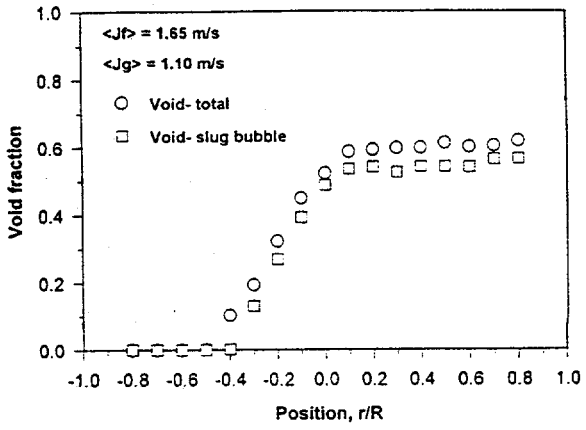


Figure 4. Local void fracation distributions obtained from the front sensor.

The effect of increasing gas flow rate on the total void fractions are, illustrated in Figure 5. It is evident from this figure that the radial profiles of the void fraction are all similar, and, the total void fractions increase with increasing gas flow rate. Comparisons of the void fraction measurement data of the resistivity probe against the hot-film anemometer measurements of References [13, 14] are shown on Figure 6. The agreement between these two data sets are reasonably good, confirming the fact that the four-sensor probe measurements can be performed with a high degree of confidence.

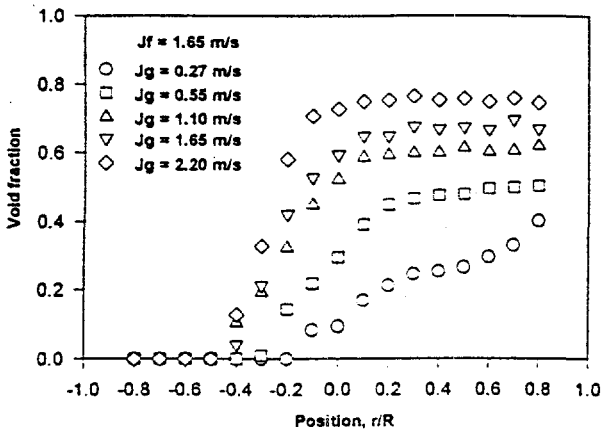


Figure 5. Influence of gas flow on the total local void fraction distributions

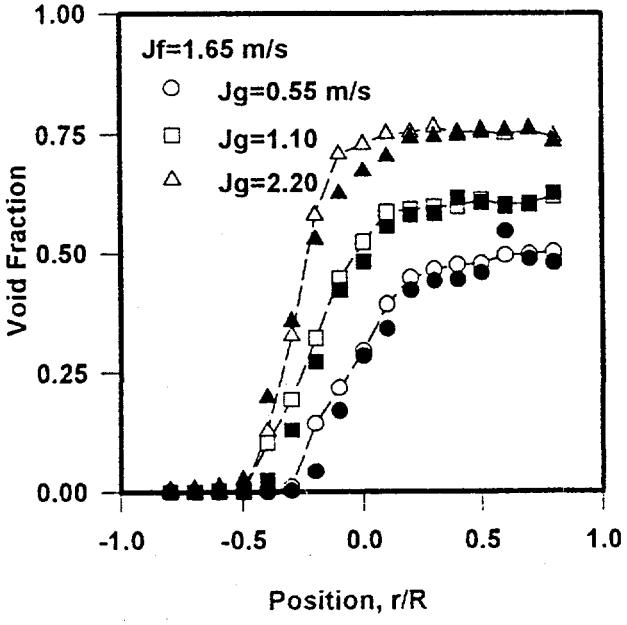


Figure 6. Total void fraction distributions (open symbols-four-sensor resistivity probes; full symbols-hot-film anemometer probe).

4.2 Local Interfacial Area Concentration

The interfacial area concentration is calculated from Eqs. (2) and (3). Figure 7 illustrates a typical interfacial area concentration profile. It is evident from this figure that the contribution from small bubbles is significant. In fact it becomes a dominating factor underneath slug bubbles and toward the top of liquid slugs. These observations are not surprising since the major generation of small bubbles occurs on the lower surface of a slug bubble where the interfacial instabilities exist. On the other hand, gravitational migration of small bubbles within the liquid slug causes an increase in bubble-population, and hence, an increase in the interfacial area concentration toward the top of liquid slugs. It is also interesting to note that interfacial area concentration for large slug bubbles goes through a maximum toward the bottom of slug bubbles. This may be attributed to the curvature effects and interfacial waviness around this location. It has been reported by Tomida et al. [16] that in the slug flow-pattern the waviness of the liquid film surrounding the long slug bubble drastically increases, suggesting an increase in interfacial area concentration.

Figure 8, demonstrate the effect of gas flow on the total interfacial area concentrations. As expected the gas flow has a significant effect on interfacial area concentrations. However, the effect of gas flow on the slug flow area contribution is confined to the bottom portion of slug bubbles. It is to be noted that interfacial area of slug bubbles toward top of the pipe cannot be measured due to finite size of the probe.

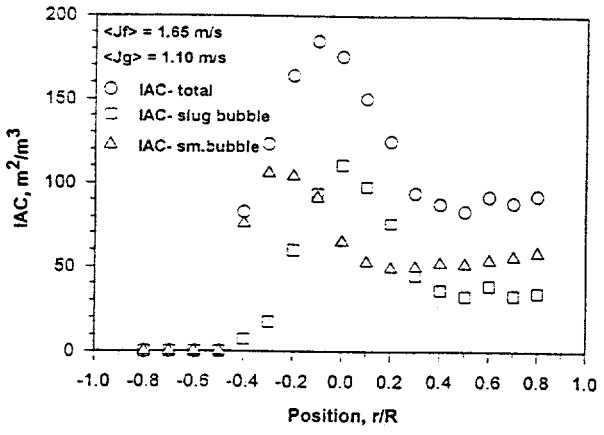


Figure 7. Local Interfacial area distributions

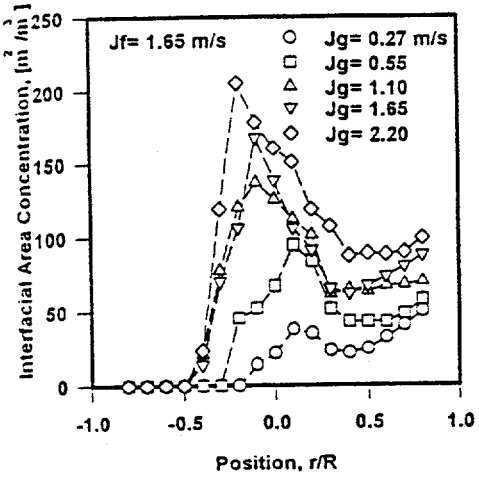


Figure 8. Influence of gas flow on the total local interfacial area concentration distributions.

4.3 Local Interfacial Velocity

The interfacial velocity components in three directions were calculated using Eq. (3). Then the axial velocity component was calculated from the resultant interface velocity, and the results are illustrated inFigure 9. The hot-film anemometer measurements of local liquid velocity [19] are also shown for the purpose of comparisons.

For the gas superficial velocity of 0.55 m/s, the interface velocity profiles were relatively flat over the whole range of investigated liquid superficial velocities. This is not surprising since $j_g \sim$ 0.55 m/s actually falls into the transition from plug to slug flow patterns. It was observed that plugs were moving smoothly with slightly higher velocity as compared to the flow of liquid slugs. As it is shown on the figure there exists a consistent peak at about r/R = 0.0~ 0.2. This was the typical case for all the flow conditions that were considered in the present study.

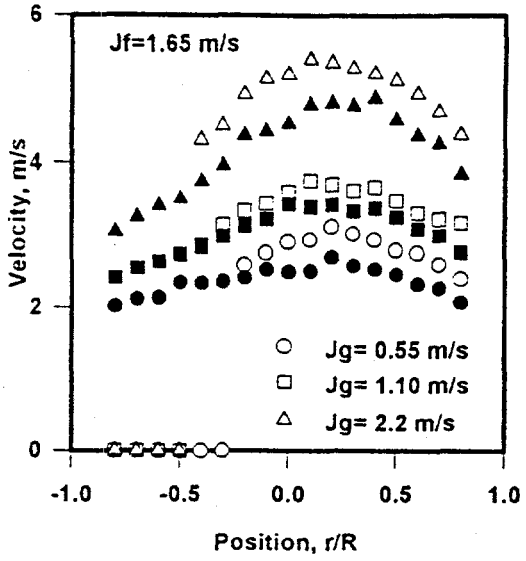


Figure 9. Local interface velocity distributions (open symbols-local interface velocity; full symbols-local liquid mean velocity by hot-film anemometry).

5. SUMMARY AND CONCLUSIONS

The internal phase distribution of air-water flow in a 50.3 mm diameter transparent pipeline has been investigated by using four-sensor resistivity probe technique. The local values of void fraction, interfacial area concentration and interface velocity distributions were measured.

The shapes of the total void fraction and void fraction due to slug bubble profiles for different gas flow rates were all similar, and higher void fractions were recorded with an increase in the gas flow rate. The local measurements obtained from the four-sensor probe compared very well with the hot-film anemometer measurements. In the region of transition from plug to slug flow patterns, with increasing gas flow rate contribution of small bubbles significantly increased.

The interfacial area concentration profiles for slugs clearly show a higher interfacial area along the bottom surface where increased interfacial waviness was reported. Experiments indicated that contribution from the increased number of small bubbles to the overall interfacial area concentration was substantial in the lower part of slug bubbles and in the upper part of liquid slugs.

ACKNOWLEDGMENT

The work reported in this paper was performed under the auspices of the U. S. Department of Energy, Office of Basic Energy Sciences. The authors gratefully acknowledge the support of the U.S. DOE/DES under the direction of Dr. O. P. Manley and Dr. R. Goulard.

REFERENCES

1. J. A. Boure, "Mathematical Modeling of Two-Phase Flow," Proceedings of CSNI Specialist Meeting, Vol. 1 (Edited by S. Banerjee and K. R. Weever), p. 85, AECL, Toronto, Ontario, (1978).

2. M. Ishii and G. Kocamustafaogullari, "Two-Phase Flow Models and Their Limitations," NATO Advanced Research Workshop on Advances in Two-Phase Flow and Heat Transfer, Spitzingsee, Germany.

3. J. Landau, J. Boyle, H. G. Gomaa, A. M. Al Tawell, "Comparison of Methods for Measuring Interfacial Areas in Gas-Liquid Dispersions," *Canadian J. of Chem Eng.*, 55, 13-28 (1977).

4. J. M. Veteau, "Contribution à l'etude des techniques de mesure de l'aire interfaciale dans les ècoulements à bulles, Thèse de Doctorat ès Sciences, Universitè Scientifique et Mèdicale et Institut National Polytechnique de Grenoble, (1981).

5. M. Ishii, K. Mishima, "Study of Two-Fluid Model and Interfacial Area, Argonne National Laboratory Report, ANL-80-111, NUREG/CR-1873, (1980).

6. J. S. Chang, B. Donevski, D. C. Groeneveld, "Examination of Correlations for Interfacial parameters in the Bubbly, Slug and Churn Flow Regimes," in *Heat Transfer Science and Technology,* B. Xuan Wang, Ed. (1988).

7. J. S. Chang, E. C. Morala, "Determination of Two-Phase Interfacial Areas by an Ultrasonic Technique," *Nuclear Engineering and Design,* 122, pp. 143-156 (1990).

8. Z. Wang, W. D. Huang, S. Srinivasmurthy, G. Kocamustafaogullari, 1990 "Interfacial Characteristic Measurements in Horizontal Bubbly Two-Phase Flow," *University of Wisconsin-Milwaukee Report*, DOE/NE/13764-4, October (1990).

9. S. T. Revankar, M. Ishii, "Theory and Measurement of Local Interfacial Area using a Four Sensor Probe in Two-Phase Flow," *Int. J. Heat and Mass Transfer,* 36, 12, 2997-3007 (1993).

10. Z. Wang and G. Kocamustafaogullari, "Interfacial Characteristic Measurements in a Horizontal Two-Phase Flow," *6th Proc. Nuclear Thermal Hydraulics*, 342-351 (1990).

11. G. Kocamustafaogullari, and Z. Wang, "An Experimental Study on Local Interfacial Parameters in a Horizontal Air-Water Bubbly Two-Phase Flow," *Int. J. Multiphase Flow,* 17, 553-572, (1991).

12. J. Riznic and G. Kojasoy, "An Experimental Study on Local Interfacial Parameters in Horizontal Slug Flow," 1997 National Heat Transfer Conference, Baltimore, Maryland, August 10-12, 1997.

13. S. Lewis, M. Herre, V. Davenport, J. R. Riznic, G. Kojasoy, "Use of Hot Film Anemometry in Horizontal Cocurrent Gas-Liquid Slug Flow," *31st ASME/AIChE/ANS/AIAA 1996 National Heat Transfer Conference,* Houston, TX, August, 3-6, ANS Proceedings, HTC-Vol. 9, 326-335, (1996).

14. S. Lewis, "Use of Hot-Film Anemometry in Horizontal Gas-Liquid Slug Flow," *MS Thesis*, University of Wisconsin-Milwaukee, May (1996).

15. S. Sharma, S. Lewis and G. Kojasoy, "Local Studies in Horizontal Gas-Liquid Slug Flow," OECD/CSNI Specialist Meeting on Advanced Instrumentation and Measurement Techniques, Sante Barbara, California, March 17-20, 1997.

16. T. Tomida, F. Yuse, O. Okezaki, "Effective Interfacial Area and Liquid-Side Mass Transfer Coefficient in the Upward Two-Phase Flow of Gas-Liquid Mixture, *Chem. Eng. J.*, 16, 81-88 (1978).

MEASURING THE INSTRUMENT FUNCTION OF RADIOMETERS

Roland Winston
Enrico Fermi Institute
The University of Chicago
Chicago, Illinois 60637, U.S.A.

Robert G. Littlejohn
Department of Physics, University of California
Berkeley, Ca 94720, U.S.A.

ABSTRACT

The instrument function is a function of position and angle, the knowledge of which allows one to compute the response of a radiometer to an incident wave field in any state of coherence. The instrument function of a given radiometer need not be calculated; instead, it may be measured by calibration with incident plane waves.

INTRODUCTION

Since the pioneering work of Walther [1], it has been widely understood that the state of an optical field, including its coherence properties, can be described by a single function (the generalized radiance) which depends on position (x) and angle (or, equivalently, transverse wave number k). The generalized radiance contains all the information available in the various second order correlation functions. The subject of radiometry and coherence has been developed and reviewed by Wolf and coworkers [2]. However, the generalized radiance presents some conceptual difficulties because it can take on negative values in certain regions of the (x,k) phase space, in spite of the fact that the results of all physical radiometric measurements are nonnegative.

In a recent paper [3], we showed how the result of a physical measurement can be represented as the phase-space overlap or scalar product of the generalized radiance with a function which characterizes the instrument. We call the latter function the "instrument function"; like the generalized radiance, it is a function of (x , k). Explicitly, we have

$$Q = 1/2\pi \int dx\, dk\, W(x,k)\, M(x,k) \qquad (1)$$

where Q is the result of the measurement, W (x ,k) is the generalized radiance (according to Walther [1]) and M(x ,k) is the instrument function. Our use of the instrument function is similar to an approach taken earlier by Gase, Ponath and Schubert [4]. As in Ref.3, we keep to a one-dimensional model to bring out the physical ideas without complicating the mathematics.

Thus, the result of a physical measurement is expressed in terms of two functions, one of which (W) characterizes the radiation field, and the other of which (M) characterizes the instrument. Furthermore, the mathematical theory of the Wigner function [5] makes it manifest that the physical result is nonnegative-definite.

In the same paper [3] we calculated the instrument function for a simple radiometer. However, the calculation does not apply universally to all radiometers; in particular not to nonimaging concentrators whose uniform phase-space acceptance properties are attractive for radiometric measurements [6]. Rather than leave the impression that each radiometer design requires a potentially tedious calculation, we show in this paper how the instrument may be calibrated. Calibration is, after all, the common practice and preferred means of using radiometers in the first place.

In this paper we will not discuss any specific radiometer design, since it is our purpose to treat the radiometer as a black box, which merely transforms the incident wave field into a nonnegative definite output signal. Nor is there any assumption that geometrical optics is valid in the inner workings of the instrument; the effects of diffraction, interference, and partial coherence are fully incorporated. All of this is made evident by the theory of the instrument function which was presented in Ref. 3.

SINGLE PLANE WAVE METHOD

Calibration with a single plane wave does not suffice to measure the instrument function of a radiometer for all values of (x , k), because the generalized radiance of a plane wave is uniform over the aperture of the radiometer. However, the information so obtained is sufficient to determine the response of the radiometer in all other circumstances in which the generalized radiance is essentially uniform across the radiometer, including many cases of incoherent, distant sources. We note that whenever the 2-point correlation function of the field, $\Gamma(x,x')$, factors into a slow function of $x+x'$ times a rapid function of $x-x'$, the generalized radiance will itself be a slow

function of x , perhaps sufficiently slow to be sensibly constant over the aperture of the instrument. The condition in question is the same as that which appears in Carter's and Wolf's [7] definition of "quasihomogeneity", except that as far as we can see those authors are thinking of the state of the radiation field at a source, whereas we are referring to the state of the field at the entrance aperture of the instrument.

Varying the incident direction sweeps out the beam-pattern response of the instrument. Let the plane wave in the x,z plane be $\Psi(x,z) = A\exp[i(k_t x + k_l z)]$, where k_t and k_l are respectively the transverse and longitudinal wave numbers. The generalized radiance W is determined from the transverse part of the wave field $\psi(x)$ according to

$$W(x,k) = \int ds\, e^{-iks} \langle \psi(x+s/2)\, \psi^*(x-s/2) \rangle, \qquad (2)$$

where the angle brackets represent the statistical average. In the present case,

$\psi(x) = A\exp(ik_t x)$, so

$$W(x,k) = 2\pi |A|^2 \delta(k-k_t). \qquad (3)$$

The statistical average is unnecessary in this case because the light is coherent.

Then according to Eq. (1), the result of the measurement is

$$Q = |A|^2 \int dx\, M(x,k_t), \qquad (4)$$

where 2 a is the aperture of the radiometer. This shows that the spatial integral of the instrument function M can be determined by calibrating with a single plane wave.

DOUBLE PLANE WAVE METHOD

There are important circumstances in which the generalized radiance is not constant over the aperture of the instrument. For example, often coherent wave fields will generate interference patterns on a scale comparable to or smaller than the aperture. For another example, consider an edge back-

illuminated by an incoherent source, or, equivalently, an incoherent source with a sharp edge. In this case, the best compromise between spatial and angular resolution is obtained by setting $a \approx \sqrt{\lambda z}$ and $\theta \approx \sqrt{\lambda/z}$, i.e., making the size of the aperture comparable to the blurring of the edge from Fresnel diffraction, and making the instrument accept a single cell in phase space. Yet another counter-example is any nearby source. For such applications, we need more detailed information concerning the instrument function.

Therefore we now consider calibrating with a pair of mutually coherent plane waves, which produce the transverse wave field

$$\psi(x) = A\exp(ik_1 x) + A\exp(ik_2 x) , \tag{5}$$

where k_1 and k_2 are the transverse wave numbers. Such a pair of plane waves might, for example, be prepared by a single plane wave incident on a beam-splitter. Then according to Eq. (2) we find

$$W(x,k) = 2\pi \{ |A_1|^2 \delta(k - k_1) + A_2|^2 \delta(k - k_2) + [A_1 A_2^* \exp i(k_1 - k_2) + c.c.] \delta[k - (k_1 + k_2)/2] \} \tag{6}$$

The first two terms are the generalized radiance functions of the two incident plane waves taken one at a time, while the third term contains the interference effects. By subtracting the results of two single plane wave calibrations as described above, we can measure the quantity

$$Q_{reduced} = \int dx\ M(x,k_+)[A_1 A_2^* \exp i(k_- x) + c.c.] \tag{7}$$

where

$$k_+ = (k_1 + k_2)/2 , k_- = (k_1 - k_2) \tag{8}$$

First consider the case that A_1 and A_2 are in phase, say, $A_1 = |A_1| \exp i\alpha$ and $A_2 = |A_2| \exp i\alpha$. Then by holding k_+ fixed and varying k_-, we obtain

$$Q_c = 2|A_1||A_2| \int dx\ M(x,k_+) \cos(k_- x) , \tag{9}$$

which is the cosine transform of M in x for a fixed value of k. On the other hand, if A_1 and A_2 are $\pi/2$ out of phase, say, $A_1 = |A_1| \exp i\alpha$ and $A_2 = |A_2| \exp i\alpha + i\pi/2$, then we obtain

$$Q_s = 2|A_1||A_2| \int dx \; M(x,k_+) \sin(k_- x) \qquad (10)$$

which is the sine transform. Then the quantity $Q_c + i\,Q_s$ is proportional to the Fourier transform of M in x , which may be inverted to find M itself as a function of x and k. As was discussed in Ref. 3, the instrument function is appreciable in the acceptance region of the radiometer in phase space, as determined by geometrical optics; but it also has nonzero values outside this region, due to diffraction effects. On the other hand, M does fall off as one moves away from the geometrical acceptance region so the range of k values which must be examined in the calibration we propose is limited.

CONCLUSIONS

In classical radiometry, the instrument is routinely calibrated by measuring the beam-pattern angular acceptance of a distant (plane-wave) source. In the context of generalized raidometry we find that a complete specification requires measurement of a two-point angular correlation. Two-point angular correlation measurements are important in astrophysics in, for example, measurement of inhomogeneities in the cosmic three degree background radiation. It would be interesting to consider whether the interpretation of such measurements would require the kind of detailed information suggested in this paper.

ACKNOWLEDGMENT

This work was supported in part by the U.S. Department of Energy, Basic Energy Science under contracts Nos. DOE-DEFG02-87ER-13726 and DE-AC03-76SF0098.

REFERENCES

1. A. WALTHER, "Radiometry and Coherence," J. Opt. Soc. Am. 58, 1256-1259 (1968).

2. EMIL WOLF, "Coherence and Radiometry," J. Opt. Soc. Am. 68, 6-17 (1978).

3. ROBERT G. LITTLEJOHN and ROLAND WINSTON, "Generalized Radiance and Measurement," J. Opt. Soc. Am. A 12, 2736-2743 (1995).

4. R. GASE, H.-E. PONATH and M. SCHUBERT, "On a Temporal-Spatial Functional and its Measurement," Annalen der Physik 7, 487-498 (1986).

5. E. WIGNER, "On the Quantum Corrections for Thermodynamic Equilibrium," Phys. Rev. 40, 749-759 (1932).

6. W. T. WELFORD and R. WINSTON, High Collection Nonimaging Optics (Academic Press, New York and London, 1989).

7. W. H. CARTER and E. WOLF, "Coherence and Radiometry with Quasihomogeneous Planar Sources," J. Opt. Soc. Am. 67, 785-796 (1977).

Applications of Nonimaging Optics for Very High Solar Concentrations

J. O'Gallagher and R. Winston
Enrico Fermi Institute
The University of Chicago

Abstract

Using the principles and techniques of nonimaging optics, solar concentrations that approach the theoretical maximum can be achieved. This has applications in solar energy collection wherever concentration is desired. In this paper, we survey recent progress in attaining and using high and ultrahigh solar fluxes. We review a number of potential applications for highly concentrated solar energy and the current status of the associated technology. By making possible new and unique applications for intense solar flux, these techniques have opened a whole new frontier for research and development of potentially economic uses of solar energy .

INTRODUCTION

In 1991, the solar energy research group at the University of Chicago measured solar flux levels exceeding those at the surface of the sun itself [1]. This remarkable achievement brought solar concentrator research into a new domain and served as a dramatic demonstration of the power of the techniques of nonimaging optics which allows the design of optical devices that approach the thermodynamic (ideal) limit, that is, the maximum possible concentration achievable for a given angular field of view. There are a whole family of single element nonimaging concentrators designed for relatively large acceptance angles, most notably the so called Compound Parabolic Concentrator or CPC. These are useful for achieving effective collection and concentration of sunlight at moderate to intermediate levels without tracking. The principles and applications of these designs are now well known and will not be discussed here. However for small angular acceptances and corresponding large concentration ratios, these single stage devices become extremely long and narrow and as such are not practical for attaining very high concentration, so that two-stage systems must be employed. In this report, we survey recent applications of a variety of two-stage designs using nonimaging secondaries which make possible the generation of high and ultra high solar fluxes .

THE THERMODYNAMIC LIMIT AND THE CONCENTRATION SOLAR THERMAL ENERGY

The limit to which a beam of light can be concentrated (or deconcentrated) is fundamentally connected to its angular divergence, θ_i, by the well known **sine law of concentration** [2]. This is

$$C_{max} = \frac{n^2}{\sin^2(\theta_i)} \qquad (1)$$

where n is the index of refraction at the target surface. Here θ_i is the total angular subtense of the source including all optical and pointing errors. For the case $n = 1$ and $\theta_i = \theta_s = 4.7$ milliradians, the half angular subtense of the sun, Eqn. (1) yields the well known limit of 46,000 on the concentration of terrestrial sunlight . This allows for no optical or tracking errors.

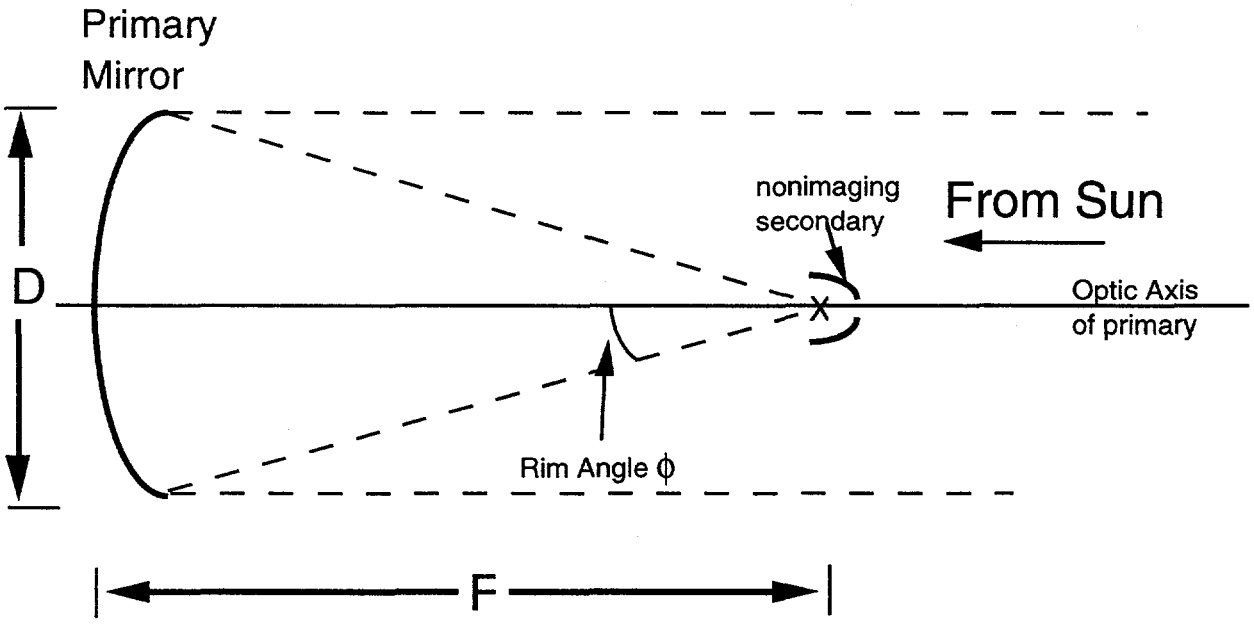


Figure 1. Schematic summary of configuration for a two-stage nonimaging concentrator.

THE LIMITATIONS OF CONVENTIONAL FOCUSING CONCENTRATORS

An early impetus to the development of nonimaging optics was the realization that conventional imaging optics falls far short of the sine law limit. For example, a parabolic reflector achieves, at best, one quarter of the sine law limit. It is easily shown [2] that the concentration of a single stage paraboloid or any other conventional reflecting focussing primary falls well short of the limits of Equation 2. In particular, a primary with a convergence angle half-angle (or rim angle) ϕ, as schematically illustrated in Figure 2, can attain a geometrical concentration ratio of at most

$$C_{1,\,max} = \frac{\cos^2\phi \sin^2\phi}{\sin^2(\theta_i)} \qquad (2)$$

and thus falls short of the maximum limit by a factor of $\frac{n^2}{\cos^2\phi \sin^2\phi}$. The maximum value of Equation 2 occurs at $\phi = 45^o$, and thus, even for this best single stage concentrator, the shortfall with respect to the limit of Equation 1 is at least a factor of $4\,n^2$. However by a suitable choice of a secondary concentrator deployed in the focal zone of the primary or "first stage" one can recover much of this loss in concentration.

Advantages and Features of Two-Stage Concentrating Systems The most effective designs (currently in use) for attaining very high concentrations and approaching the thermodynamic limit with a practical system are the so-called two-stage configurations [3,4] comprised of a focusing first stage (or primary) and a nonimaging second stage (or secondary) deployed in the focal zone of the primary as indicated schematically in Figure 1. From Eqn. (1), we know that a nonimaging secondary can achieve a geometrical concentration factor of

$$C_{2,max} = \frac{n^2}{\sin^2\phi} \tag{3}$$

where ϕ is the rim angle of the primary. Combining this with the concentration of a single stage focusing primary (Eqn. 3) we find that the practical geometric limit (with a fixed acceptance secondary) is

$$C_{2\text{-stage, max}} = C_1 \cdot C_2 = \frac{n^2 \cos^2\phi}{\sin^2\theta_i} \ . \tag{4}$$

It should be recognized that the single stage limit of Equation 3 applies not only to paraboloids, but to *any concave focusing primary* (Welford and Winston, 1980). The two stage limit of Equation 5 comes close to the ideal limit of Equation (1) for small ϕ, i.e., for large focal length to diameter (F/D) ratios. The rim angle ϕ is related to the focal ratio f = F/D by

$$\phi = \tan^{-1}\left[\frac{1}{(2f - 1/(8f))}\right] \tag{5}$$

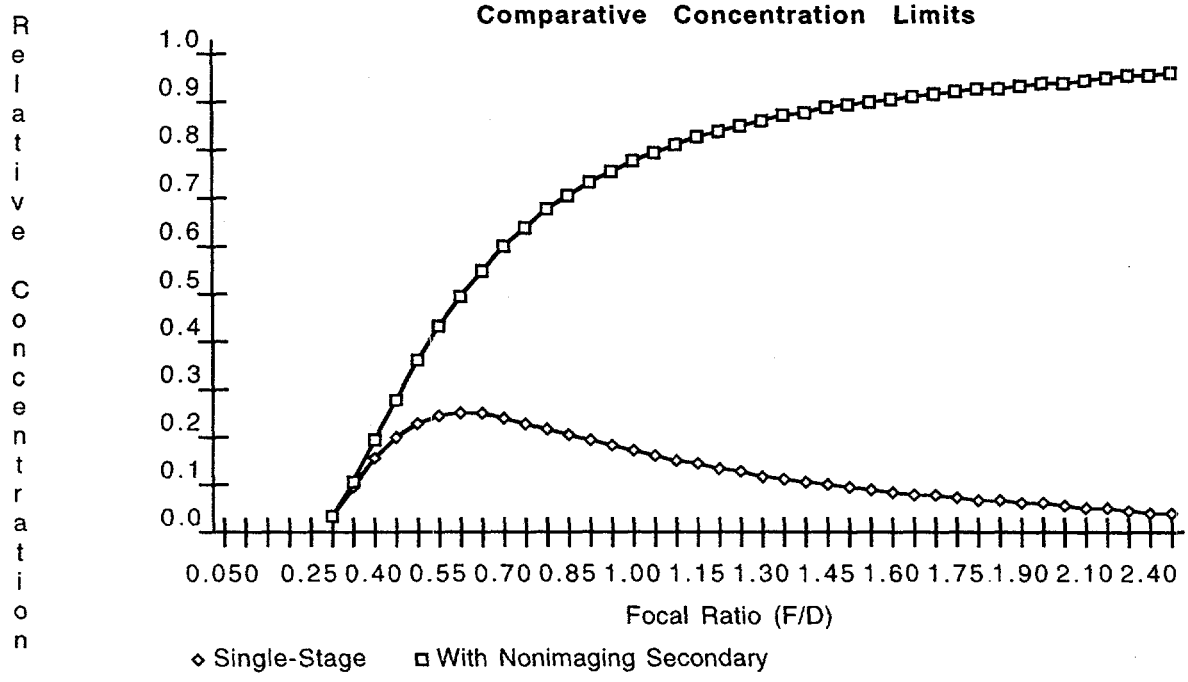


Figure 2. Comparison of the maximum geometric concentration attainable by one and two stage systems with the ideal limit.

Design Considerations for Nonimaging Secondary Concentrators. There are three different regimes for applications. These are 1) ultra-high flux applications in air, 2) ultra-high flux applications in a refractive medium, and 3) conventional dish-thermal retrofit applications. Corresponding secondaries of choice in each of these regimes respectively are 1) Reflecting CPCs (or CECs) [3] , 2) Refracting Hi-Index Nonimaging concentrators; either dielectric filled CPCs (DCPCs) or Dielectric Totally Internally Reflecting Concentrators (DTIRCs) [5], and 3) flow-line or "trumpet" concentrators (Winston and Welford, 1979). Each of these types of secondaries has advantages and disadvantages and particular features that must be borne in mind in optimizing its design.

Demonstration and Measurement of Ultrahigh Solar Fluxes. Following through with a long-standing desire to explore the development of these techniques for larger scale, higher power applications, the National Renewable Energy Laboratory (NREL), formerly the Solar Energy Research Institute (SERI) designed and constructed a scaled-up solar concentrating furnace facility specifically intended to make use of nonimaging optics [6]. This High Flux Solar Furnace (HFSR) concept uses a modified long focal length design and is capable of delivering up to 10 Kilowatts to the focal zone and in particular to the entrance aperture of a secondary concentrator.

The measurement of ultrahigh fluxes may seem straightforward in principle, but it is difficult to implement in practice. For this reason, most of the high flux measurements based on these techniques rely on fluid based calorimetric measurements. Because of difficulties in measuring high fluxes in large scale high-index mediums, new techniques were developed [7,8,9] to extract light from higher index materials into a lower index ones. A new way to measure the flux passing through a large dielectric aperture developed from research at the Weizmann Institute of Science (WIS) in Israel and at the University of Chicago. These investigations studied the use of long extensions attached to the exit apertures of secondary nonimaging concentrators to extract light into air. A relatively simple device was developed that extracts light into air from a borosilicate medium (n=1.46) with an efficiency greater than 99%. New DTIRC secondaries were fabricated with such "extractor tips" . These secondaries were then combined with the cold water calorimeter developed for the air measurements and a successful measurement of the power delivered through a small aperture surrounding the "waist" of the extractor tip was carried out.

APPLICATIONS USING HIGHLY CONCENTRATED SUNLIGHT

Recent experiments and others performed at the National Renewable Energy Laboratory's (NREL) High Flux Solar Furnace (HFSF) have demonstrated the effectiveness of using concentrated sunlight and advanced nonimaging secondaries to pump lasers and produce fullerenes (potentially useful new forms of molecular carbon). The recently developed techniques that allow more flexibility in design [10] have been used to develop two new configurations, each of which couples the high solar flux available at the HFSF to unusually shaped targets which impose unusual constraints.

Solar-Pumped Lasers There are two methods for pumping lasers using sunlight, a) end-pumping and b) side-pumping. . There are advantages and disadvantages for each. The end-pumping scheme can be used on very small scales, but the pump light enters the laser crystal entirely from one end. Previous high-flux measurements show that this system can not be scaled up indefinitely as the optical coupling between dielectric surfaces degrades when exposed to ultrahigh solar fluxes at the kW power scale. An end-pumping scheme has a maximum output power to about 5 W. The side-pumping scheme doesn't face this limitation, because the lasing medium is excited sunlight entering over a much larger aperture (the side walls of the laser crystal). Current research suggests that space-based applications hold the most promise for solar

lasers. In space, the solar insolation level, the input pump source, is more intense and much more stable. This in turn allows more stable lasing configurations that increase laser output efficiency and brightness (and building an electrically-pumped laser is difficult).

Fullerene Production. Researchers at Rice University announced the first creation and isolation of macroscopic quantities of C_{60} and C_{70} molecules and other less-common fullerenes by placing a carbon arc inside a partially evacuated inert atmosphere (Helium gas at around 1/6 of an atmosphere) and supplying a high voltage. However, there turned out to be a problem with using arc lamps to produce fullerenes in that, as the size of the arcs increased, the yield relative to total soot vaporization decreased. This led to the consideration of using highly concentrated sunlight. Sunlight on earth has a relatively low UV content and mirrors can be made less reflective at lower wavelengths, making it a promising source to produce fullerenes. Using solar flux to produce small amounts of fullerenes was demonstrated simultaneously by both researchers at Rice University and researchers at NREL. The method at NREL used a nonimaging secondary CPC in conjunction with the High Flux Solar Furnace. More recent work is underway to develop mass-production techniques. The scalability of solar as opposed to small scale arc lamp systems may make this the preferred method of fullerene production.

Solar Processing of Materials. Ideally, concentrated sunlight can cause materials to reach equilibrium temperatures approaching those found on the sun's surface (5,800 K). This allows solar to cover a wide range of applications at all temperatures below this level. The high fluxes also lead to extremely high heating rates in non-equilibrium setups because the heating rate is proportional to the incident flux minus the reradiation losses, which are small for low temperatures. Surfaces of materials can be superheated to induce chemical reactions that modify the properties and composition on a material's surface while leaving the bulk material unchanged.

Dish-thermal Applications (Reflecting Trumpet Secondary) A project is being carried out by he University of Chicago to design a practical trumpet secondary concentrator to be use in combination with a faceted membrane primary concentrator for dish-stirling applications. This is a retrofit design for a dish which was originally designed as a single stage and so it cannot attain the full power of a fully optimized two stage concentrator. Nevertheless, preliminary ray trace studies show that the addition of a small trumpet to the receiver aperture will allow a reduced aperture size and corresponding lower thermal losses at the operating temperature of 675° C.

Solar thermal Applications of High-Index Secondaries. There is interesting research at the Weizmann Institute of Science (WIS) in Israel, studying the use of large-scale dielectric secondary concentrators and an efficient extractor tip to create high-temperature gas turbine engines. A summary of their system is given in Ref. [9], but has relatively little on their extractor design. More information on the extractor tip operation can be found in Ref. [8]. The use of high-index concentrators with TIR reflection conditions (no losses from reflection that occur on metal surfaces) minimize optical losses in the system and increases the concentration limit by n^2. In general, the use of a higher-index material to form the aperture of a gas turbine will not give higher operating temperatures. This is because the reradiation is proportional to n^2 times the area of the secondary exit aperture, which normally cancels the n^2 gain in concentration obtained by using higher-index materials. All gas turbines operate at temperatures much less than the sun's (6000 K) and therefore, the spectrum of reradiation is considerably red-shifted into the infrared (IR) region. By applying selective coatings onto the dielectric concentrator and extractor optics, one may prevent IR radiation from reradiating and it may be possible to reduce heat losses inside high-temperature electrical generation systems. Research is in progress at the Weizmann Institute of Science in Israel [9]to study the coupling of dielectric secondaries with gas turbines. They use a 6-sided extractor tip to let light out of the high-index secondary and into the high-temperature gas environment of the turbine (n=1).This is another possible use of high-index secondaries other than end-pumping solid-state laser crystals.

Using Highly Concentrated Sunlight in Space The techniques of nonimaging optics are particularly valuable in space or lunar environments where the use of solar thermal energy has obvious advantages. Earlier preliminary studies have explored this concept for the production of cement from lunar regolith and for solar thermal propulsion in space. For example, extremely high temperatures, in the range 1700 to 1900°C, are necessary for the production of cement from lunar minerals. Such temperatures, will in turn require very high levels of solar flux concentration. Energy budgets for the support of permanent manned operations on the lunar surface are expected to be limited. For high temperature thermal (i.e. >300° C) end uses, direct solar energy has obvious advantages over most other practical power sources. Conventional combustion processes are clearly impractical and conversion of electricity (either solar or nuclear generated) to high temperature heat represents a very wasteful use of high quality energy. On the other hand solar radiation is abundant and non-depletable. Most importantly, it is readily converted to heat with high efficiency, although at high temperatures this requires high concentration as will be discussed below.

Earlier work illustrated the feasibility of some particular two-stage configurations and indicated that the corresponding solar thermal conversion efficiency can be about 2.5 times that of the corresponding conventional design at 1500°C. A preliminary design configuration for such a high flux nonimaging solar concentrating furnace for lunar applications was proposed . It employs a tracking heliostat, a fixed, off-axis, two-stage concentrator with a long focal length utilizing a nonimaging trumpet or CPC type secondary deployed in the focal zone of the primary. An analysis of the benefits associated with this configuration employed as a solar furnace in the lunar environment shows that the thermal conversion efficiency can be about 3 to 5 times that of the corresponding conventional design at 2000°C. Furthermore this configuration allows several other advantageous practical design options in addition to high performance. For instance, the furnace itself and associated support structure and equipment need not shade the primary collecting aperture and spherical or faceted primaries may be able to be used.

Solar thermal propulsion systems in space will require very high temperatures to generate necessary levels of thrust by the direct solar heating and resulting expansion and expulsion of the propellent material. The generation of such temperatures, in the range 1400°C to 2200°C, will in turn require very high levels of solar flux concentration. In practice, to attain such levels it may be useful and perhaps even necessary to incorporate some form of ideal or near ideal nonimaging concentrator.

DISCUSSION

Applications of the techniques of nonimaging optics in the design of solar thermal concentrators has resulted in significant performance improvements in conventional short focal length dish concentrators, dramatic changes in the approach to solar furnace design, and the establishment of new solar flux concentration records in air and in refractive media. These accomplishments in just over ten years of active experimental development are indications of what can be achieved with these powerful methods. The wide variety and range of applications make this a very exciting field. It extends the bounds of solar energy research into solar manufacturing and even moves us in and into outer space.

ACKNOWLEDGMENTS: This work was supported in part by the U.S. DOE under Grant DEFG02-87ER-13726, under NREL subcontract Nos. XK-4-04070-03 and XX-6-06019-02 and by Sandia National Laboratories.

List of References

1) Cooke, D., Gleckman, P., Krebs, H., O'Gallagher, J., Sagie, D., and R. Winston, 'Brighter than the Sun", *Nature* **346**, 802, 1990.

2) Welford, W. and R. Winston (1989), High Collection Nonimaging Optics, Academic Press, New York.

3) Welford, W. and R. Winston (1980), "Design of Nonimaging Concentrators as Second Stages in Tandem with Image Forming First-Stage Concentrators," *Appl. Opt.* **19**(3), 347-351.

4) O'Gallagher, J., Winston, R., and A. Lewandowski (1993), "The Development of Two-stage Nonimaging Concentrators for Solar Thermal Applications,"Proceedings of the American Solar Energy Society Annual Conference, Washington, D.C, 203-209.

5) Ning, Xiaohui, O'Gallagher, J. and R. Winston (1987), "The Optics of Two-Stage Photovoltaic Concentrators with Dielectric Second Stages," *Applied Optics,* **26**, 1207.

6) Lewandowski A., Bingham, C., O'Gallagher, J., Winston, R., and D. Sagie (1991), "Performance Characterization of the SERI High Flux Solar Furnace," *Solar Energy Materials* **24**, 550-563.

7) Jenkins, D., Winston, R., Bliss, J., O'Gallagher, J., Lewandowski, A. and C. Bingham (1996a), "Solar Concentration of 50,000 Achieved with Output Power Approaching 1 kW," submitted to *J. Sol. Energy Eng.*

8) Karni, J., Ries, H., Segal, A., Krupkin, V., and A. Yogev (1994), "Delivery of Radiation for a Transparent Medium," Israel Patent 109,366 and international patent application PCT/US95/04915.

9) Karni, J., Ries, H., Segal, A., Krupkin, V., and A. Yogev (1995), "The DIAPR: A High-Pressure, High-Temperature Solar Receiver," Intl. Solar Energy Conf., Hawaii, 591-596.

10) Jenkins, D. and R. Winston (1996), "Integral Design Method of Nonimaging Optics," accepted for publication in *J. Opt. Soc. Am. A.*

11) Krupkin, V., Kagan, Y., and A. Yogev (1993), "Non Imaging Optics and Solar Laser pumping at the Weizmann Institute,"Nonimaging Optics: Maximum Efficiency Light Transfer II, Proceedings SPIE 2016, 50-60.

LOCALIZATION OF ENERGY ON THE MOLECULAR SCALE

Katja Lindenberg and David W. Brown

Department of Chemistry and Biochemistry and Institute for Nonlinear Science
University of California, San Diego
La Jolla, CA 92093-0340, U.S.A.

ABSTRACT

We discuss the spontaneous localization of vibrational energy in translationally invariant anharmonic chains at finite temperatures. In addition to the familiar energy-driven coherent mechanisms, which are rapidly degraded by thermal fluctuations, we identify the entropy-driven phenomenon we call "stochastic localization", within which we include a number of characteristics of soft anharmonic oscillators in thermal equilibrium. Principal among these are a tendency for soft oscillators to spend more time at higher energies than comparable harmonic oscillators, and for high-energy fluctuations in soft oscillators to persist for longer times than lower-energy fluctuations, leading to a tendency for energy fluctuations to be organized into "bursts" separated by intervals of relative quiet. We illustrate the effects of stochastic localization on a bistable impurity embedded in a chain of soft oscillators by comparing it to an impurity embedded in a harmonic chain. Effects on transition rates at a given system energy can be quite dramatic.

The spontaneous localization of mobile vibrational energy in molecular materials as an energy-focusing mechanism has long been an intriguing idea. This possibility has arisen, for example, in the context of materials in which such localization and/or transport may in turn lead to switching and other thresholding phenomena, detonation, chemical reactions, or local melting and other deformational effects. The mechanisms whereby energy in low-frequency modes can be transferred to higher energy modes (e.g. from vibrational to electronic) pose many interesting questions. For example, consider the events leading up to the regime where chemistry begins following a mechanical shock. Because the principal effect of an impinging shock is to promptly compress a material passing under it, energy is first delivered into low-lying vibrational modes associated with volumetric compression; these include the acoustic modes associated with translations of the center of mass of a unit cell, librational modes associated with rigid-body rotations of principal unit cell constituents, and the lowest-lying optical modes associated with relative translations of unit cell constituents. On the other hand, chemical reactions are associated with deformations of individual molecules and with their electronic excitations. Therefore, understanding such chemical reaction problems leads us to focus on the dynamics of their low-lying vibrational mechanical precursors and their coupling to higher energy modes.

Spontaneous energy localization in molecular materials may arise in qualitatively different ways. The most thoroughly studied energy localizing mechanism is static disorder through compositional variability or through the presence of defects. This mechanism is ubiquitous and quite well understood, and is *not* the subject of our work. While disorder promotes energy localization, it inhibits energy transport. We are interested in localization mechanisms that can coexist, at least in principle, with transport mechanisms. Such energy-localizing mechanisms arise even in translationally invariant aggregates (e.g., even nearly defect-free energetic materials are detonable) and are made possible by *anharmonicities*. The very localization of energy

in turn may drive the system (at least locally) into even more nonlinear regimes.

In the absence of thermal fluctuations a wide variety of translationally invariant anharmonic systems are capable of supporting long-lived, localized excitations under appropriate conditions. In this characterization we include not only completely integrable systems that support robust, spatially coherent, infinitely-long-lived solitons and solitary waves, but also a number of systems that support localized excitations for experimentally relevant times. Although many such systems have been identified in principle, the experimental observation of molecular solitons has been less successful. There are several reasons for this difficulty: (1) Few practical situations approach a degree of idealization consistent with the existence of infinite-lived or even very long-lived excitations such as solitons. A particular challenge in this regard is arriving at a practical and quantitative characterization of a "meaningful lifetime." (2) It may be difficult to create initial conditions that lead to soliton-like behavior; this particular statement will be illustrated below. (3) Most experimental techniques available for probing soliton structure are indirect and nondiscriminating in the sense that bulk properties of materials are probed; thus, most measurements average signals over substantial volumes of space thus simultaneously including coherent and incoherent motions and presenting serious challenges to deconvolution.

Thermal fluctuations tend to degrade the careful *energy-driven* balance that leads to coherent localization. On the other hand, we have recently identified an *entropy-driven* localization mechanism that we have called *stochastic localization* [1]. This mechanism, which occurs in soft anharmonic oscillators, is more robust than energy-driven localization. Below we describe the origin of this localization and some of its signatures. Its practical observation and/or importance remain to be ascertained.

Consider a chain of N oscillators labeled $n = 1, 2, \cdots, N$ with periodic boundary conditions ($N+1 \equiv 1$). Each oscillator is subject to an on-site potential $V(x_n)$ and to harmonic nearest-neighbor interactions. The Hamiltonian of the system is

$$H\{\dot{x}, x\} \;=\; \sum_n \;\left[\; \frac{1}{2}\dot{x}_n^2 \;+\; \frac{1}{2}c^2(x_n - x_{n-1})^2 \;+\; V(x_n)\right] \;. \tag{1}$$

We consider two types of chains. In one the on-site potential is harmonic with unit frequency, $V(x) = x^2/2$. This we call the *harmonic chain.* In the other the on-site potential is soft anharmonic,

$$V(x) \;=\; \frac{1}{2}\frac{x^2 \;+\; \alpha x^4}{1 + x^2}, \tag{2}$$

with $\alpha < 1$. This we call the *anharmonic chain.* Note that the harmonic chain is subsumed in Eq. (2) if we set $\alpha = 1$. If we set $\alpha = 0$ the on-site potential saturates and the oscillators can "run away." We will therefore always take $\alpha > 0$.

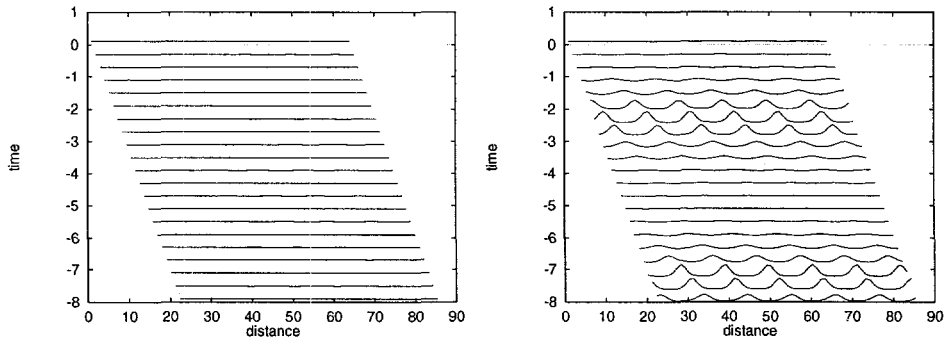


Figure 1: Energy density profile for harmonic (left panel) and anharmonic (right panel) chains with modulated initial condition.

The harmonic chain has N normal modes whose behavior is clearly understood. The anharmonic chain, on the other hand, does not have normal modes and in fact supports soliton and other solitary wave solu-

tions. An example that illustrates this behavior is shown in Fig. 1. The figure shows the result of initializing each chain of 64 oscillators with a nearly-uniform initial energy distribution upon which is imposed a weak spatially-periodic modulation. The left panel is the harmonic chain, while the right panel shows the anharmonic chain. The chains have been "opened" for graphing purposes. The vertical axis in each case indicates the local energy density, the horizontal axis indicates distance, and successive offsets indicate the progression of time. The modulation in the harmonic chain persists in time, but is hardly visible on the scale shown. On the other hand, the modulation in the soft anharmonic chain grows periodically in time, transforming the original nearly-uniform energy distribution into a train of pulses, each of which spans a number of unit cells within which a several-fold amplification of the original energy density is achieved for a non-trivial interval of time [2]. The rather coherent concentration of energy achieves local deformations that may assist in the initiation of chemistry [3, 4]. The dramatic focusing in this particular case is a transient phenomenon ("instability") that is eventually degraded.

We are ultimately interested not only in the performance of such a medium at zero temperature, where nonlinear collective excitations that lead to energy localization are often stable (at least in some parameter regimes), but also at finite temperatures, where such coherent collective excitations must compete with thermal fluctuations.

We have recently identified an energy-localizing mechanism that is *driven* by thermal fluctuations, that is, it is entropy driven rather than energy driven [1]. This mechanism, which we call *stochastic localization*, is even more general and ubiquitous than other energy localization mechanisms that have been proposed for regular molecular aggregates. Stochastic localization does not require spatial coherence and is robust in the face of thermal effects — indeed, it can be argued to be in competition with forces that foster spatial organization.

In order to describe stochastic localization, consider first a *single* oscillator described by the Hamiltonian $H\{\dot{x}, x\} = \frac{1}{2}\dot{x}^2 + V(x)$, in which $V(x)$ is the on-site potential introduced earlier. When such an oscillator is in contact with a thermal bath at temperature T, a number of simple thermodynamic arguments applied to an ensemble of such independent oscillator immediately lead to the following conclusions:

1. The average energy of a soft anharmonic oscillator is higher than that of a harmonic oscillator at a given temperature: they share the same average kinetic energy at a given temperature, but the soft oscillator has greater excursions and therefore a higher potential energy.

2. The energy fluctuations in a soft anharmonic oscillator are larger than those of a harmonic oscillator at the same temperature. This conclusion is easily reached from a simple analysis of the specific heat which is directly connected to these fluctuations.

3. In an ensemble of harmonic oscillators (or in a single oscillator over time) half of the total energy on average is in kinetic form and half in potential form. In an ensemble of soft oscillators, on the other hand, on average a greater fraction of the total energy is in potential form.

4. The previous item immediately leads to the conclusion that a high energy fluctuation in a soft anharmonic oscillator persists for a longer time than does an equal fluctuation in a harmonic oscillator. This is due to the fact that the rate of energy dissipation is directly proportional to the *kinetic* energy.

All of the foregoing considerations point to the notion of statistical localization, that is, *a spatial and temporal localization of energy in anharmonic oscillators.* This localization is entropy-driven and incoherent: it comes about because the density of states of a soft anharmonic oscillator increases with energy so that it is entropically favorable for the system to concentrate higher energies in some members of the ensemble (and lower energies in others) than it would in an ensemble of harmonic oscillators, where the favorable distribution is more uniform. Furthermore, high-energy fluctuations survive for a longer time in the anharmonic chain.

What happens when oscillators are connected as a network such as described by the Hamiltonian (1)? The effect of the nearest neighbor interactions is to drive neighboring oscillators toward the same phase and thus to establish coherence on the smallest spatial scale. Dynamically, this local drive toward spatial coherence has the effect of dispersing any localized energy distribution, since such a distribution constitutes a *deviation* from perfect alignment. The main dynamical consequence of building spatial coherence in vibrational systems is *dispersion.* Thus at zero temperature two opposing tendencies coexist: that of the soft anharmonicity to *increase inhomogeneity*, and that of the coupling to lead to the *dispersion* of energy

and hence to a *decrease in inhomogeneity.* Optimal solutions achieve some degree of balance between these opposing tendencies, often giving rise to solitary wave solutions and their idealization in the concept of the soliton [2]. These optimal solutions may or may not survive the effects of thermal fluctuations and dissipation. On the other hand, thermal fluctuations bring with them the complementary effect of stochastic localization. This localization mechanism becomes stronger with increasing thermal fluctuations, but the coupling between oscillators acts to further redistribute energy from a "hot" oscillator to its neighbors. It is difficult to know what the final balance might be. Although both the coherent and the incoherent localization mechanisms are a direct consequence of the (softening) anharmonic character of the oscillators, they are in a sense opposite to one another and also to the dispersive action of increased coupling.

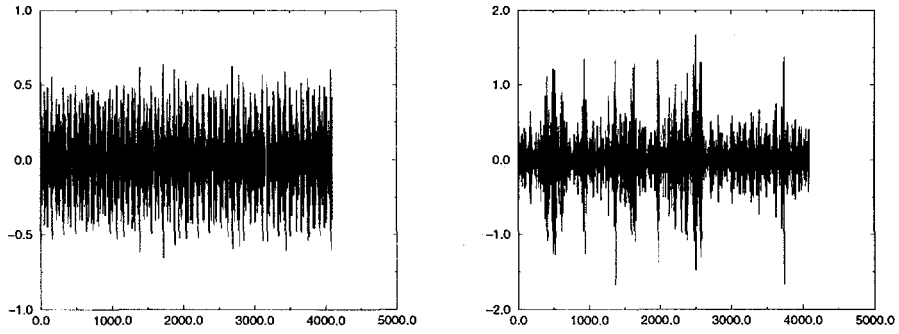


Figure 2: Position of oscillator 55 as a function of time in the harmonic chain (left panel) and the anharmonic chain (right panel) of Figure 1. Note the scale differences.

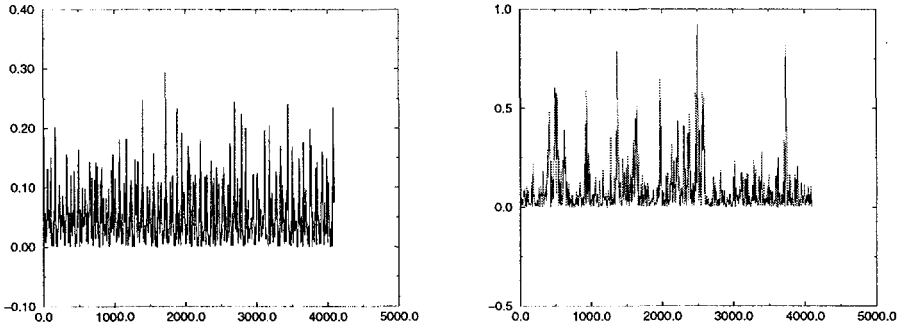


Figure 3: Energy of oscillator 55 as a function of time in the harmonic chain (left panel) and the anharmonic chain (right panel) of Figure 1. Note the scale differences.

At this time most of our available information comes from simulations. We have carried out simulations on harmonic and anharmonic chains (the latter with $\alpha = 0.05$). The value of the dispersion parameter c plays a central role in determining the nature of the ultimate self-consistent state in the absence of thermal fluctuations; large values tend to result in broad solitary waves spanning many lattice sites, and small values

tend to result in compact localized vibrations that may be "pinned" to particular lattice sites [5]. In our work we have mostly used the value $c = 0.5$ (this is also the value used to generate Fig. 1), which besides being "intermediate" (the allowed range is $c \in (0, 1)$), is motivated by consideration of the actual dispersion relations of low-lying librations in certain organic molecular crystals [6]. The values of c typically chosen for the study of anharmonic arrays designed to exhibit nonlinear behavior are usually much smaller, of the order of 10^{-2} [7].

In Fig. 1 we showed the evolution of chains of 64 coupled oscillators with an initial condition that led to clear coherent focusing in the anharmonic chain. In Figs. 2 and 3 we show for exactly the same systems a number of other features of these chains. The first panel in Fig. 2 shows the trajectory of one of the oscillators (number 55) in the chain as a function of time, and the second panel shows the trajectory of this oscillator in the anharmonic chain. Although the quality of coherence is a bit less clear than in Fig. 1, the differences are evident. They are also evident in Fig. 3, which shows the energy of this oscillator in the chains as a function of time.

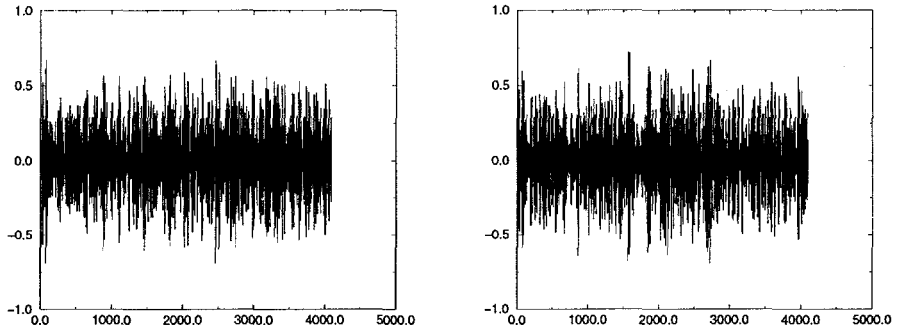


Figure 4: Position of oscillator 55 as a function of time in the harmonic chain (left panel) and the anharmonic chain (right panel). The chains have the same initial energy as in Figure 2 (0.1 per oscillator) but distributed randomly.

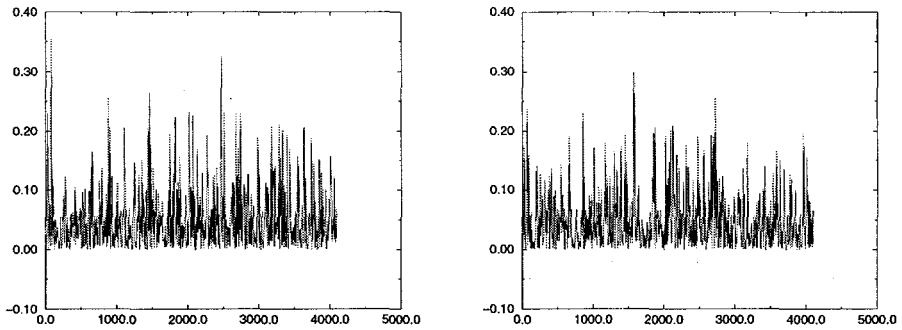


Figure 5: Energy of oscillator 55 as a function of time in the harmonic chain (left panel) and the anharmonic chain (right panel). The chains have the same initial energy as in Figure 3 but distributed randomly.

The importance of the initial condition is evident when we compare these figures to those shown in Figs. 4 and 5. In these four panels we show exactly the same oscillator in exactly the same chains with exactly the same total energy, but now the energy is initially randomly distributed. Although this is a microcanonical (fixed energy) system, in this set of figures the systems behave essentially as they would in thermal equilibrium. Now the energy is distributed quite uniformly in both chains, and there is no evidence of coherent energy focusing. The initial energy is too low to show the effects of stochastic localization - the average energy per oscillator is 0.1, which places it near the bottom of the on-site potentials where both potentials are very similar. Again, we stress the importance (and difficulty) of creating the proper initial condition in order to observe the effects of coherent localization.

How might one observe stochastic localization and/or the way it may coexist with coherent localization mechanisms? A number of possible signatures come to mind. One is to connect the chains to a heat bath of sufficiently high temperature, so that energies where the harmonic and anharmonic potentials differ substantially come into play. The two panels in Fig. 6 show such a case. It is difficult without further analysis to conclude whether or not there is a substantial difference in the texture of these two chain histories – the sharp persistent ridge in the anharmonic chain differs from those in the harmonic chain in a number of ways, including the fact that the anharmonic ridge is clearly mobile and keeps its integrity during the entire time history shown in the figure.

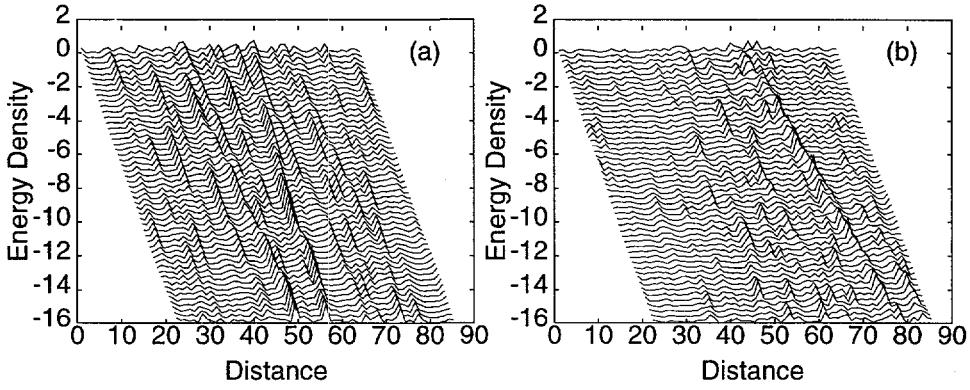


Figure 6: Time dependence of the energy density in a chain of oscillators evolving according to (1) with $c = 0.5$. The chain is in contact (via appropriate Langevin terms) with a heat bath at temperature 0.5. Vertical axes represent energy, horizontal axes represent distance along the chain of 64 sites. Time increases from top to bottom. Left panel: harmonic chain; right panel: anharmonic chain.

Another probe of stochastic localization, more directly related to our goal of identifying localization mechanisms that might trigger other events, is to place a bistable "impurity" in the chain. In this case our chain consists of 64 oscillators (harmonic or anharmonic) connected as before and one bistable on-site potential of the form

$$V(x) \;=\; V_0 \left(\frac{x^4}{4} - \frac{x^2}{2} \right) \;, \tag{3}$$

also connected via harmonic springs to its neighbors (we label this particle number 65 and connect it to its neighbors, oscillators 1 and 64, by the same harmonic springs as connect the other oscillators). If we think of this bistable system as a configuration coordinate, then one of the wells might represent one configuration (or "reactant") and the other another configuration (or "product"). The characteristic frequencies of the bistable impurity are chosen to fall within the band of frequencies of the chain so that any energy localization cannot be ascribed to an energy mismatch as occurs in many disordered systems. In our simulations we take $V_0 = 0.75$. If localization is observed, it is certainly due to the anharmonicity of the chain; if it occurs at sufficiently high "temperature," then it is due to stochastic localization rather than coherent localization.

In Fig. 7 we show a typical trajectory of the bistable system in a harmonic chain (left panel) and anharmonic chain (right panel) with a random initial distribution of energies (microcanonical system). The average energy per oscillator is 0.23. We point to the following difference in texture: in the harmonic panel the impurity tends to oscillate in one of the bistable wells (either around -1 or around 1) with single passages

from one to the other. If the impurity gains enough energy to make the transition, it loses it quickly, remains in the other well for a while, and then gains sufficient energy to cross again. In the anharmonic panel, on the other hand, if the bistable particle gains enough energy it tends to *retain* it for a long time, thus oscillating back and forth at a high energy above the barrier separating the wells. Once this energy is lost, the bistable particle remains rather cold for a while, that is, its oscillations have a smaller amplitude than those of the harmonic system. This is the signature of stochastic localization: the anharmonic chain causes the impurity to experience persistent hot and cold periods relative to the harmonic chain. In Fig. 8 we show the phase space portrait of the bistable system in the two cases. The right panel has a greater concentration of both cold and hot trajectories than the left. In Fig. 9 we show another example of the trajectories of the bistable impurity; here the energy per oscillator is 0.23. Clearly, the transition rate statistics in the two chains are quite different.

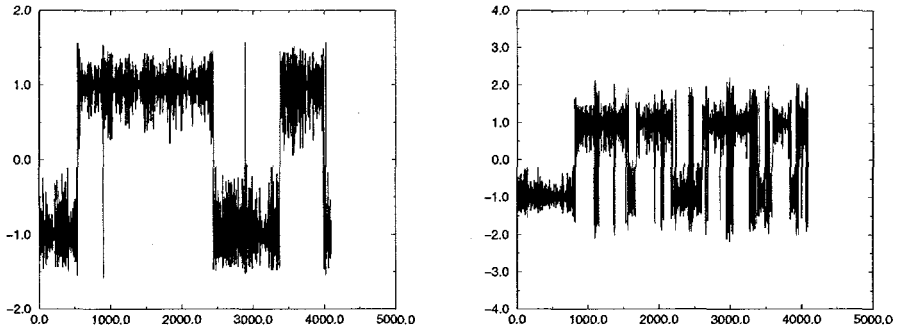


Figure 7: Time dependence of the trajectory of a bistable impurity in a chain, with initial energy 0.23 per site. Left panel: harmonic chain; right panel: anharmonic chain. Note the scale differences.

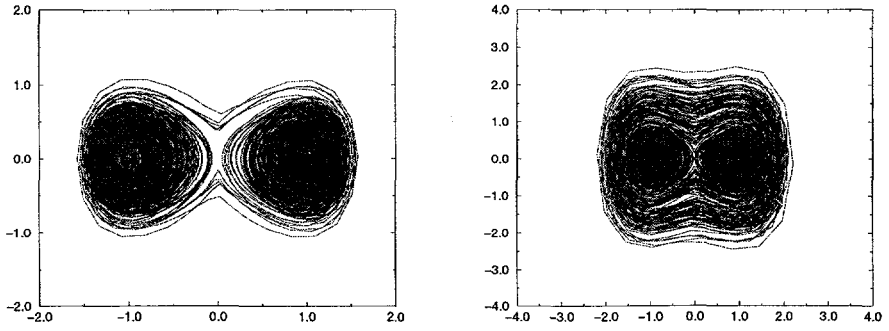


Figure 8: Phase space portrait of a bistable impurity in a chain, with initial energy 0.23 per site. Left panel: harmonic chain; right panel: anharmonic chain. Note the scale differences.

We have compared and contrasted two mechanisms for energy localization in translationally invariant anharmonic chains. The more familiar one is energy driven and leads to coherent localization in the form of solitons, solitary waves, and other such coherent structures. These structures are sensitive to (and usually disappear in the presence of) thermal fluctuations. The other, more robust, mechanism of energy localization in (soft) anharmonic chains relies on the presence of thermal fluctuations and is entropy-driven. We

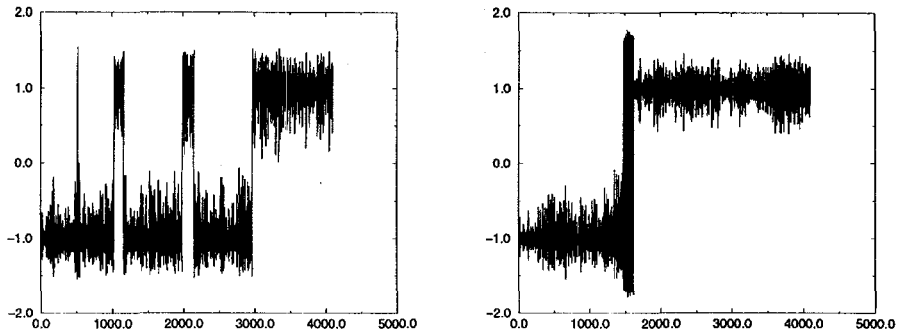


Figure 9: Time dependence of the trajectory of a bistable impurity in a chain, with initial energy 0.20 per site. Left panel: harmonic chain; right panel: anharmonic chain.

discussed the subtleties involved in observing and discriminating between these mechanisms. We continue to develop these ideas in classical systems and have made some preliminary progress in quantum systems, where translational invariance places even greater burdens on the stationary states of the system [8].

ACKNOWLEDGEMENT

This work was performed under the auspices of the U. S. Department of Energy. Portions were carried out in collaboration with Lisa Bernstein and with Jose M. Sancho.

REFERENCES

[1] D. W. BROWN, L. BERNSTEIN, and K. LINDENBERG, "Stochastic Localization," *Phys. Rev. E* **54**, 3352 (1996).

[2] G. B. WHITHAM, *Linear and Nonlinear Waves*, John Wiley & Sons, New York (1974); R. K. DODD, J. C. EILBECK, J. D. GIBBON, and H. C. MORRIS, *Solitons and Nonlinear Wave Equations*, Academic Press, New York (1982); A. NEWELL, *Solitons in Mathematics and Physics*, SIAM, Philadelphia (1985).

[3] R. A. MARCUS and N. SUTIN, *Biochim. et Biophys. Acta* **811**, 265 (1985), and references therein.

[4] L. SANCHE, *IEEE Trans. Elec. Insulation* **28**, 789 (1993), and references therein. lattice sites

[5] R. S. MACKAY and S. AUBRY, "Proof of existence of breathers for time-reversible or Hamiltonian networks of weakly coupled oscillators," *Nonlinearity* **7**, 1623 (1994).

[6] A. M. MICU, D. DURAND, M. QUILICHINI, M. J. FIELD, and J. C. SMITH, "Collective vibration in crystalline l-alanine," *J. Phys. Chem.* **99**, 5645 (1995).

[7] D. HENNIG, K. O. RASMUSSEN, and G. P. TSIRONIS, "Breatherlike impurity modes in discrete nonlinear lattices," *Phys. Rev. E* **52**, R4628 (1995).

[8] D. W. BROWN and K. LINDENBERG, "Looking for soliton signatures in polaron Bloch states," to appear in *Physica D* (1997).

QUANTITATIVE PHOTOGRAPHY OF INTERMITTENCY IN SURFACE WAVE TURBULENCE

W.Wright, R. Budakian, S.J. Putterman

Physics Department, University of California, Los Angeles, CA 90095

ABSTRACT

At high amplitudes of excitation surface waves on water distribute their energy according to a Kolmogorov type of turbulent power spectrum. We have used diffusing light photography to measure the power spectrum and to quantify the presence of large structures in the turbulent state.

INTRODUCTION

One of the key problems of science and engineering is to understand the fate of energy which is injected into a fluid so as to drive it very far from equilibrium. As the energy flows through phase space it experiences a competition between randomization and structure formation which are issues that come into play in the study of turbulence. The properties of turbulent flow are of importance in phenomena ranging from turbomachinery to airplane design, to weather prediction. A good perspective on the problem of turbulence is provided by the following three quotes:

> "Big whorls have little whorls
> Which feed on their velocity;
> And little whorls have lesser whorls,
> And so on to viscosity
> (in the molecular sense)"

> "The wind comes in gusts."

"I am an old man now and when I die and go to heaven there are two matters on which I hope for enlightenment. One is quantum electrodynamics and the other is the turbulent motion of fluids. And about the former I am really rather optimistic."

The first quote is due to Richardson (1926)[1] and expresses the concept of the turbulent cascade. Energy which enters a fluid at long wavelength scatters from itself to generate motion on shorter and shorter wavelengths. This is the way in which turbulent

nonlinear interactions create a randomization of the input energy. However, turbulence is not so simple. Turbulence is not merely scaled up thermal noise. In the midst of the random motion there can appear unexpected concentrations of energy, which are typically surfaces on which the dissipation is concentrated[2]. This "intermittency" is the basis for the second quote which has been attributed to Landau. The difficulties connected with the interplay between intermittency and the cascade are pinpointed by the third quote due to Sir Horace Lamb[3].

Looking back at this prophetic remark we can note the mathematical structure that has been developed to explain 'qed' and marvel at the lack of progress in defining, probing, and solving the problem of turbulence. This situation persists despite the huge interest in 'chaos' and the experimental opportunities openned up by modern technology. We propose that the continuing validity of Lamb's quote is due in part to the inability of experimentalists to provide quantitative measurements of the wavenumber or 'k' spectrum of turbulent motion. Only from the 'k' spectrum can one diagnose the complex modal structure of high amplitude motion. As we will demonstrate below such measurements provide a quantification of intermittency and randomness in turbulence and they form the basis for advances in the analytic theory.

The physical system which we have used to gain insight into turbulence consists of capillary waves that run around on the surface of a fluid. At low amplitude the principle of superposition applies and the waves run through each othe. But at high amplitudes nonlinear terms cause the waves to scatter. This exchange of energy eventually leads to wave turbulence.

EXPERIMENT

Figure 1 shows the apparatus that we have used to excite and measure the large amplitude distortions of a fluid surface[4]. A vibration exciter oscillates a container of fluid in the vertical direction at sufficient amplitude that instabilities determined by the Matthieu equation come into play. The resulting surface motion is made visible by suspending into the water a .04% solution of polyballs. This concentration is sufficiently dense that light traveling through the water is so strongly scattered that it diffuses. A charge coupled device (CCD) records the light to exit the fluid. Typically the surface is broken up into one million pixels [1024x1024] where each pixel is capable of recording a dynamic range of 65,000 gray scales [or 16 bits]. This image is converted into the surface height with the help of a calibrated measurement of the amount of light to exit the surface as a function of fluid depth. The deeper the fluid the smaller is the amount of light to make it to the surface at that location.

At low amplitudes of excitation the motion is regular and exhibits the square wave patterns so characteristic of parametrically excited ripples[5]. At high amplitudes there is a transition to stochastic motion characteristic of the turbulent state that we wish to understand. Renderings of the data for low and high amplitude motion are shown in Figure 2. The broad band power spectrum of high amplitude motion [converted to a function of ω] is shown in Figure 3b. The slope of the observed spectrum [on a log-log plot] is close to that predicted by Kolmogorov scaling laws [6].

Theories which are based upon a random phase approximation yield a power spectrum in agreement with the Kolmogorov dimensional analysis and of course also yield a description of turbulence that is devoid of intermittency. Since phase is a physical quantity of importance equal to power [or amplitude squared] we have used the instantaneous photos to investigate the phase coherence, if any, of the turbulent state. To do this we digitally filtered the photo to include only the contributions from a range of wavelengths. In this case those wavelengths corresponding to the frequency range 373Hz to 429Hz [as indicated on Figure

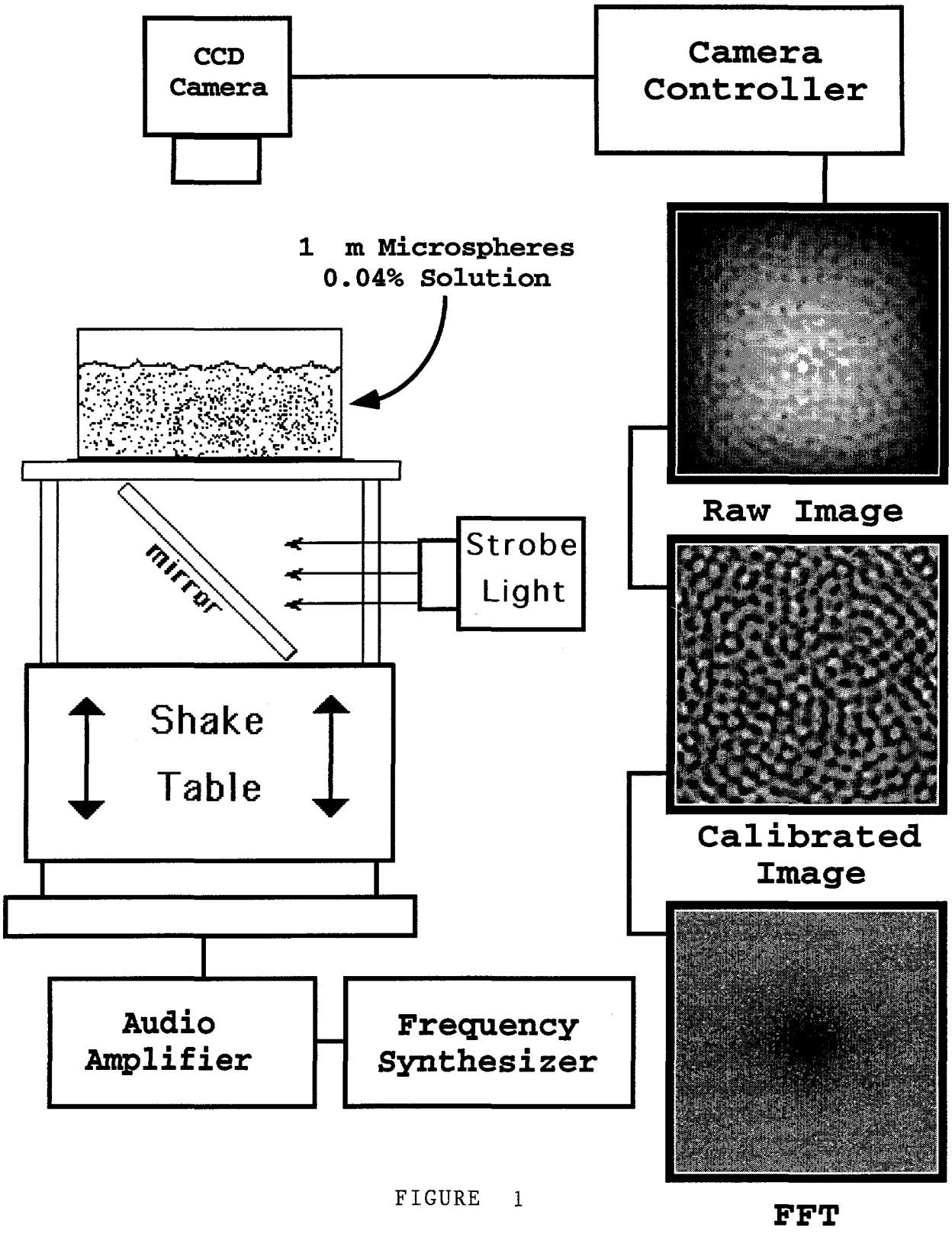

CCD Camera
Camera Controller
1 m Microspheres
0.04% Solution
Raw Image
mirror
Strobe Light
Shake Table
Calibrated Image
Audio Amplifier
Frequency Synthesizer
FFT

FIGURE 1

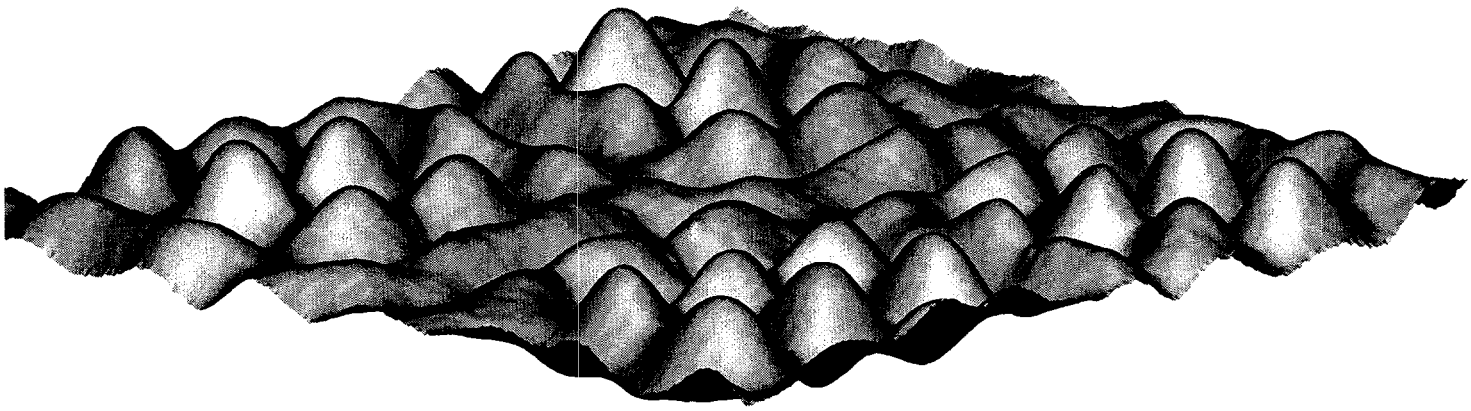
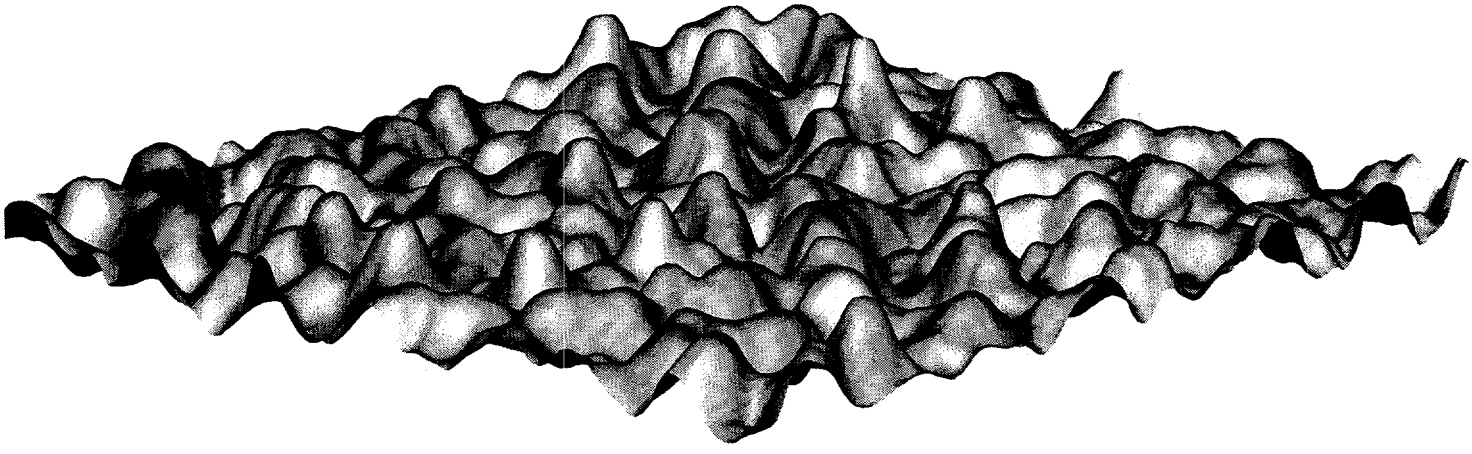
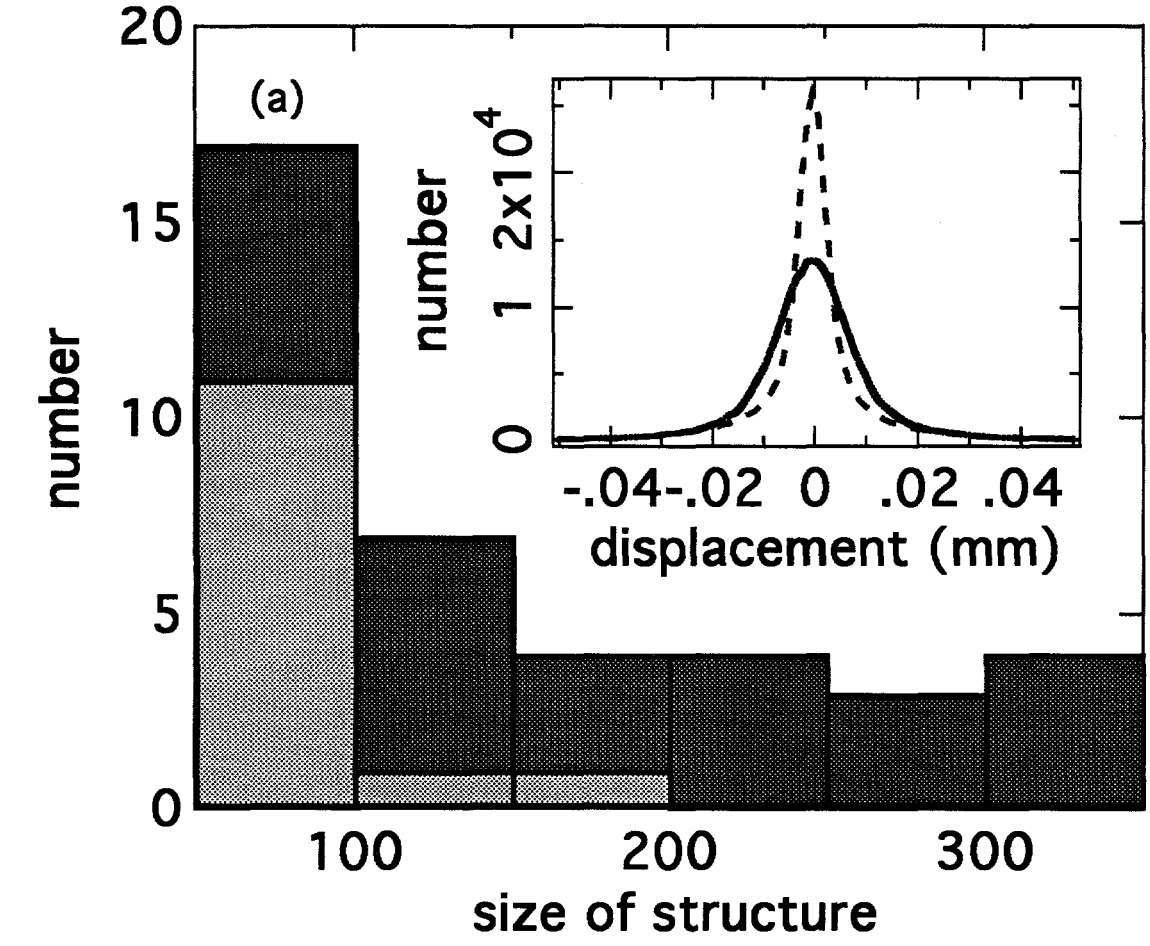


FIGURE 2




(a)
number
20
15
10
5
0
100
200
300
size of structure
number
0
1
2x10^4
-.04
-.02
0
.02
.04
displacement (mm)

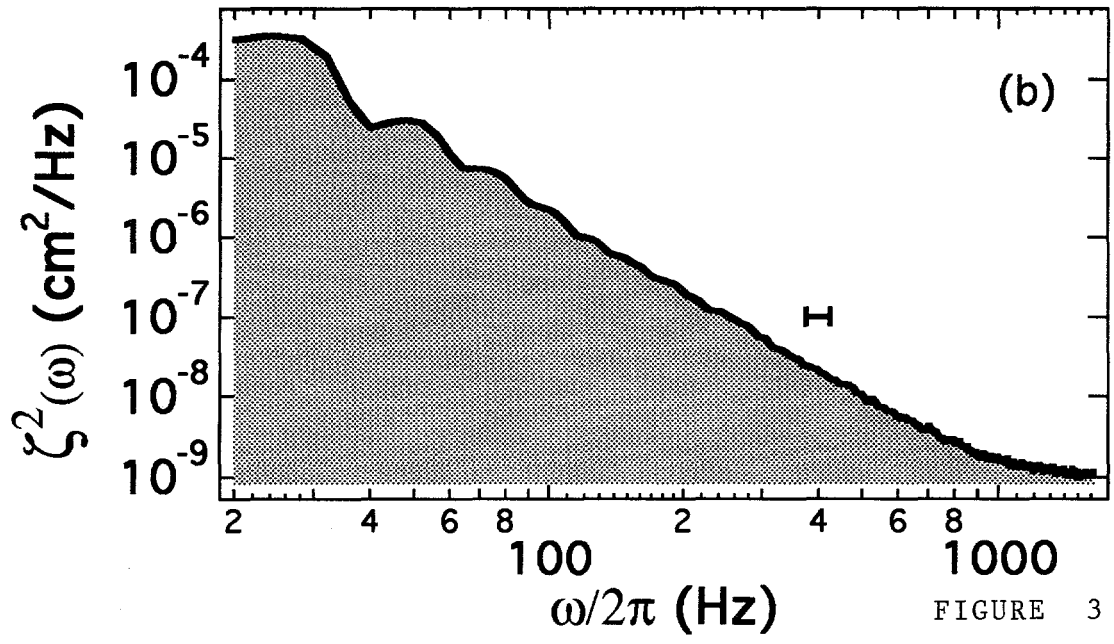

(b)
ζ²(ω) (cm²/Hz)
10^-4
10^-5
10^-6
10^-7
10^-8
10^-9
H
2
4
6
8
100
2
4
6
8
1000
ω/2π (Hz)

FIGURE 3

3b]. From this filtered photo we then constructed the dissipation function which is proportional to the square of the Laplacian of the surface height. A rendition of this data is shown in Figure 4. Note that the turbulence is localized onto surfaces. It is not uniformly spread out. According to these experiments the turbulent state is filled with structures or so-called "intermittency".

In order to obtain a quantitative handle on intermittency we have assigned a size to the structures according to the number of pixels that have common sides and exceed the rms power by at least a factor of 5. One now verifies that turbulent motion is characterized by large fluctuations as is shown in Figure 3a where the distribution of sizes of these structures is displayed with a dark shading. For data with randomized phases the [but same rms power] the structures are much smaller. The randomized data is displayed with a light shade. The inset to Figure 3a shows the distribution of amplitudes in a single filtered photo [solid line] as compared to the same power spectrum but with randomized phases [dashed line].

These measurements provide a quantitative basis for the development of a theory which unifies the insights that while turbulence involves a statistical description of many independent modes of fluid motion its power spectrum appears to be a finely balanced average of large fluctuations. The experimental challenges are to achieve a longer range of turbulence and resolve the time as well as spatial dependence of the structures that characterize turbulent motion.

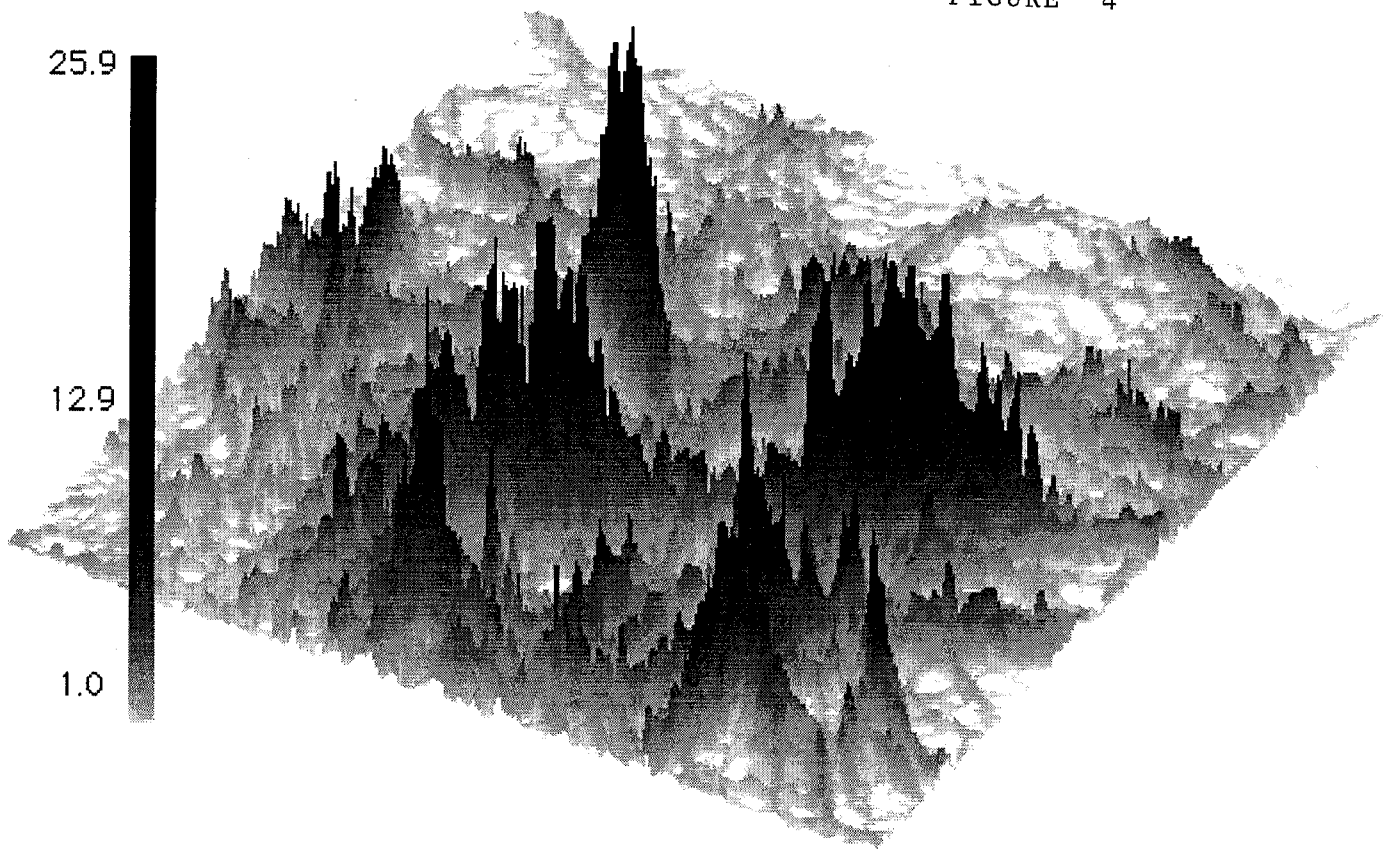
ACKNOWLEDGMENT

This research is supported by the U.S.D.O.E. Division of Engineering and Geophysics.

REFERENCES

1. L.F. Richardson, Weather Prediction by Numerical Process, Cambridge University Press, (1922); & Proc. Roy. Soc. A110, 709, (1926)
2. L.P. Landau, E.M. Lifshitz, Fluid Mechanics (Pergamon, N.Y.) 1955.
3. Quoted in K.R. Sreenivason, Nature 344, 192 (1990).
4. W. Wright, R. Budakian, S. Putterman, Physical Review Letters, 76, 4528 (96).
5. N.B. Tufillaro, R. Ramshankar, J.P. Gollub; Order-Disorder Transition in Capillary Waves: Phys. Rev. Lett. 62,422-425, (1989).
A.B. Ezersky, S.V. Kiyashko, P.A. Matusou, M.I. Rabinovich; Domain Walls and Dislocations in Capillary Ripples: Europhys. Lett. 26,183-188 (1994).
S.T. Milne; Square Patterns and Secondary Instabilities in Driven Capillary Waves: J. Fluid Mech. 225, 81-100, (1991). B. Christianson, P. Alstrom, M. T. Levinsen; Ordered Capillary Wave States: Quasicrystals, Hexagons and Radial Waves: Phys. Rev. Lett. 68, 2157-2160 (1992).
R.Blümel,I.H. Davidson,W.P.Reinhardt,H.Lin, M.Sharnoff; Quasilinear Ridge Structures in Water Surface Waves: Phys Rev A45 ,2641-2644, (1992).
B.J. Gluckmen, P. Marq, J.Bridges, J.P. Gollub; Time-Averaging of Chaotic Spatiotemporal Wave Patterns: Phys.Rev. Lett. 71,2034-2037 (1993).
E.Bosch, W.van de Water; Spatiotemporal Intermittency in the Faraday Experiment: Phys Rev Lett: 70, 3420-3423(1993).
6. V.E. Zakharov, V.S. L'vov, G. Falkovich; Kolmogorov Spectra of Turbulence: [Springer-Berlin] 1992.

FIGURE 4



IN-FLIGHT BEHAVIOR OF DISSIMILAR CO-INJECTED PARTICLES IN THE SPRAYING OF METAL-CERAMIC FUNCTIONALLY GRADIENT MATERIALS

J.R. Fincke, W.D. Swank, and D.C. Haggard

Optical and Plasma Physics
Idaho National Engineering Laboratory
Idaho Falls, ID 83415-2211

ABSTRACT

In the spraying of functionally gradient coatings the particle ensemble delivered to the substrate can vary from a relatively low melting point metallic particle to a significantly higher melting point ceramic particle. At various stages in the spray process the particle ensemble can be either predominantly metallic, ceramic, or an intermediate combination. For co-injected particles the mixtures do not behave as a simple linear superposition of the spray patterns of the individual particle types. The particle temperature, velocity, size distributions, and pattern characteristics of the resulting spray fields is examined for all ceramic particle sprays (ZrO_2), all metallic particle sprays (NiCrAlY), and for a 1:1 mixture. The major particle-particle interaction occurs in the injector itself and results in a modified spray pattern which is different from that of either material sprayed alone. The particle velocity distributions generally exhibit a bimodal nature which is dependent on the size and density of the injected particles.

INTRODUCTION

Functionally Gradient Materials (FGMs) having either continuously or stepwise varying compositions and/or microstructures offer solutions to engineering problems involving coating systems with large differences in the coefficient of thermal expansion (CTE). A classic example where large differences in the CTEs are particularly troublesome are ceramic thermal barrier coatings (TBCs) which typically consist of the ceramic material applied over a metallic bond coat which, in turn, is applied to the surface of a metallic substrate. By continuously grading the composition of the coating from that of the bond coat material at the structure/coating interface to that of the ceramic TBC at the outer surface the stresses caused by mismatches in CTE are lessened.

One of the major objectives in the fabrication of FGMs is that the final structure should vary in a regular and consistent manner. Deviations from the planned gradient design may result in significant deviations in the properties of the gradient, such as the CTE. These deviations can then lead to localized stress concentrations and ultimately to crack formation and failure of the coating. Most of the deviations in the compositional gradient can be traced back to non-optimized spray parameters, resulting in lower than anticipated deposition efficiency. In plasma spraying low deposition efficiency for one species can often lead to banding. Banding can also occur when the trajectories of the various particle species diverge significantly between their injection into the plasma and impact on the substrate. Therefore, it is important that the inter-relationships between the particle size distributions, injection orientation, feed rates, etc., and the resulting particle trajectories and temperatures, etc., are understood.

By its nature thermal plasma spraying is well suited to the fabrication of FGMs. There exist a number of thermal spray processing methods[1,2] suitable for producing FGMs. The simplest, from a hardware, reliability, and manufacturing point of view, is the use of a single torch, fed by two computer-controlled powder feeders and a single common injector. The powders are mixed on-the-fly in a tee configuration upstream of the injector. While this configuration requires that the particle size distributions be chosen such that the particle trajectories coincide at impact the advantage of simplicity and reliability of this configuration in a manufacturing operation should not be ignored. It is also generally observed that over a wide range of operating conditions the co-injection of premixed powder generally yields a more uniform distribution of species within the graded layers[2].

This study addresses the development of an understanding the behavior of ensembles of dissimilar co-injected particles. The particle types examined are a metallic particle (NiCrAlY) and a ceramic particle (ZrO_2). The major particle-particle interaction occurs in the injector itself and results in a modified spray pattern which is different from that of either material sprayed alone. The spatial and statistical distributions of particle size, velocity, and temperature are also modified.

EXPERIMENTAL PROCEDURE

All testing was conducted using a Metco 9MB spray gun and two Miller model 1270 powder feeders. The powder feeders and carrier gas injection are independently computer-controlled. The spray conditions are summarized in Table 1. The powders used were a Praxair NiCrAlY, Ni-346-1, and a Sultzer-Metco stabilized zirconia, 204NS. Both powders were roughly spherical as supplied. The size distributions, obtained by sieving, appear in Figures 1 and 2. All spraying was conducted in ambient laboratory air at 86 kPa (4800 ft above sea level). A wide range of NiCrAlY:ZrO_2 ratios by volume were examined, however, only 1:1 by volume data will be presented in detail. In performing the spraying the feed rate of the ZrO_2 was held constant at 1.2 kg/hr while the feed rate of the NiCrAlY was alternately set at 1, 2, and 4 kg/hr. The two powders were also sprayed individually to provide comparison data. All particle data is acquired at an axial standoff of 100 mm.

Table 1. Thermal Spray Parameters.

Gun Type	Metco 9MB
Primary Gas	Argon (40 slm)
Secondary Gas	Hydrogen (12 slm)
Gun Voltage	75 V
Gun Amperage	600 A
Thermal Efficiency	75 %
Carrier gas	6000 sccm
	(3000 sccm each feeder for mixtures)

The experimental data acquired includes particle trajectory and pattern, particle velocity and size, and particle ensemble temperature. Particle velocity, size, and temperature statistics were also obtained and the number of cold particles estimated. Particle ensemble temperature and trajectory information consisting of the spray pattern shape, centroid position and pattern width, along with gun operating parameters and efficiency were measured by an In-flight Ltd. Torch Diagnostic System, TDS-1610, and In-flight Particle Pyrometer, IPP-2000. Particle velocity and size were measured by an Aerometrics Phase Doppler Particle Analyzer (PDPA) system. The single particle temperature diagnostic, which was designed and fabricated in-house[3,4], is integrated with the PDPA. The PDPA system does not discriminate between hot and cold particles, and only particles which generate a Doppler burst are included in the temperature statistics. This serves to tightly locate the particles spatially and by keeping track of those particles which give size and velocity information but do not yield a temperature measurement gives an estimate of the number of cold particles present. The minimum measurable particle temperature with the current experimental configuration is approximately 1300°C for 40 μm particles and somewhat higher for smaller diameter particles.

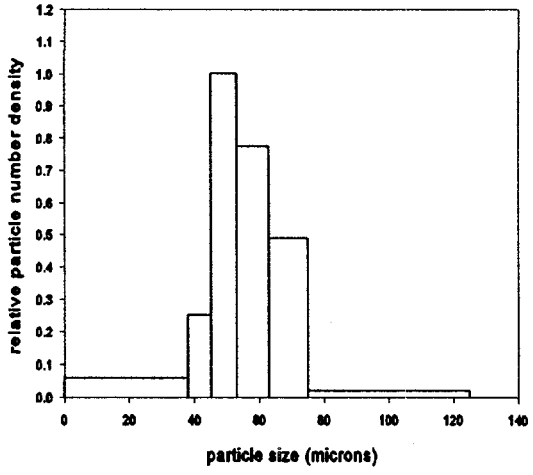


Figure 1. NiCrAlY particle size distribution obtained from sieve data.

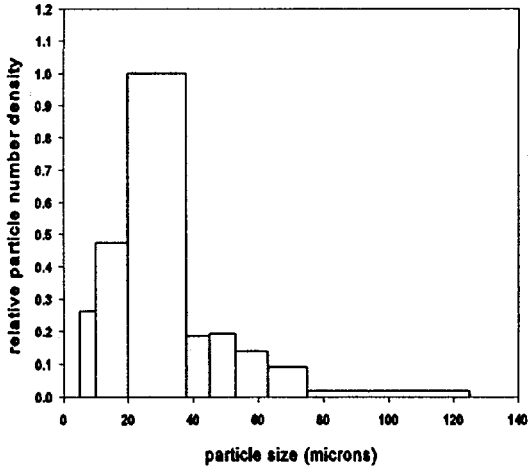


Figure 2. Zirconia particle size distribution obtained from sieve data.

RESULTS

The spray pattern shapes for the individual particle types and a 1:1 mixture (by volume) are shown in Figure 3. Also shown for reference is the sum of the two individual particle type curves. These data are obtained by imaging the light emitted by the hot particles onto the linear CCD array of a linescan camera, hence the data represent only the spatial distribution of hot particles. It is clear that the spatial distribution of the particles for the 1:1 mixture is not the simple linear super position of the individual particle types. This is because the injection velocity of the mixed system particle ensemble is different from that of either individual particle type. The average particle injection velocities and standard deviations appear in Table 2. The effect of particle-particle interactions in the close confines of the injector serves to increase the average velocity of the NiCrAlY particles while decreasing the average velocity of the ZrO_2 particles. It is interesting to note that even though the standard deviation of the injection velocity of the NiCrAlY particles is one-half that of the ZrO_2 particles the spray pattern width of the NiCrAlY powder alone is significantly broader that the zirconia only.

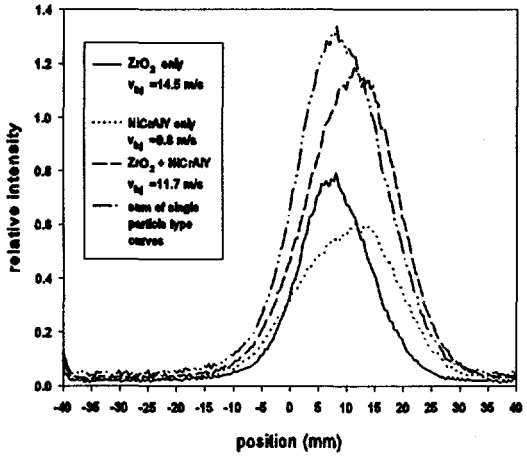


Figure 3 Spray pattern and trajectory obtained from linescan camera.

Table 2. Summary of particle injection velocities.

Particle	V_{inj} (m/s)	RMS (m/s)
ZrO_2	14.5	2.98
NiCrAlY	9.8	1.45
1:1 mix	11.7	2.40

The spatial distribution of average particle size and velocity at a standoff of 100 mm appears in Figures 4 and 5. Injection is from left-to-right. There is evidence of significant injection induced aerodynamic particle sizing, with the smallest particles being turned by the plasma jet nearer the point of injection. The larger particles with their higher momentum tend to congregate on the far side of the jet centerline. The average particle size data suggests that some aerodynamic separation of the individual particle types is also occurring, with the lighter and smaller ZrO_2 particles being preferentially closer to the point of injection and the larger, heavier NiCrAlY particles preferentially residing farther away.

The histograms of particle size on the centerline of the gun appear in Figures 6-8 for the ZrO_2, NiCrAlY, and 1:1 mix respectively. The corresponding particle velocity distributions are shown in Figures 9-11. It is interesting to note that even though the NiCrAlY particle size distribution on the centerline is much broader than that of the ZrO_2 the resulting velocity distribution is much narrower. Since there are relatively more zirconia particles than NiCrAlY particles on the centerline for the 1:1 mixture the bimodal nature of the velocity distribution is just hinted at, and the particle size distribution is skewed toward the smaller zirconia particles. The correlation between the particle velocity and particle size for the three cases is shown in Figures 12 and 13. The lighter (and smaller) zirconia particles exhibit a much stronger dependance of velocity on size than do the NiCrAlY particles.

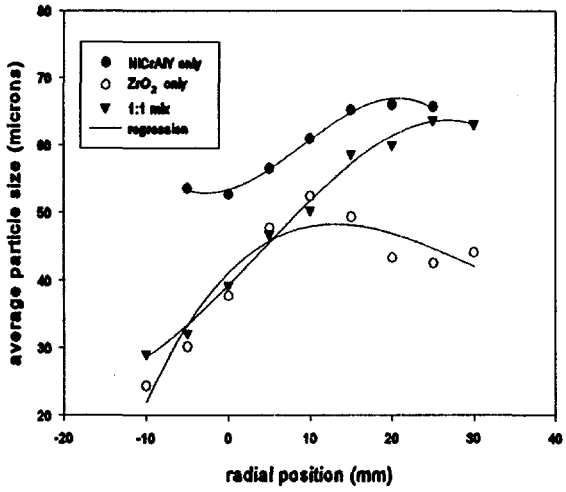


Figure 4 Radial distribution of average particle size at a standoff of 100 mm.

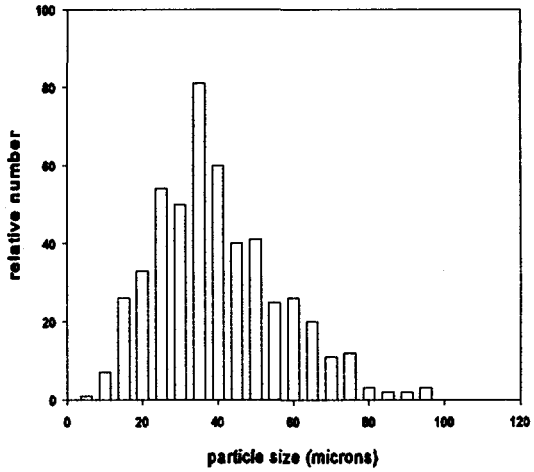


Figure 5 Zirconia particle size distribution on spray gun centerline.

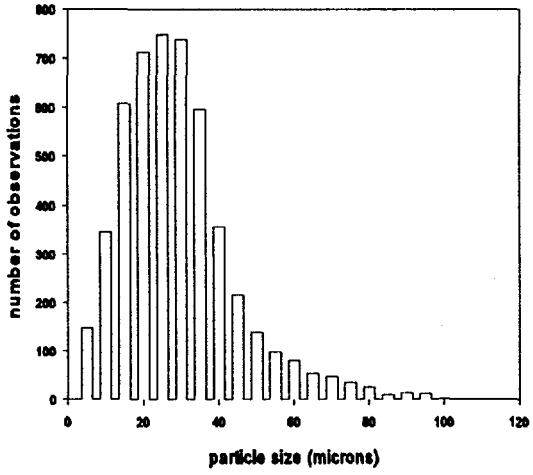


Figure 6 One-to-one by volume mixture particle size distribution on spray gun centerline.

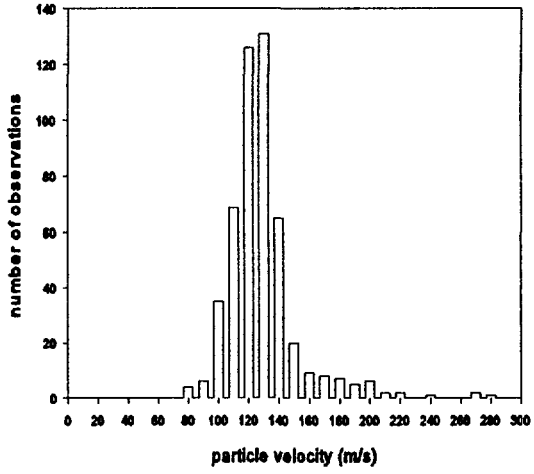


Figure 7 NiCrAlY particle velocity distribution on spray gun centerline.

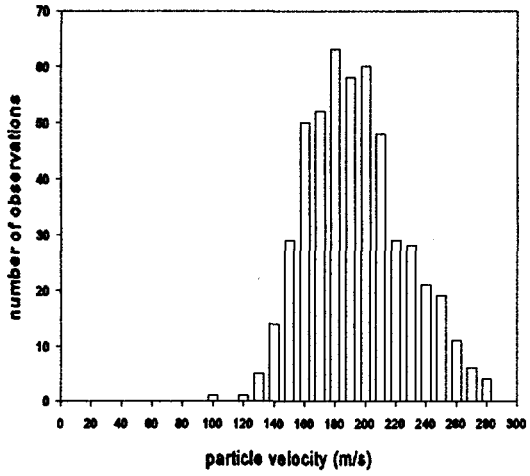


Figure 8 Zirconia particle velocity distribution on spray gun centerline.

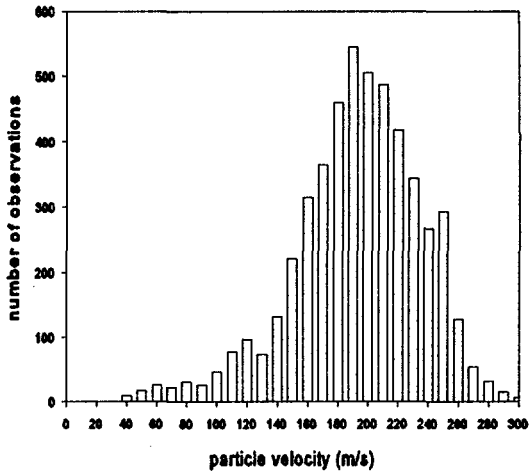


Figure 9 One-to-one by volume mixture particle velocity distribution on spray gun centerline.

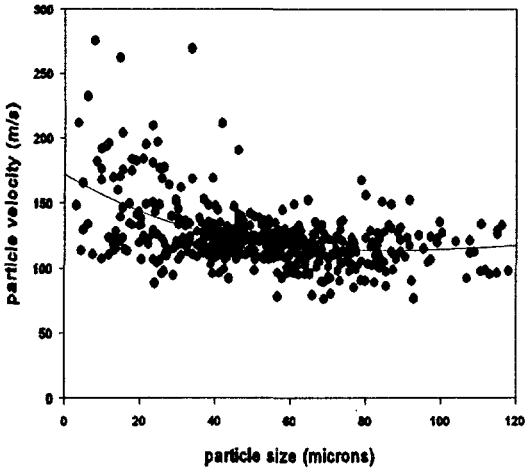


Figure 10 Correlation between particle velocity and diameter for zirconia NiCrAlY, centerline measurement.

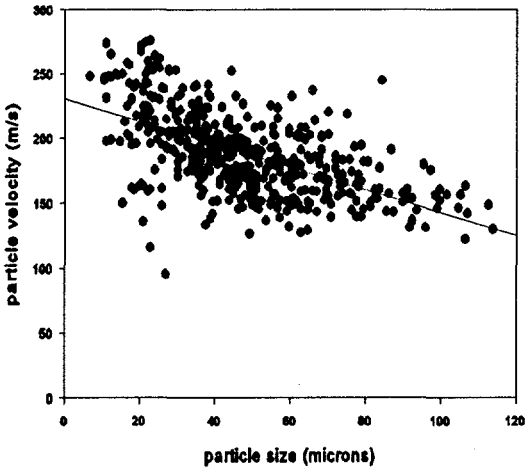


Figure 11 Correlation between particle velocity and diameter for zirconia powder, centerline measurement.

For the1:1 mixture, (Figure not shown), the range of observed particle velocities is increased with the denser NiCrAlY particles exhibiting lower velocities. The histograms of particle temperature on the gun centerline appear in Figures 14-16. The average NiCrAlY particle temperature is well above the melting point of (~1500 °C) and is significantly lower than the mean temperature of the zirconia particles which is very near their melting point of 2620 °C. The temperature distribution for the 1:1 mixture is very similar to that for the zirconia due to their relatively larger number. The data indicate that approximately 65% of the NiCrAlY particles observed were

hot, that is, yield a measurement of temperature, while approximately 50% of the zirconia particles were hot. Of the hot NiCrAlY particles >95% were above the melting temperature, while only approximately 50% of the zirconia particle observed were greater than the melting temperature. Near the maximum particle concentration (spray pattern center) at approximately -15mm, there are roughly equal numbers of NiCrAlY and zirconia particles. Under these conditions the bimodal nature of the velocity distribution becomes more apparent, Figure 17, and the particle size and temperature distributions broaden, Figure 18 and 19. The broadening of the temperature distribution, which is strongly dependent on the size distribution is accompanied by a decrease in the number of hot particles, to less that 35% with the fraction that are molten being even less.

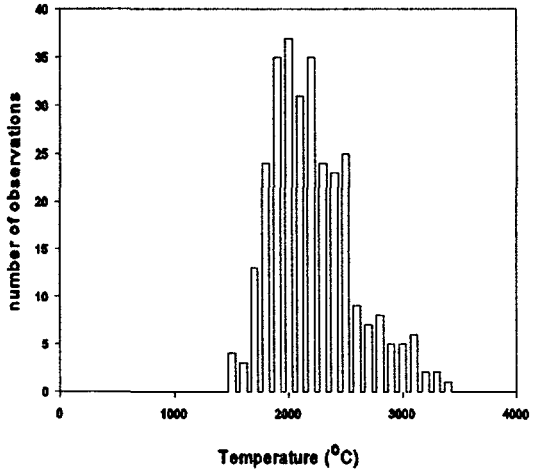


Figure 12 NiCrAlY particle temperature distribution on spray gun centerline.

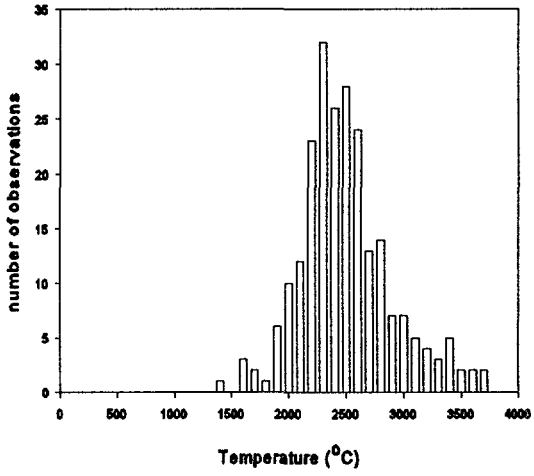


Figure 13 Zirconia particle temperature distribution on spray gun centerline.

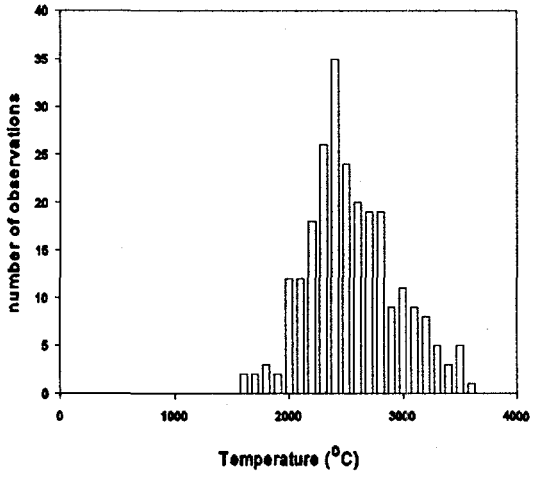


Figure 14 One-to-one by volume mixture particle temperature distribution on spray gun centerline.

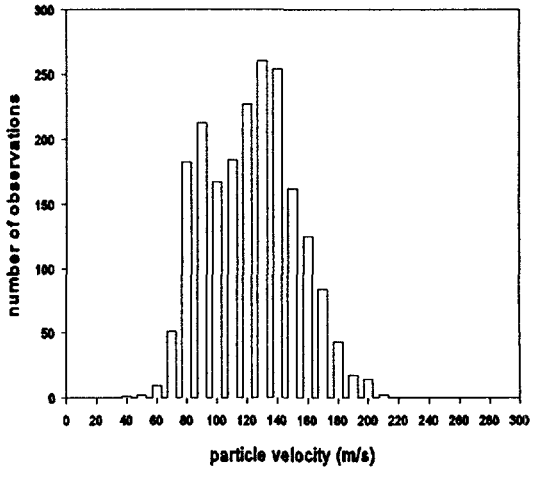


Figure 15 One-to-one by volume mixture particle velocity distribution at spray pattern center.

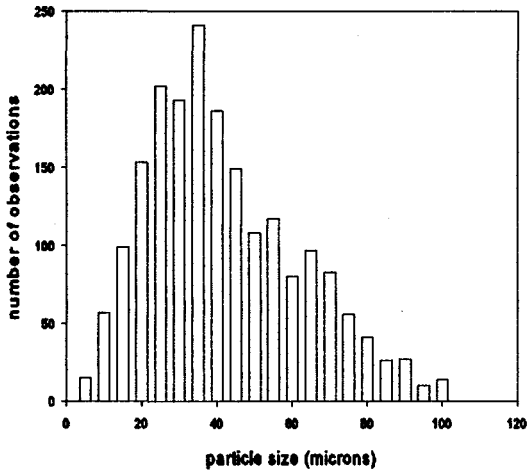


Figure 16 One-to-one by volume mixture particle size distribution at spray pattern center.

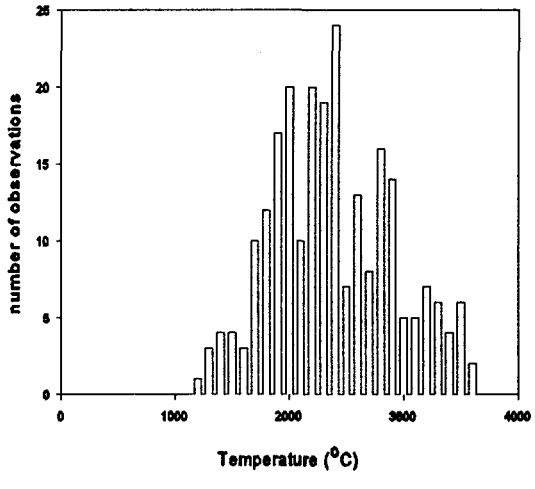


Figure 17 One-to-one by volume mixture particle temperature distribution at spray pattern center.

CONCLUSIONS

The co-injection of dissimilar particle types is complicated and significantly influenced by particle-particle interactions in the close confines of the injector. Because of this, the resulting spray field is more complicated than a simple linear superposition of individually injected single particle types. On the gun centerline approximately 65% of the metallic particles are estimated to be above the melting point, while only about 25% of the zirconia particles are above their melting point. Near the spray pattern center for 1:1 mixtures, less than 35% of the particles are "hot" and even fewer are at or above their respective melting points. In designing spray schedules for the fabrication of FGMs, the fact that deposition efficiency is influenced by spray trajectory, which in turn is influenced by mixture ratios and particle-particle interaction in the injector, must be considered and accounted for.

ACKNOWLEDGMENTS

This work was performed under the auspices of the U. S. Department of Energy under DOE Field Office, Idaho, Contract DE-AC07-94ID13223, supported by the U. S. Department of Energy, Office of Energy Research, Office of Basic Energy Sciences, Division of Engineering and Geosciences.

REFERENCES

1. Sampath, S., Herman, H, Shimoda, N., and Saito, T., MRS Bulletin, **20** (1995) 27.
2. Smith, W., Jewett, S., Sampath, S., Berndt, C., Herman, H., Fincke, J., and Wright, R., Thermal Spray: Practical Solutions for Engineering Problems, ASM international, (1996) 317.
3. Fincke, J., Swank, W., Jeffery, C., IEEE Transactions on Plasma Science, **18** (1990) 948.
4. Fincke, J., Swank, W., Jeffery, C., and Mancuso, C., Measurement Science Technology, **4** (1993) 559.

HELICAL WAVES AND NON-LINEAR DYNAMICS OF FLUID/STRUCTURE INTERACTIONS IN A TUBE ROW

F.C. Moon, M. Thothadri

Cornell University
Ithaca, New York 14853, U.S.A.

Abstract

The goal of this study has been to investigate low-dimensional models for fluid-structure dynamics of flow across a row of cylindrical tubes. Four principle results of this experimental-theoretical study are discussed. i) Experimental evidence has shown that the dynamic instability of the tube row is a subcritical Hopf bifurcation. ii) The critical flow velocity decreases as the number of flexible cylinders increases. iii) The linear model exhibits coupled helical wave solutions in the tube dynamics. iv) A nonlinear model of the tube motions shows a complex subcritical Hopf bifurcation with a secondary bifurcation to a torus or quasi-periodic oscillation. In this analysis the tools of center manifolds, normal forms and numerical simulation are used.

INTRODUCTION

It is well known that fluid-elastic instabilities and vibrations are responsible for numerous failures in heat exchange systems in the power industry. Linear models in fluid-structure dynamics are well known (e.g., Chen (1978)). However only recently have nonlinear models and analysis been used to explore the post-flutter dynamics. Our focus in this study is the cross flow of air across a row of flexible cylinders.

Connors (1970) studied the instability in a row of circular cylinders, measured the quasi-static forces and put forth a simple empirical criterion for the critical flow velocity. Research over the past couple of decades has been concentrated on predicting this critical flow velocity for different configurations. Approximate empirical results exist for certain array configurations, but no general result applicable to all configurations has been found.

The lack of understanding of the actual mechanics of the fluid-structure interaction has prevented most attempts from predicting fluid forces in sufficient detail. A few attempts have been made to get at the basis of these fluid forces, Paidoussis & Price (1988) , Chen (1978), Leaver & Weaver (1982). Although a very good insight has been obtained through these studies, their inability to predict the structural dynamics for a general system has led numerous studies to experimentally determine the fluid forces, Tanaka & Takahara (1981), Chen *et al.* (1993).

Paidoussis (1987) provides a review of the work in fluid-elastic instabilities due to internal flow, external axial flow, annular flow and cross-flow. A similar review on fluid-elastic instabilities due to cross-flow is provided by Weaver *et al.* (1987), Chen (1989).

Some nonlinear models have been put forth to capture the global behavior of the structure, Bishop and Hassan (1964), Hartlen and Currie (1970), Price *et al.* (1990). The most popular one being the Hartlen-Currie model or the wake-oscillator model. The chaotic dynamics due to impacting of the cylinders with their supports have been studied by some researchers recently; Chen *et al.* (1993). Paidoussis has reported that a rotated array on cross-flow loses stability in a supercritical Hopf bifurcation. Muntean (1995) at Cornell reported that a single cylinder in a row of rigid tubes loses stability by a negative damping mechanism in a sub-critical Hopf bifurcation and put forth a nonlinear model capturing some of the characteristics.

DESCRIPTION OF EXPERIMENTS

The wind tunnel in this study is a standard blower type low turbulence air tunnel. The test section is 91.5 cm long and is made of plexiglass. Figure 1 shows an isometric view of the test section. The test section is 256 mm $\times$ 256 mm in cross-section. The row of cylinders is positioned about 30.5 cm downstream from the upstream end of the test section. The cylinder row, consisting of nine cylinders, has seven oscillating cylinders and a fixed cylinder at each end.

Each oscillating cylinder has two degrees-of-freedom, displacements inline and transverse to the flow. Each of the oscillating cylinders is made of a hollow plexiglass tube, 19 mm OD – 16 mm ID, 256 mm long and has a mass per unit length of 0.23 kg/m. Fixed cylinders at the ends are solid 19 mm OD rods of Aluminum and are bolted to the structure containing the cylinders, thus holding the entire structure in place. A 91.5 cm long, 1.6 mm diameter steel rod passes through the cylinder caps and support their gravity. These steel rods are pinned at the top and the bottom ends, providing the cylinders with the required degrees of freedom. The pitch-to-diameter ratio is 1.35. The relative static position of the cylinders in the inline direction can be adjusted by increasing or decreasing the length of the springs, keeping all the cylinders aligned. The natural frequency of the cylinders in the inline direction is 7.4 Hz and 6.8 Hz in the transverse direction. The damping ratio in still air was found to be close to 0.013 in both directions.

Figure 2 shows evidence for a *subcritical Hopf bifurcation.* This bifurcation diagram for the *central cylinder* is obtained with all the seven cylinders oscillating, i.e. the responses of only the central cylinder were used to obtain Figure 2. The bifurcation diagram for the rest of the cylinders seemed to be qualitatively similar to the one found for the central cylinder. For small flow velocities, the cylinder oscillations were small amplitude turbulent buffetting due pressure fluctuations, flow noises, etc. The amplitude of these oscillations are found to increase almost linearly with flow velocity as has been reported in earlier studies, Muntean (1995), Blevins (1977). Moreover, as the flow velocity is increased, the stochastic component of the cylinder response decreases. No modal patterns of oscillations could be detected in the cylinder row for turbulent buffetting.

As the flow velocity is increased further, small scale buffeting of the cylinders persists, but *sufficiently* large perturbations of the cylinders result in large amplitude, periodic limit

cycle oscillations. The frequency of these limit cycle oscillations was found not to vary much with flow velocity. The reduced flow velocity at the sub-critical Hopf bifurcation was found to be, $U_{cr} \approx 115$. Beyond this critical flow velocity only large amplitude oscillations are present and further increase in the flow velocity leads to impacting of the cylinders with the support structure. To estimate the amplitude of the unstable limit cycle the central cylinder was perturbed from the buffetting oscillations by giving it a known displacement in the direction of the flow (initial velocity is zero). The smallest such displacement that resulted in the large amplitude limit cycle oscillations (upper branch) was recorded as the amplitude of the unstable limit cycle for the corresponding flow velocity.

The variation in the critical reduced flow velocity as the number of oscillating cylinders was found. This experiment was performed by allowing a certain number of cylinders to oscillate while restraining the motions of the rest. The number of cylinders of cylinders was increased in a symmetric fashion around the central cylinder. As the number of oscillating cylinders is increased, the *critical flow velocity* was found to *decrease*. The onset of instability for a single oscillating cylinder was far removed from those for more than one cylinder. Thus coupling between the cylinders drastically reduces the onset of instability.

THEORETICAL MODEL

The equations of motion of n^{th} the cylinder, with respect to a frame fixed at its equilibrium position, is given by

$$M \begin{Bmatrix} \ddot{x}_n \\ \ddot{y}_n \end{Bmatrix} + \begin{bmatrix} C_x & 0 \\ 0 & C_y \end{bmatrix} \begin{Bmatrix} \dot{x}_n \\ \dot{y}_n \end{Bmatrix} + \begin{bmatrix} K_x & 0 \\ 0 & K_y \end{bmatrix} \begin{Bmatrix} x_n \\ y_n \end{Bmatrix} = \boldsymbol{f}_n \tag{1}$$

where the terms on the left-hand side are the inertial, damping and stiffness terms of the cylinder structure and $\boldsymbol{f}_n$ on the right-hand side is a vector containing the fluid forces on the n^{th} tube in the x and y directions. The displacements of the n^{th} cylinder are denoted by x_n and y_n.

The profile of the flow (upstream) across the cylinder row was experimentally found to be fairly uniform and the cylinders are coupled only by the fluid flowing past them, i.e. there is no structural coupling.

As a result of the uniformity of the flow across the cylinder row and the periodicity of the cylinders, any set of three cylinders has a symmetry along the flow direction about the middle cylinder. The fluid force on the n^{th} cylinder due to the relative motions of the (n−1)th and the (n+1)th cylinder are related by $\boldsymbol{f}_n^{n-1} = \boldsymbol{f}_n^{n+1}$ for the direct coupling of the x and y directions and by $\boldsymbol{f}_n^{n-1} = -\boldsymbol{f}_n^{n+1}$ for the cross coupling between the x and y directions. Using these symmetry conditions and linearizing the fluid forces about the cylinder equilibrium positions we have,

$$\begin{aligned} \boldsymbol{f}_n = & -\tfrac{1}{2}\rho D U \left(\begin{bmatrix} \alpha' & -\sigma' \\ -\tau' & \beta' \end{bmatrix} \begin{Bmatrix} \dot{x}_n - \dot{x}_{n-1} \\ \dot{y}_n - \dot{y}_{n-1} \end{Bmatrix} + \begin{bmatrix} \alpha' & \sigma' \\ \tau' & \beta' \end{bmatrix} \begin{Bmatrix} \dot{x}_n - \dot{x}_{n+1} \\ \dot{y}_n - \dot{y}_{n+1} \end{Bmatrix} \right) \\ & + \rho U^2 \left(\begin{bmatrix} \alpha'' & -\sigma'' \\ -\tau'' & \beta'' \end{bmatrix} \begin{Bmatrix} x_n - x_{n-1} \\ y_n - y_{n-1} \end{Bmatrix} + \begin{bmatrix} \alpha'' & \sigma'' \\ \tau'' & \beta'' \end{bmatrix} \begin{Bmatrix} x_n - x_{n+1} \\ y_n - y_{n+1} \end{Bmatrix} \right) \end{aligned} \tag{2}$$

The single-primed quantities, $(\alpha', \sigma', \beta', \tau')$ are the fluid-damping coefficients and the double-primed quantities, $(\alpha'', \sigma'', \beta'', \tau'')$ are the fluid-stiffness coefficients. The notation used here is slightly different from the one used by Chen *et al.* (1993b).

The coefficients α'', α' couple the x-motions of adjacent cylinders and β'', β' couple y-motions. The coefficients τ'', σ'', τ', σ' couple the x (y)-motions of one cylinder to the y (x)-motions of the adjacent cylinders. The direct coupling coefficients were measured by using the ARMA process (Auto Regressive Moving Average) and the cross coupling coefficients were determined through heuristic fits based on instability mechanisms.

ANALYTICAL RESULTS

We consider a traveling wave solution to Equations 1 and 2 of the form

$$\left\{ \begin{array}{c} x_n \\ y_n \end{array} \right\} = \left\{ \begin{array}{c} a_1 \\ a_2 \end{array} \right\} e^{i(qn-\omega t)} \tag{3}$$

where, $q = 2\pi\kappa P$; κ is the wave number (i.e. $\frac{2\pi P}{q}$ is the wave length), ω is the angular frequency and P (pitch) is the periodicity associated with the structure. In the above traveling wave solution, the $e^{-i\omega t}$ characterizes the temporal behavior of the solution for the initial data given by e^{iqn}. The quantity q, is the phase difference between adjacent cylinders and hence characterizes the spatial distribution of the cylinders at any instant of time.

The traveling wave solution, Equation 3, written compactly as $\boldsymbol{x} = \boldsymbol{a}e^{i(qn-\omega t)}$, implies that $\boldsymbol{x}_{n+1} = \boldsymbol{x}_n e^{iq}$. This implies that the frequency $\omega(q)$ has a period 2π in q, i.e. $\omega(q) = \omega(q+2\pi)$ and as $q = 2\pi\kappa P$, the above periodicity leads to $q(\kappa) = q(\kappa + \frac{1}{P})$. If we restrict q between $[-\pi, \pi]$ then $-\frac{1}{2P} \leq \kappa \leq \frac{1}{2P}$ which means the wavelength $\lambda \geq 2P$.

If q is complex the real part of q contributes to the spatial distribution of cylinders (wave patterns) and the imaginary part results in *spatial* growth or decay(depending on the sign) as $n \rightarrow \infty$ at any given time. These waves with complex wave numbers are called *evanescent waves* or *cutoff waves.*

Figure 3 shows a helical wave mode for $u_r = 90$.

By this representation, the wave in Figure 3 *at any instant in time*, is a helix that circles around the $+x$-axis in the counter-clockwise direction. The points of constant phase advance in the $+x$-direction with velocity, ω_r/q_1. Moreover, any cylinder (i.e. for a fixed n) will circle the $+x$-axis in the clockwise direction for this wave.

For the fixed-ends boundary condition (i.e. $\boldsymbol{x}_o = \boldsymbol{x}_{N+1} = 0$), the constraint of zero displacements for the end cylinders implies that the wave numbers q, can no longer be purely real. The imaginary part of the wave numbers are positive and this ensures that the waves are spatially decayed to satisfy the boundary conditions at the ends for all time.

NONLINEAR BIFURCATION ANALYSIS

Our ultimate goal is to understand what the nature of the nonlinear forces are that lead to a subcritical dynamic instaiblity. These instabilities are very dangerous in practice. The subcritical phenomenon however depends on the nonlinear nature of the fluid coupling forces. In the absence of dynamic measurements of the nonlinear part of the fluid forces, we posit a plausible nonlinear model. In this model we consider two tubes; one constrained to

vibrate transverse to the flow, and the other constrained to vibrate in the flow direction. We represent the dynamics of this sytem by the equations;

$$\begin{Bmatrix} \dot{x} \\ \dot{y} \\ \ddot{x} \\ \ddot{y} \end{Bmatrix} = \boldsymbol{A}(u_r) \begin{Bmatrix} x \\ y \\ \dot{x} \\ \dot{y} \end{Bmatrix} + (\frac{u_r}{u_{cr}})^2 \begin{Bmatrix} 0 \\ 0 \\ C_{30}\dot{x}^3 + C_{31}\dot{x}^2\dot{y} + C_{32}\dot{x}\dot{y}^2 + C_{33}\dot{y}^3 \\ D_{30}\dot{y}^3 + D_{31}\dot{y}^2\dot{x} + D_{32}\dot{y}\dot{x}^2 + D_{33}\dot{x}^3 \end{Bmatrix} + \frac{u_r}{u_{cr}} \begin{Bmatrix} 0 \\ 0 \\ C_{50}\dot{x}^5 \\ D_{50}\dot{y}^5 \end{Bmatrix} \tag{4}$$

where, $\boldsymbol{A}(u_r) = \begin{bmatrix} \boldsymbol{O} & \boldsymbol{I} \\ -\boldsymbol{K}(u_r) & -\boldsymbol{C}(u_r) \end{bmatrix}$

is the matrix governing the linearized equations, $\boldsymbol{K}$ is the non-symmetric stiffness matrix, $\boldsymbol{C}$ is the damping matrix and u_r is the parameter in the system, the reduced flow velocity. The various elements of the stiffness and damping matrices were obtained for the two degree-of-freedom system from the linear model obtained earlier.

In this model u_r represents a normalized flow velocity, while u_{cr} is the critical flow velocity for flutter when two of the eigenvalues of A cross the imaginary axis.

The *center manifold theorem* ensures that for the full nonlinear equations there exists a subspace (center manifold) which is tangential to the subspace spanned by the eigenvectors, at the equilibrium point. The flow close to the equilibrium point approaches the center manifold asymptotically. The theorem also ensures that the stability of the equilibrium point in the full nonlinear equations is the same as the stability when restricted to flow on the center manifold. Moreover, any additional equilibrium point or limit cycles which occur in a neighborhood of the given equilibrium point on the center maifold are gauranteed to exist in the full nonlinear equations.

Near-identity transformations of increasing order were used to reduce the flow on the center manifold, obtained from the center manifold analysis described above, to a normal (canonical) form. After reducing the equations to a normal form a further transformation into polar coordinates was carried out (see Rand and Armbruster (1987)).

The system of Equation 4 was studied through numerical simulation. The amplitude of the limit cycles in the Bifurcation diagram was obtained by using Newton's root finding method on the Poincare Map, obtained by taking a Poincare section at $\dot{y} = 0$. The stability of the limit cycles were obtained from the eigenvalues of the Jacobian evaluated at the limit cycle. See Figures 4 and 5.

Figure 4 shows the subcritical branch of the unstable limit cycle that occurs at the critical flow velocity predicted by linear theory. At lower flow velocities both a stable limit cycle of large amplitude is shown as well as a torus on quasi-periodic vibration (characterized by two incommensurate frequencies). These results were obtained by using a Poincare map technique shown in Figure 5, which plots $\{x,\ y,\ \dot{x}\}$ when $\dot{y} = 0$. A point or dot indicates a limit cycle or periodic oscillation, while a closed curve indicates a quasi-periodic oscillation.

CONCLUSIONS

This study has shown that experimental dynamic instabilities in tubes in cross flow can exhibit subcritical limit cycle oscillations (Hopf bifurcation). These subcritical Hopf bifurcations are dangerous in practice because large amplitude oscillations can be induced at

much lower flow velocities than those predicted by classical linear models for fluid-structure instability. The nature of these subcritical instabilities depends on the nonlinear fluid and structural forces. We have used an ad-hoc nonlinear model to show the possible complexities in the secondary bifurcations as a function of flow velocity. This study has demonstrated the need to obtain better measurements of dynamic fluid structure forces in order to provide realistic predictive models of nonlinear fluid-structure instabilities.

REFERENCES

1. R.D. BISHOP and A.Y. HASSAN, "The lift and drag forces on a circular cylinder in flowering field". In *Proceedings of the Royal Society of London, Series A, 277*, pages 51–75, (1964).

2. R.D. BLEVINS, *Flow-Induced Vibrations.* Van Nostrand Reinhold Company, (1977).

3. S.S. CHEN, "Crossflow-Induced Vibrations of Heat Exchanger Tube Banks," *Nuclear Engineering and Design*, 47:67–86, (1978).

4. S.S. CHEN, S. ZHU, and J.A. JENDRZEJCZYK, "Motion-dependent Fluid Forces Acting on Tube Arrays in Crossflow," Research Report ANL-93/15, Argonne National Laboratory, (1993).

5. H. CONNORS, Jr., "Fluid-elastic Vibration of Tube Arrays Excited by Cross Flow," in D.D. Reiff, editor, *Flow-Induced Vibration in Heat Exchangers*, pages 42–56. ASME, (1970).

6. J.H. LEAVER and D.S. WEAVER, "A Theoretical Model for Fluid-elastic Instability in Heat Exchangers," *ASME Journal of Pressure Vessel Technology*, 104:147–158, (1982).

7. G. MUNTEAN, *Influence of Fluid Wake Structure on the Dynamics of a Flexible Tube Row in Cross Flow*, Ph.D. Thesis, Cornell University, (1995).

8. M.P. PAIDOUSSIS, "Flow Induced Instabilities of Cylindrical Structures," *Applied Mechanics Review*, 40(2):163–175, (1987).

9. M.P. PAIDOUSSIS and F.C. MOON, "Nonlinear and Chaotic Fluid-Elastic Vibration of a Flexible Pipe Conveying Fluid," *Journal of Fluids and Structures*, 2:567–591, (1988).

10. M.P. PAIDOUSSIS and S.J. PRICE, "The Mechanisms Underlying Flow-induced Instabilities of Cylinder Arrays in Cross Flow, *Journal of Fluid Mechanics*, 187:45–59, (1988).

11. R.H. RAND, and D. ARMBRUSTER, *Perturbation Methods, Bifurcation Theory and Computer Algebra*, Springer-Verlag, (1987).

12. H. TANAKA and S. TAKAHARA, "Fluid-elastic Vibration of tube Array in Cross Flow," *Journal of Sound and Vibration*, 77(1):19–37, (1981).

13. D.S. WEAVER And J.A. FITZPATRICK, "A Review of Flow Induced Vibrations in Heat Exchangers," In *Proceedings of the International Conference in Flow-Induced Vibrations*, pages 1–17, (1987).

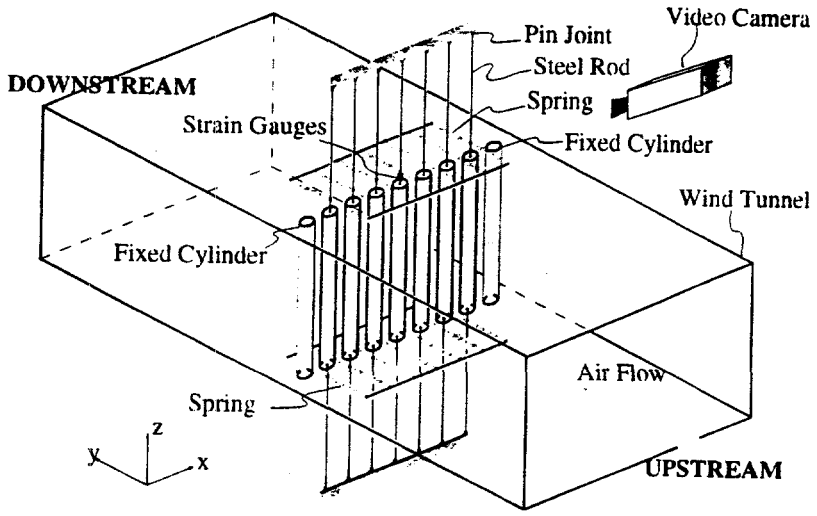


FIGURE 1

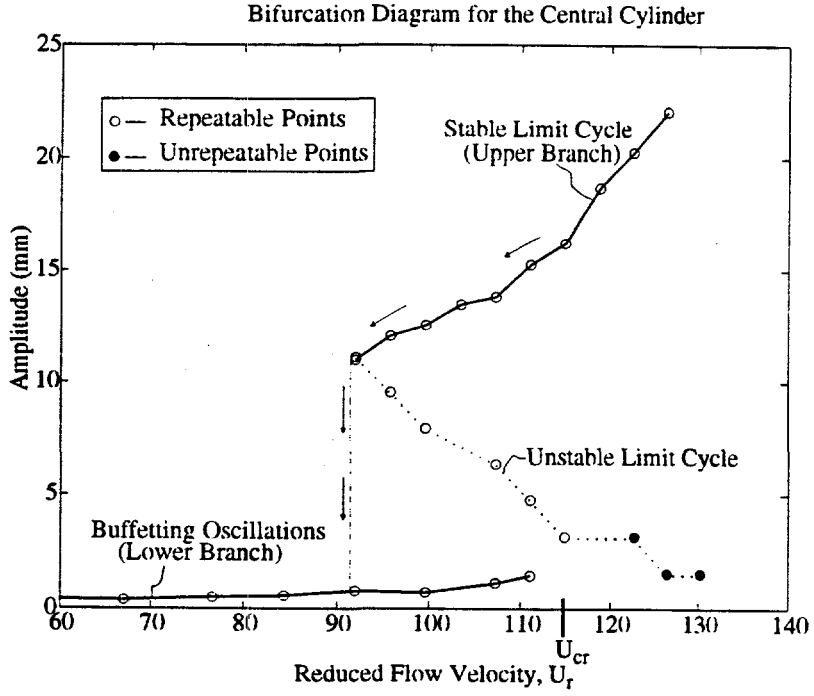
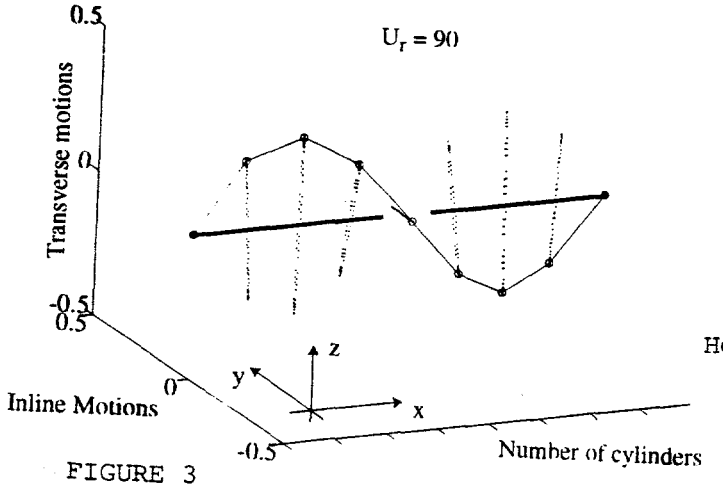


FIGURE 2



Helical Wave Mode

FIGURE 3

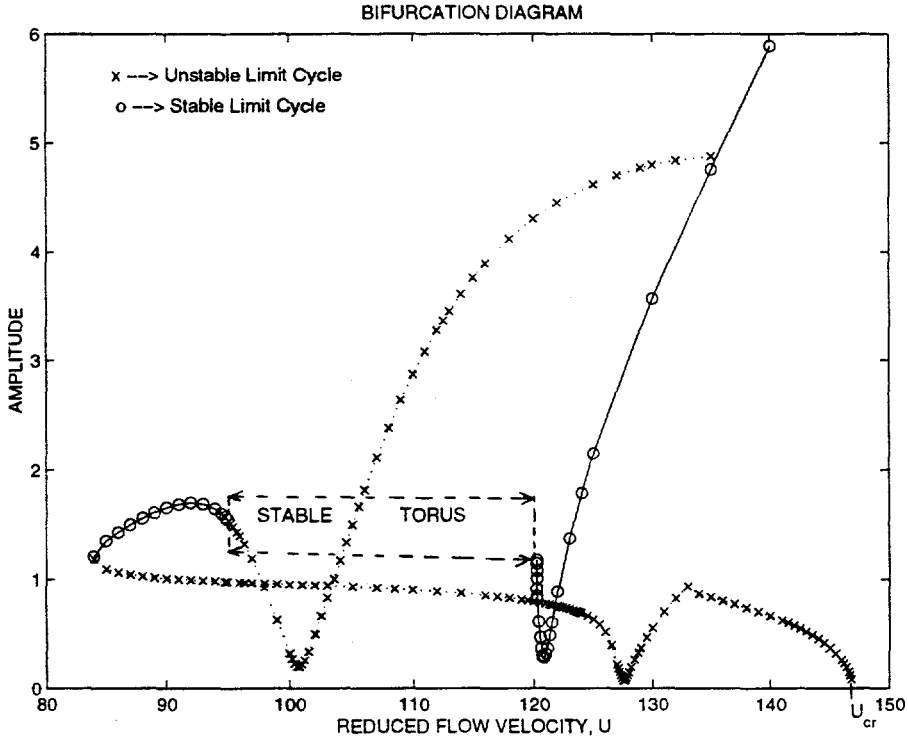
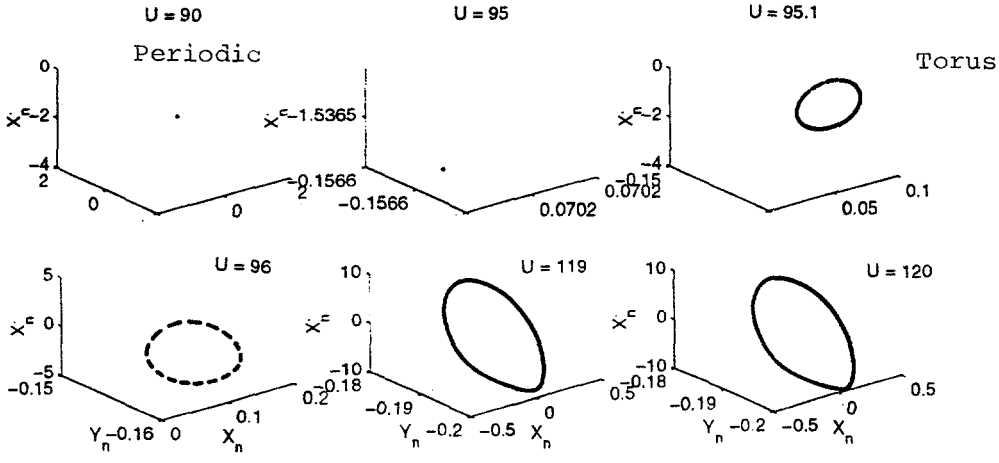


FIGURE 4



Periodic

Torus

FIGURE 5 Poincare Maps

FUNDAMENTAL STUDIES OF SPRAY COMBUSTION

S.C. Li, P.A. Libby and F.A. Williams

Center for Energy and Combustion Research
and
Department of Applied Mechanics and Engineering Sciences
University of California, San Diego
La Jolla, CA 92093-0411

ABSTRACT

Our research on spray combustion involves both experiment and theory and addresses the characteristics of individual droplets and of sprays in a variety of flows: laminar and turbulent, opposed and impinging. Currently our focus concerns water and fuel sprays in two stage laminar flames, i.e., flames arising, for example from a stream of fuel and oxidizer flowing opposite to an air stream carrying a water spray. Our interest in these flames is motivated by the goals of reducing pollutant emissions and extending the range of stable spray combustion. There remains considerable research to be carried out in order to achieve these goals. Thus far our research on the characteristics of sprays in turbulent flows has been limited to nonreacting jets impinging on a plate but this work will be extended to opposed flows with and without a flame. In the following we discuss details of these studies and our plans for future work.

WATER SPRAYS IN TWO STAGE LAMINAR FLAMES

To improve understanding of the chemical and fluid mechanical interactions between water droplets and the formation of NOx in flames we have carried out studies of two stage laminar flames. To do so we have employed our counterflowing burner which is the principal device used in our experimental effort. In these studies one stream is a fuel-rich mixture of methane and air with a stoichiometric ratio denoted Φ while the opposed stream involves air and a spray consisting of relatively small water droplets. Different droplet loadings and different flow conditions permit us to investigate the influence of strain rates and water concentration on NOx formation in flames.

Under the conditions we have studied there exists a vaporization plane, i.e., a plane close to the flame by which all the water has vaporized. Because of the relatively low temperature of the gas mixture at this plane, the distribution of temperature in the flame is altered by the water which therefore has both a chemical and a thermal influence.

In these flames there occur a green premixed reaction zone in which the oxidizer in the methane-air stream is consumed, a blue diffusion zone in which the residual fuel components namely CO and HO are consumed by the air in the air-droplet stream and finally a vaporization plane as noted earlier. In the neighborhood of the axis all three of these surfaces are flat and parallel to one another. The separation distance between them decreases with increasing rates of strain and equivalence ratio of the fuel-air stream. In experiments on these flames temperature profiles are measured with thermocouples and concentration profiles of the stable species are measured by gas chromatography.

Complementing in an important way these experiments are theoretical-numerical studies in which relatively complete descriptions of the fluid mechanics and chemistry of these flows from the exit plane of one jet to the exit plane of the other are considered. The present computations involve 140 elementary reactions among the following 38 species:

CH_4, O_2, CO, CO_2, H_2, H_2O, H, OH, O, HO_2, H_2O_2, CH, CHO, CHCO, $^1CH_2$, $^3CH_2$, CH_2O, CH_3, C_2H, C_2H_2, C_2H_3, C_2H_4, C_2H_5, C_2H_6, N, NH, NH_2, NH_3, NO, HNO, HCN, HNCO, NCO, CN, NO_2, N_2, N_2H, N_2O.

Both experiment and computation show that the methane and oxygen in a fuel rich mixture react rapidly in the premixed zone to form CO and H whose concentrations reach maxima as the methane disappears. The further oxidation of CO and H to form CO and HO occurs in the diffusion zone. The C_2 species such as C_2H_2, C_2H_4 and C_2H_6 are predicted to be produced rapidly in the early stages of the premixed zone but these species are consumed in the diffusion zone. The concentrations of the C_2 species strongly depend on the stoichiometry of the methane-air mixture; the larger the value of Φ the higher the concentrations of these species, a finding consistent with the observed soot emissions in fuel rich methane-sir flames. Water is the most important agent for the consumption of the radical CH which leads to prompt NOx. Thus the addition of water to flames not only reduces prompt NOx because of its reduction of CH but reduces thermal NO_x because of its reduction of the maximum temperature. This dual effect has been widely known but our study has established the rate of the crucial reaction $CH + H_2O \rightarrow CH_2O + H$.

Figures 1 and 2 present some of the results of our studies. These figures show the distributions across the flame of the temperature and mole fraction of NO for various fuel-air ratios. Figures 1 and 2 relate to these flames without water and with 10% water addition respectively. The important result shown here is that there is a 50% reduction in peak NO concentration with water addition. Further results are given in Fig. 3 where the emission index E is plotted versus equivalence ratio for various rates of water addition. Again we see the significant benefit realized from the addition of water.

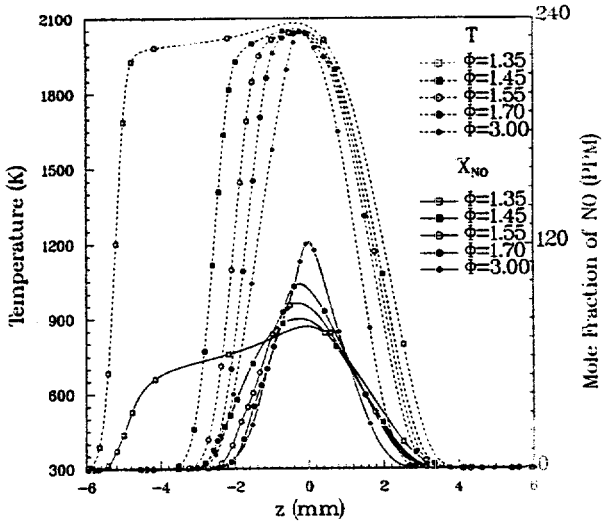


Figure 1. Computed profiles of temperature and NO concentration without water.

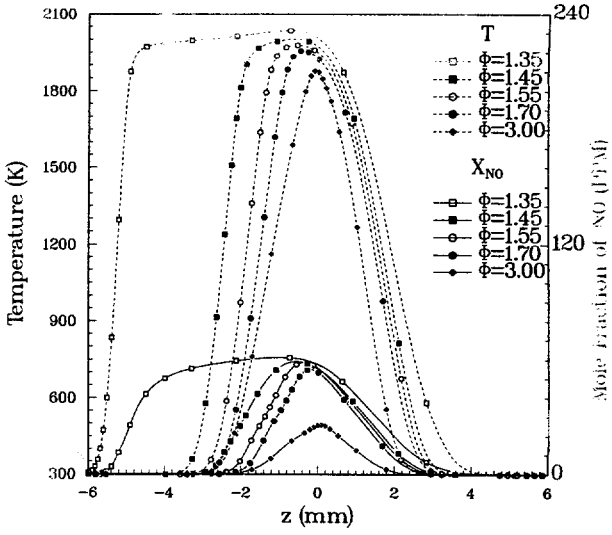


Figure 2. Computed profiles of temperature and NO concentration with 10% water added in air stream.

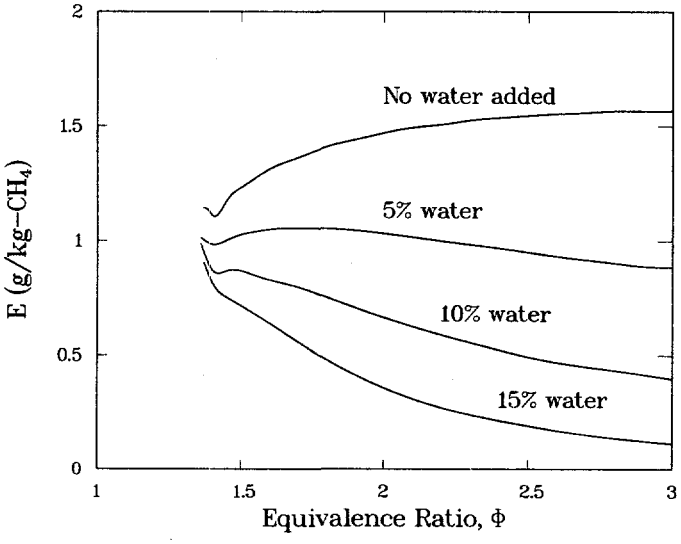


Figure 3. Predicted NO emission index as a funciton of equivalence ratio and mass percentage of water added to the air stream

TWO-STAGE METHANOL SPRAY FLAMES

In earlier studies reported in [1, 2] we have described the combustion of fuel droplets carried in a nitrogen stream counter to an air stream. This configuration results in the usual spray diffusion flame but we have recently been studying two-phase, two-stage flames which are analogous to those discussed in the previous section but with water replaced by fuel. We thus consider fuel droplets carried by air flowing counter to an airstream. The droplet loading is such

that the fuel-air stream is fuel rich, i.e., has a value of Φ greater than unity. Such flames have several distinct advantages in particular that they permit the chemistry of fuel sprays under a wide rage of reaction zone conditions to be studied both experimentally and computationally. In particular coordinated studies of these flames permit the identification of the reactions which are critical to the combustion of liquid fuels in much that same fashion as we were able to identify the critical reactions for the chemical inhibition of NOx formation by water. Moreover in experiments the increased thickness of the reaction zone results in improved spatial resolution and thus increased data accuracy.

Our recent research on two phase methanol flames has involved a stream of droplets transported in air with an equivalence ratio ranging from 1.6 to 3.0 and a fuel side rate of strain varying from 50 s^{-1} to 100 s^{-1}. The droplet sizes are such that again there is a vaporization plane. Moreover the flame also involves a pale green premixed zone and a blue diffusion zone. The same experimental techniques as described earlier are employed to measure the distributions of temperature and concentrations of the stable species and the same computational methods described earlier are brought to bear. Here comparison of experiment and computation permit values of several key reaction rates of uncertain value to be determined.

The distribution along the axis of the computed axial velocity for a typical flow is shown in Fig. 4. We have data only for the portion of the flow in which droplets exist but for that limited region excellent agreement is seen. Also shown are the experimental data on the Sauter mean diameter; the rapid decrease in droplet diameter in the neighborhood of the edge of the flame is consistent with the existence of a vaporization plane as noted earlier. The interesting increase in the mean diameter upstream of that plane is a consequence of the larger diameters being decelerated to a lesser extent than the smaller droplets in the decelerating gas stream. In due course even these larger droplets are vaporized.

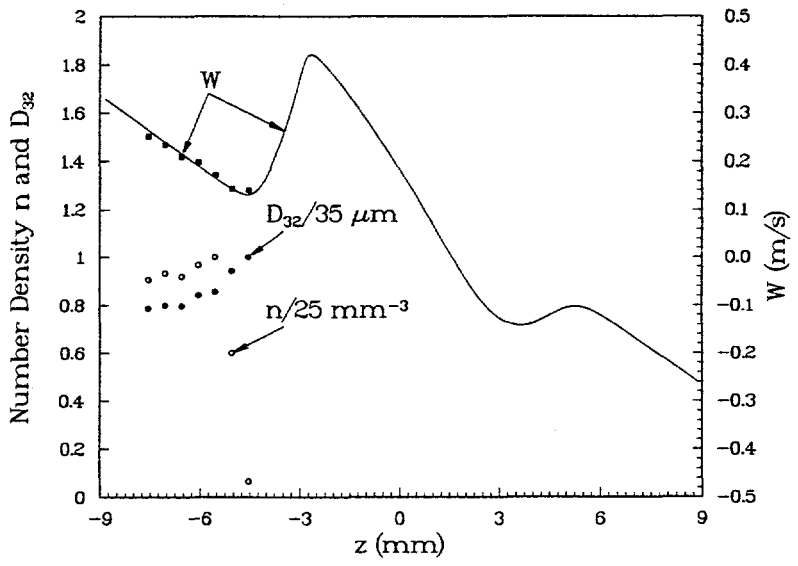


Figure 4. Profiles of axial velocity, number density and Sauter mean diameter on the burner axis.

In Fig. 5 we compare computed and measured distributions of the temperature and major species. We see two high temperature regions corresponding to the two reaction zones discussed earlier, namely the diffusion zone at z = 2 mm and the lower temperature premixed zone at z = 3 mm. Comparison of Figs. 4 and 5 establishes that the peaks in the velocities coincide with the temperature maxima. The experimental and computed distributions of the concentrations of various species shown in Fig. 5 are in excellent agreement. Of particular interest is the location of the peaks in the CO and H_2 concentrations implying agreement as to the location of the premixed zone.

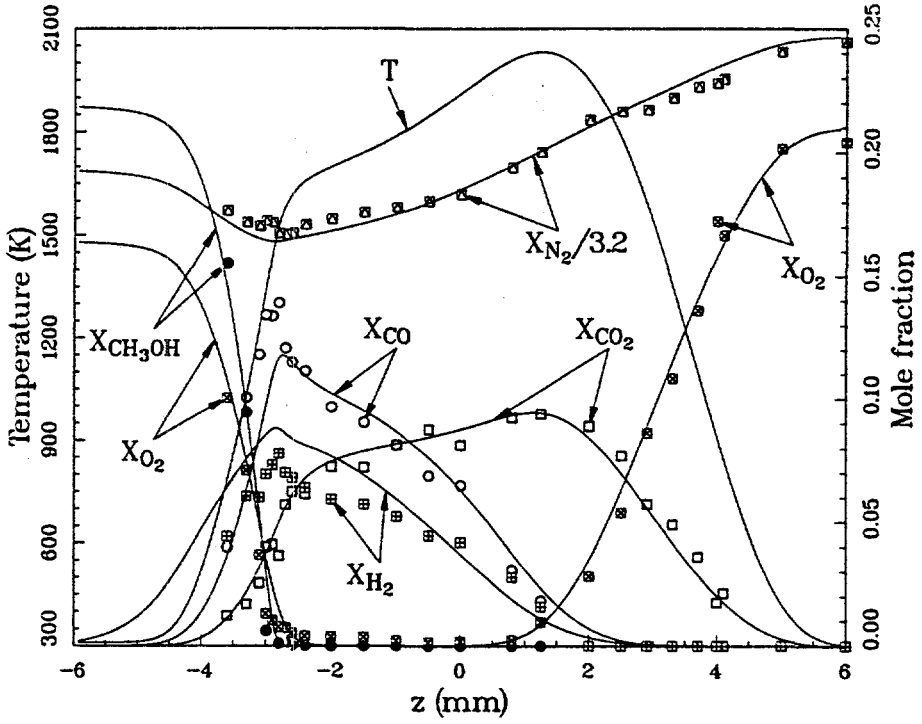


Figure 5. Comparison between measurement and prediction of profiles of temperature and concentration for $\Phi = 2$ with $a = 50\ s^{-1}$.

In Fig. 6 we show the experimental and computed distributions of the mole fraction of the various C_2 species, namely C_2H_2, C_2H_4 and C_2H_6. Two sets of computed results are shown; for the solid curves the chemistry is the same as that used in Fig. 5 while for the dashed curve the isomerization reaction $CH_3O + M \rightarrow CH_2OH + M$ has been removed. This removal is seen in Fig. 6 to displace the premixed zone appreciably toward the stagnation plane and to increase the peak concentrations of methane and the C_2 species. This influence of one crucial reaction demonstrates the value of coordinated experimental and computational studies in establishing the correct chemistry in these systems. Because of its significantly greater concentration of the OH radical, this figure also demonstrates the importance of H atoms in these flames. By comparing experiment and computation we are able to deduce a rate constant for the reaction $CH_2OH + H \rightarrow CH_3 + OH$ which is the dominant reaction in the formation of CH_4 and C_2 species. Previous values for this constant were found to be in error.

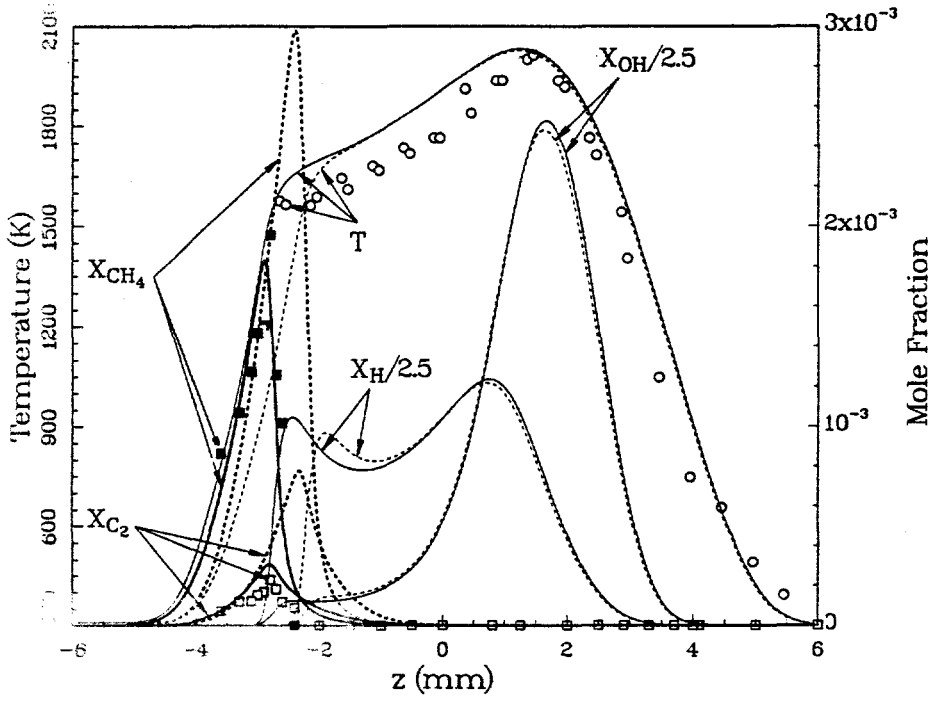


Figure 6. Comparison between measurement and prediction of profiles of temperature and concentrations of minor species for $\Phi = 2$ with $a = 50\ s^{-1}$.

TURBULENT SPRAYS IN STAGNATION FLOWS

Our studies of turbulent sprays have involved a stream impinging on a wall. In this case a grid is installed near the exit plane of the jet. We again use our PDPA to measure two velocity components and droplet diameter. Our spray generator is operated so as to produce a relatively wide range of droplet diameters such that for our flow conditions the smallest droplets can be considered to follow the fluctuations in gas velocity. We are thus able to obtain at various positions in the flow the statistics of the velocity of the gas and of droplets of various sizes. A more direct measurement of relative velocities would involve simultaneous measurements of gas and droplet velocities at a particular point in space and time. This is beyond the capability of our current instrumentation but in the near future a particle image velocimetry system will be brought to bear on this problem.

Despite this limitation we have used our data to construct a simplified description of the response of droplets of different sizes to turbulence. A feature of this description is that two bounding droplet sizes can be estimated: The largest droplets that effective respond to turbulent

velocity fluctuations and the smallest droplets that are immune to such fluctuations. Figure 7 presents these results for our experiment in which the Kolmogorov wave number is estimated to be 185 and the wave number corresponding to the integral scale is three; plotted are the droplet diameter in microns against the normalized wave number of the velocity fluctuations. Thus for a given wave number we obtain the two diameters in question. The theory reported in Fig. 7 can be used to make similar plots for other turbulent flows, i.e., those with different spectral ranges.

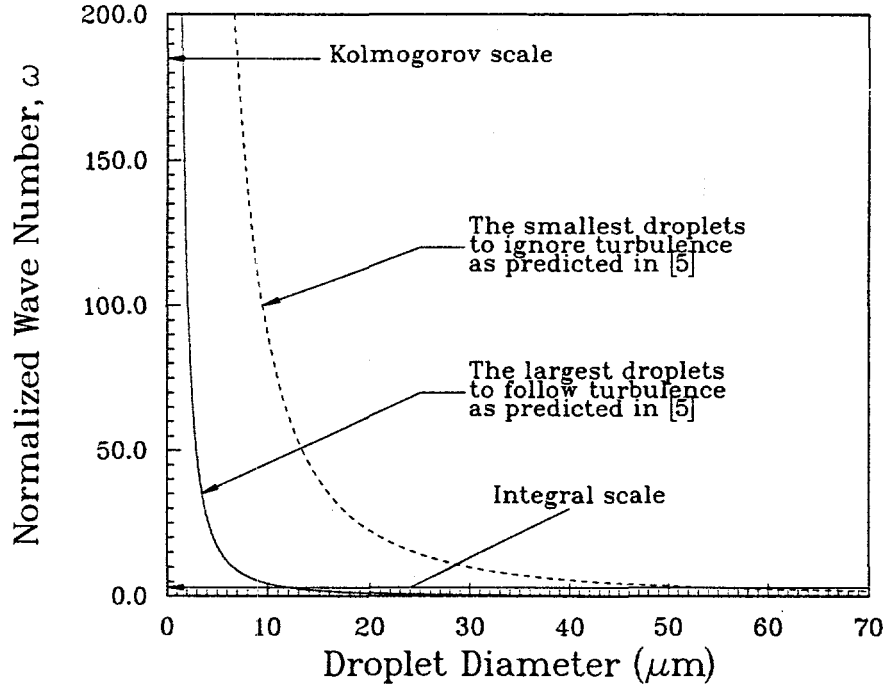


Figure 7. The largest droplets to follow and smallest droplet to ignore turbulent fluctuations of various scales.

LAMINAR COUNTERFLOWING STREAMS IN A NARROW CHANNEL

In one of our earlier investigations we studied the dynamics of droplets in non-reactive opposed laminar flows [1]. Of particular interest in the present context was the finding that for a given set of flow conditions, i.e., jet spacing and jet velocities, droplets of a well defined range of sizes would oscillate across the stagnation plane. Provided a drag law is assumed the dynamics of such droplets are readily studied since the flow field is simply and accurately described. However, there is an essential experimental difficulty of photographing the motion of such droplets in axisymmetric jet flows, namely that the droplets always leave the focal plane of the camera and are lost for purposes of observation. The remedy we found was to photograph the response of the droplets in an opposed slot flow. In this case opposed flows enter the passage between two plates and leave in orthogonal directions. The resultant flow has three planes of symmetry. While this configuration permitted the oscillations in question to be readily photographed, the flow field in

which the droplets move must be analyzed if calculated and observed droplet trajectories are to be compared. As a consequence we have been studying the flow field in question.

From a fundamental point of view this slot flow is an addition to the various relatively simple fundamental solutions to the Navier-Stokes equations and is therefore of interest on its own. By invoking the various symmetry requirements and restricting attention to narrow slots we obtain a set of three partial differential equations with two independent variables: the streamwise coordinate and the coordinate normal to the walls. The dependent variables are the streamwise velocity, the transverse gradient of the transverse velocity and the pressure. Not surprisingly a single parameter appears in these equations, namely a Reynolds number. As is typically the case in fundamental flows, solutions to these equations are readily obtained for two limiting cased, i.e., for small and large Reynolds numbers. For intermediate values numerical solutions involving iteration must be constructed.

At the present time the numerical analysis of the limiting cases is completed and examples for intermediate Reynolds numbers are being considered. When the analysis of the flow field in these slots is completed, comparison of the observed and photographed trajectories will be undertaken.

REFERENCES

1. S.C. LI, "Spray Stagnation Flames," invited review paper by *Progress in Energy and Combustion Science* (accepted), (1997).

2. S.C. LI, P.A. LIBBY and F.A. WILLIAMS, "Spray Impingement on a Hot Surface in Reacting Stagnation Flows," *AIAA J.* 33, 1046 (1995).

3. S.C. LI, N. ILINCIC and F.A. WILLIAMS, "Reduction of NO_x Formation by Water sprays in Strained Two-Stage Flames," ASME paper No. 96-GT-545, presented at ASME Turbo-Expo '96, Birmingham, England UK, June 10-13, 1996 and accepted for publication in *ASME Journal of Engineering for Gas Turbines and Power* (accepted) (1996). Also *ASME* No. 96-GT-545.

4. S.C. LI, and F.A. WILLIAMS, "Experimental and Numerical Studies of Two-Stage Methanol Flames," *Twenty-Sixth Symposium (International) on Combustion*, The Combustion Institute, Pittsburgh, PA, pp. 1017-1024 (1997).

5. S.C. LI, P.A. LIBBY and F.A. WILLIAMS, "Turbulent Sprays in Stagnation Flows," *Atom. Sprays* 5, 287 (1995).

FILM COOLING IN A PULSATING STREAM

H. Fasel, A. Ortega, I. J. Wygnanski
Department of Aerospace and Mechanical Engineering,
The University of Arizona,
Tucson AZ 85721
USA

ABSTRACT

The mean flow and stability characteristics of a plane, laminar wall jet were investigated experimentally, theoretically, and numerically for a constant wall temperature boundary condition. The streamwise mean velocity and temperature profiles and the downstream development of the hydrodynamic and thermal boundary layer thicknesses were obtained through simultaneous hot and cold wire measurements. Even at relatively low temperature differences, heating or cooling of the surface sufficiently altered the mean velocity profile in the inner region to produce significant effects on the jet stability. Selective forcing of the flow at the most amplified frequencies produced profound effects on the velocity and temperature fields and hence the time-averaged shear stress and heat transfer. Large amplitude excitation of the flow at high frequencies resulted in a **reduction** in the maximum skin friction by as much as 65% with an **increase** in the maximum wall heat flux as high as 45%. The skin friction and wall heat flux were much less susceptible to low frequency excitation.*

NOMENCLATURE

Re_j	Reynolds number measured at exit plane $= \rho U_j d/\mu$	Y_m	maximum normal distance from the wall (for the integration domain), m
T	local mean temperature, K	c_p	specific heat of air at free-stream, J/kg-K
T_o	free-stream temperature, K	d	wall jet slot width, m
T_w	wall temperature, K	k	thermal conductivity, W/m-K
U	local streamwise mean velocity, m/s	$\dot{m}$	mass flow rate, kg/s
U_j	jet exit velocity, m/s	t'	fluctuating temperature, K
U_m	local maximum streamwise velocity, m/s	u'	fluctuating streamwise velocity, m/s
U_o	free-stream velocity, m/s	x,y	coordinates, m

* Parts of this paper are reprinted from ASME Heat Transfer Division Vol. 330, 1996 National Heat Transfer Conf. Vol. 8, pp. 105-113, with permission of the American Society of Mechanical Engineers.

Greek symbols

δ	denotes boundary layer thickness, m
δ_v	local hydrodynamic boundary layer thickness, m
δ_t	local thermal boundary layer thickness, m
μ	free-stream dynamic viscosity, N-s/m^2
ρ	free-stream density, kg/m^3

Subscripts

j	jet exit plane
m	maximum
o	free-stream
t	thermal
v	hydrodynamic
w	wall

INTRODUCTION

The wall jet is a thin jet of fluid introduced tangentially along a surface. The free-stream can either be co-flowing or quiescent. This flow field consists of two primary unstable shear layers that are associated with two different kinds of instability modes: a viscous (inner) mode associated with the inner near-wall region, and an inviscid (outer) mode associated with the inflection point in the outer region. Wall jets have important technological applications such as in de-icing or de-fogging of windshields. In film cooling of gas turbine components, a turbulent wall jet is used to shield blades and other surfaces exposed to high temperature freestream flow. In previous explorations of the wall jet, Katz et al. (1992) and Zhou et al. (1993) found that among other effects, significant reductions in average skin friction can be obtained in a turbulent wall jet by excitation of the flow at its dominant mode. Given the importance of the wall jet in the transport of heat and mass to or from surfaces, it is important to question whether forcing may have similar related effects on the convective heat transfer. The possibility of augmenting or reducing the wall heat transfer, with minimal energy expenditure for excitation, would lead to innovative methods for controlling heat transfer. The wall jet thus provides a fundamental yet practical flow for investigating the possibility of enhancing or suppressing rates of transport by both active and passive flow control.

Theoretical investigations of the incompressible, isothermal, laminar wall jet were first performed by Tetervin (1948) and Glauert (1956) who obtained a closed form solution. Results were substantiated by the hot wire measurements of Bajura and Szewezyk (1970). Cohen et al. (1992) theoretically investigated the effects of blowing and suction on an incompressible laminar wall jet. They found a new family of self-similar solutions in which Glauert's solution is a member. These self-similar solutions were later confirmed experimentally by Amitay & Cohen (1993). The self-similar solutions of the temperature field for a low speed, laminar, incompressible wall jet with constant physical properties were obtained by Schwarz & Caswell (1961) for the special cases of constant wall temperature, varying wall temperature, and constant wall heat flux. Mitachi & Ishiguro (1974) performed a theoretical investigation of the laminar wall jet with different wall thermal boundary conditions. The linear stability of the wall jet was investigated by Chun & Schwarz (1967), Bajura & Szewezyk (1970) , Bajura & Catalano (1975), and Mele et al. (1986). They showed that the initial stages of transition are two-dimensional in nature, and are dominated by the mechanism of vortex pairing, which is commonly observed in free shear flows.

We present recent experimental and numerical results to show the effects of low and high excitation levels on the velocity and temperature fields of a two-dimensional, laminar wall jet flowing over a constant temperature surface. Experiments and numerical computations were performed for a steady, unforced wall jet followed by a wall jet with low amplitude forcing. Surprisingly, it was found that under most circumstances, forcing decreased the time-averaged wall shear stress while the wall heat transfer was increased.

DESCRIPTION OF EXPERIMENTS

The experiments were performed in a thermally controlled, closed return, low speed, air wind tunnel as shown in Fig. 1. The test section was 711 mm wide, 165 mm high and 2,362 mm long. A slot type wall jet

apparatus was used which essentially duplicates the design of Zhou et al. (1993) which was used in the study of a turbulent wall jet with forcing. This apparatus consisted of a 0.5 HP centrifugal blower, acoustic muffler, air-water heat exchanger, diffuser, a plenum chamber fitted with an acoustic speaker, three screens of 30, 40 and 50 mesh size and a contraction nozzle having a variable area ratio. In these experiments the area ratio was 28:1 resulting in a jet slot width of 3.2 mm and a jet exit velocity of 1.5 m/s. The apparatus was designed to fit in the test section floor of an existing closed return wind tunnel which will provide the free-stream flow in future experiments. The flow exited the wall jet apparatus tangentially over an isothermal heat transfer surface. The surface consisted of a 19 mm thick highly polished aluminum tooling plate which was held at constant temperature by means of cooled/heated water from a recirculating chiller that counter flowed through milled slots inside the plate. Surface temperatures were measured with Type K thermocouples that were potted into the surface with an aluminum epoxy. The measurements of the streamwise velocity and temperature were conducted using two side-by-side (7 mm apart) standard Disa Model 55P11 single hot and cold wire probes with a 5 μm diameter tungsten sensor having length-to-diameter ratio of 300. The hot wire was kept at an over-heat ratio of 1.6 and had a maximum frequency response of 30 kHz. The frequency response of the cold wire was 600 Hz. The cold wire was calibrated against a NIST traceable lab standard thermistor probe. The external forcing was introduced using acoustic perturbations produced by a 12 inch speaker cone that was placed in the diffuser of the wall jet apparatus shown. A digital-to-analog converter was used to generate the forcing signal which was passed through an amplifier before it was used to activate the speaker. The forcing frequencies were determined from linear stability theory, and the excitation level measured at the jet exit plane was incrementally varied from 0.6 to 2.0 %. In this investigation the forcing frequencies were fixed at 43.9 Hz and 21.1 Hz corresponding to the inner and outer modes of instability. In order to retain phase information, the phased locked technique was used in which the hot wire, the cold wire and the function generator records, consisting of 256 points each, were digitized simultaneously and saved. At each measuring point, up to 800 events were recorded.

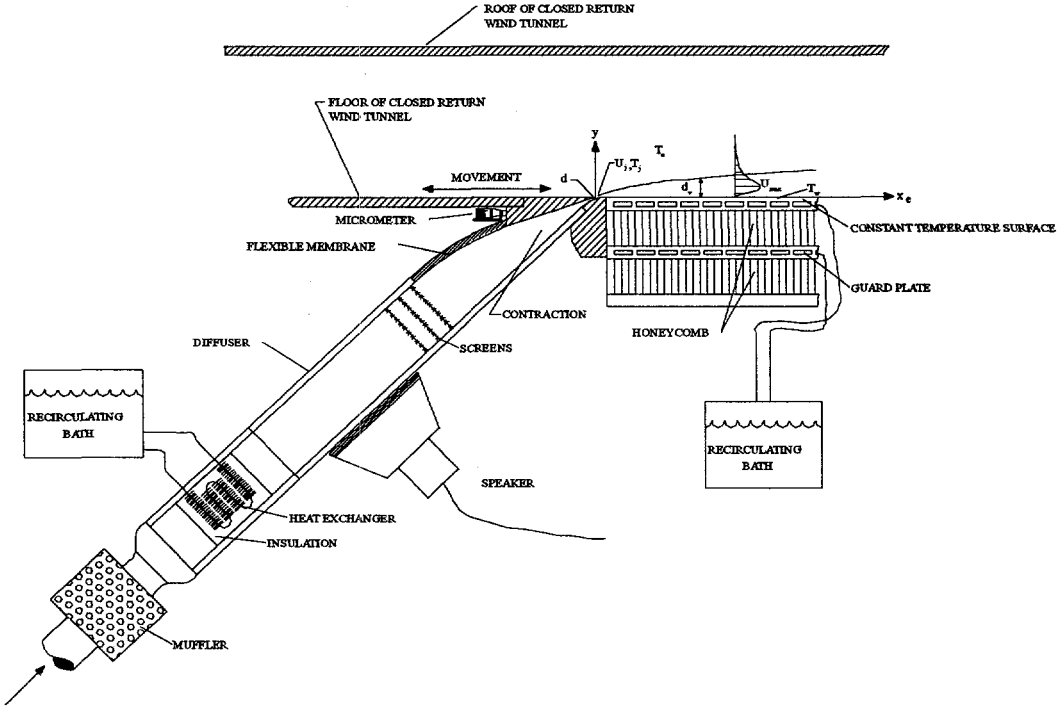


Figure 1: The experiment apparatus.

EXPERIMENTAL RESULTS

Undisturbed Mean Flow

Three sets of experiments were conducted corresponding to three temperature ratios denoted by T_w / T_o: wall cooling with $T_w / T_o = 0.98$, wall heating with $T_w / T_o = 1.03$ and the neutral case with $T_w / T_o = 1$, in which the surface, jet exit, and free-stream temperatures were all the same. Throughout the experiments, the jet exit Reynolds number, $\mathrm{Re}_j = \rho U_j d / \mu$, remained fixed at 320. The normalized mean profiles of the streamwise velocity measured for the three cases above and at three normalized streamwise locations ($x/d =$ 12, 16, and 20) are shown in Fig. 2. The local maximum velocity, U_m, and the local thickness of the hydrodynamic boundary layer, δ_v (where the mean velocity is equal to $U_m / 2$ in the outer region), were used to render all variables dimensionless. Figure 2a shows full velocity profiles for all cases, while Figs. 2b, 2c, and 2d present the inner region in more detail for the cooling, neutral, and heating cases, respectively. The symbols represent the experiments while the solid line corresponds to Glauert's (1956) theoretical prediction. The agreement between the experiments and the theory for the neutral case is good except in the inner region where the velocity gradient of the experimental data is higher, Fig. 2c. This may be due to the relatively high turbulence level (0.5%) of the jet. When cooling is applied, Fig. 2b, the inner velocity gradient increases and the maximum velocity is shifted closer to the surface. Heating the surface, Fig. 2d, appears to have the opposite effect, but the differences in the heated case compared to the neutral case cannot be claimed to be outside of the range of experimental uncertainty. Despite these small changes in the mean velocity due to heating and cooling, a more significant effect is observed on the fluctuating streamwise velocity distribution.

Mean Flow Field Subjected to External Forcing

The mean velocity and temperature fields were next investigated under conditions with externally applied forcing. Disturbances were introduced into the flow at various frequencies and amplitudes. Three cases, corresponding to the above three values of T_w / T_o, were studied. In this section only the results for the heating case are presented. The dimensional mean profiles of the streamwise velocity measured for the heating case and at three normalized streamwise locations ($x/d =$ 12, 16, and 20) are shown in Fig 3. Figure 3a presents the case in which the flow is excited with a high frequency disturbance (approximately 44Hz for inner mode forcing), while Fig 3b corresponds to a low frequency disturbance (21Hz for outer mode forcing.) The excitation level is defined at $x/d = 0$ as $100 \cdot u'_m / U_j$, where, u'_m is the maximum fluctuating streamwise velocity (measured by using the phase-locked data technique) and U_j is the jet mean exit velocity. As is shown in Fig 3a, when inner mode forcing is applied the velocity field is dramatically altered. The effects become more pronounced as the downstream distance increases. The effect of outer mode forcing is not as dramatic because the growth rates associated with the outer mode are smaller than the inner mode growth rates, as predicted by the theory (see Amitay & Cohen (1996)). In order to see the effects of forcing on the wall shear stress, the near wall region of the velocity field was plotted and is presented in Fig 4. Seven cases, corresponding to the unforced case, inner and outer mode forcing with 0.6%, 1% and 2% excitation levels, are presented. Figures 4a, 4b and 4c represent the experimental data taken at $x/d =$ 12, 16, and 20, respectively. The solid line corresponds to the velocity gradient for the unforced case. For low levels of excitation, the velocity gradient near the wall is negligibly affected. As the excitation level is increased to 2% the velocity gradient near the wall is decreased by up to 15% (with respect to the unforced case) for the outer mode forcing, and by up to 65% for the inner mode forcing.

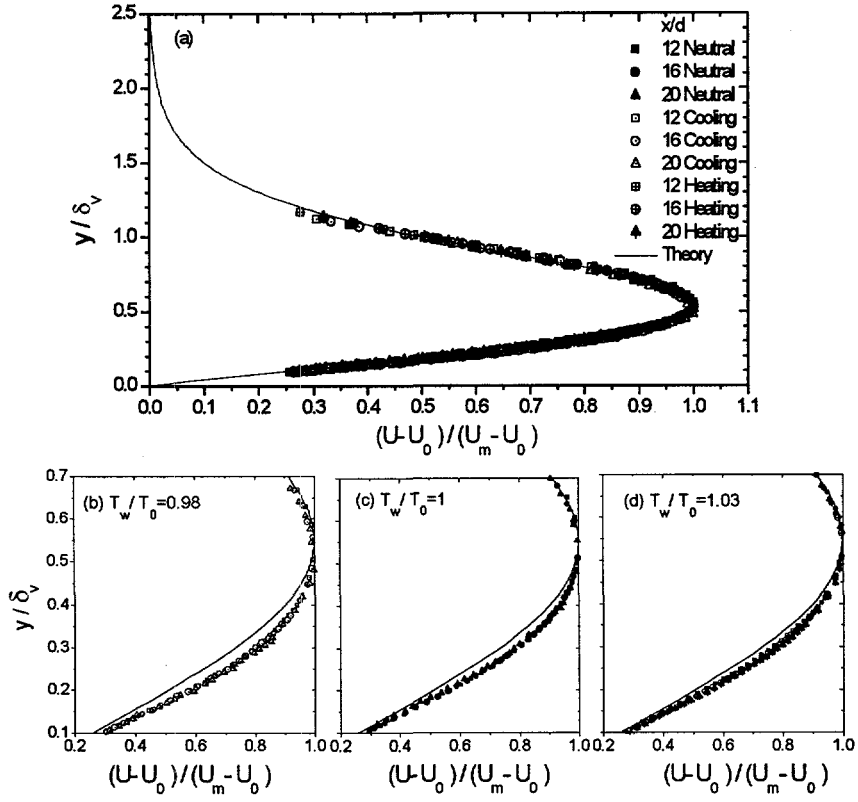


Figure 2: Comparison of normalized velocity profiles at three x/d locations and for T_w/T_o = 0.98, 0 and 1.03 (a) full profile, (b), (c) and (d) near-wall region

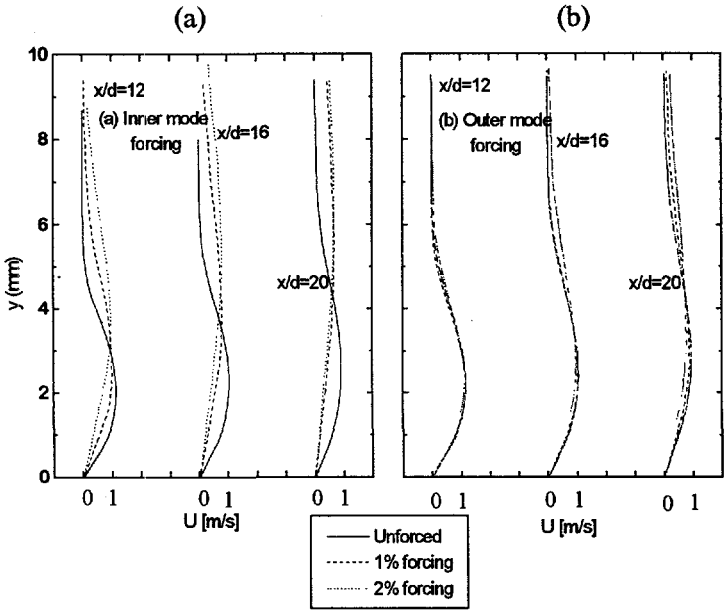


Figure 3: Dimensional mean velocity profiles for various excitation levels and downstream locations (a) inner mode forcing and (b) outer mode forcing for T_w/T_o = 1.03.

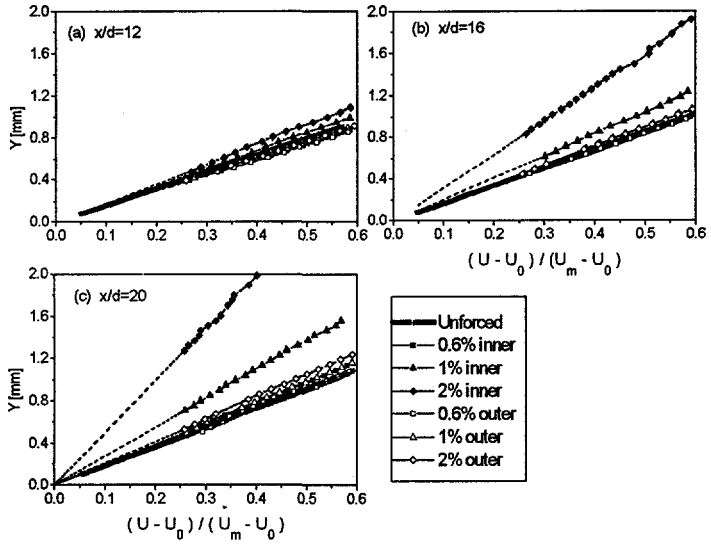


Figure 4: Near-wall velocity gradient comparison for different excitation levels and frequencies at x/d of (a) 12, (b) 16 and (c) 20 and for $T_w/T_o = 1.03$.

The normalized mean temperature profiles measured for the heating case and at x/d =16 are shown in Fig 5a and 5b for inner and outer mode forcing, respectively. As was observed for the velocity field, for low excitation levels the effect of forcing on the temperature profile is negligible. As the excitation level is increased, surprisingly, the temperature gradient near the wall *increases* until an inflection point is formed on the temperature profile. Again, the effect of the inner mode forcing is much more significant. The effect of forcing on the near-wall region of the temperature profile is presented in Fig 6. Low excitation levels do not alter the temperature gradient near the wall, while increasing the excitation level causes an increase in the temperature gradient. Outer and inner mode forcing at 2% excitation level increases the temperature gradients by up to 20% and 45%, respectively. The augmentation in the heat transfer is accompanied by a net reduction in the time-averaged wall drag, a result which could not be anticipated.

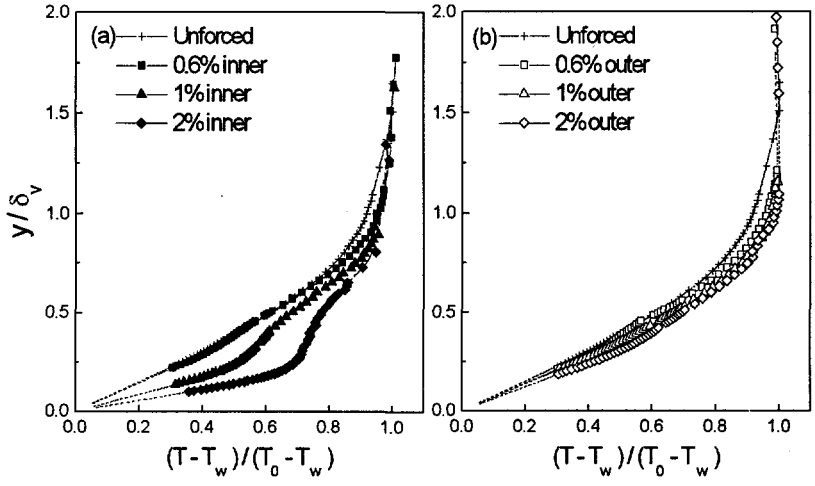


Figure 5: Comparison of normalized temperature profiles measured at x/d = 20 and for different excitation levels for (a) inner mode forcing and (b) outer mode forcing for $T_w/T_o = 1.03$.

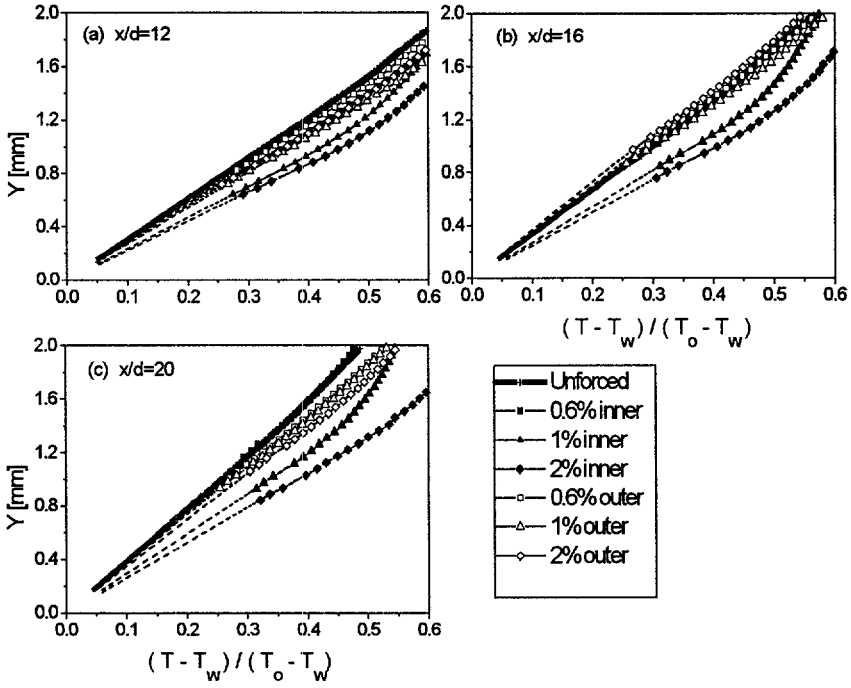


Figure 6: Near-wall temperature gradient comparison for different excitation levels and frequencies at x/d of (a) 12, (b) 16 and (c) 20 and for $T_w/T_o = 1.03$.

As shown in Fig 5 high levels of forcing introduce non-linear effects as observed by a dramatic thickening of the boundary layer. Large scales associated with the outer region become more significant. As was shown by Bajura & Catalano (1975) and Amitay & Cohen (1996) two rows of counter rotating vortices associated with the two regions of the wall jet are present in laminar to turbulent transition. It is the outer row of vortices (which rotate in the counter-clockwise direction and thus transport momentum away from the wall) that receive most of the energy under high levels of forcing and therefore, decrease the velocity gradient near the wall. We hypothesize that this same motion is responsible for the increase in the near-wall temperature gradient which was observed in Figs. 5 and 6. It is plausible that the large vortices in the outer region transport fluid away from the wall to the outer region (which is at a different temperature) and also entrain fluid towards the wall. This motion may also explain the large reduction in the temperature gradient, occurring at y/δ_v of about 0.2 as seen in Fig 5, because the unsteady upward advective transport is increased compared to diffusive transport in this region. This enhanced thermal mixing is manifested also in the observed increase in the wall temperature gradient and thus the wall heat flux.

Fluctuating Quantities

Since the evolution of controlled disturbances apparently is a primary factor in augmenting the heat transfer, the structure and evolution of the temperature and streamwise velocity fluctuations were investigated in detail. In order to obtain phase information, phase-locked data were measured and analyzed for low and high levels of excitation corresponding to the two modes of instability associated with the wall jet flow. These levels of excitation correspond to the linear and nonlinear transitional behavior of the flow. Figure 7 shows the dimensional fluctuating temperature distribution versus the dimensional distance from the wall. The data were taken using the conventional phase-locked technique, and were measured at x/d =16 for T_w/T_o = 1.03 and for three excitation levels of (a) 0.6%, (b) 1%, and (c) 2%. A single peak is present for 0.6% forcing, as is shown in Fig 7a. This peak is located at the same distance from the wall for

both inner and outer mode forcing. The magnitude of this peak for inner mode forcing is higher due to higher initial growth rates associated with this mode. As the level of forcing is increased, a second peak is formed on the distribution very close to the wall. This second peak occurs at the same location as the inflection point in the mean temperature profile, and is most likely due to a maximum in the unsteady normal velocity v' occurring at this location. Also, the width of the disturbed region is much thicker for inner mode forcing due to the increase of the boundary layer thickness. Similar trends occurred when the wall was cooled.

Figure 8 shows the normalized distribution of the fluctuating fundamental streamwise velocity for an excitation level of 0.6% and for both heating (a) and cooling (b) for outer mode forcing. The distribution contains two large peaks corresponding to the velocity gradient near the wall and the inflection point in the outer region, as was explained by Mele et al. (1986). The outer peak is initially larger and has a broader disturbed region than the inner peak because the above mentioned forcing frequency prevails in the outer region. In both cases, as the downstream distance increases, the ratio between the inner and outer peaks is increased which is consistent with linear stability theory. According to the linear stability theory, as the downstream distance is increased smaller scales governed by the near wall region become more dominant. For a fixed downstream location, the ratio between the inner and outer amplitude maxima is larger for the cooling case. In order to further understand the effect of forcing on the velocity field, the effect of the level of excitation was also addressed. In Fig 9, two cases of inner mode forcing corresponding to (a) 0.6% and (b) 2% are presented for three downstream locations. δ_v was used to normalized the normal distance from the wall which is consistent with the stability theory. The inner peak is significantly higher, as expected, because the effect of inner mode forcing is most dominant in the inner region. For the low excitation level of 0.6%, the ratio between the inner to outer maxima increases with the downstream distance as was observed when outer mode forcing was applied in Fig 8. This is once again consistent with linear stability theory. However, for the 2% forcing the trend is the opposite, with the ratio between the inner to outer maxima decreasing with downstream location. This is probably due to the fact that for high excitation levels non-linear mechanisms are introduced. At any given x/d, the boundary layer is much thicker and therefore large scales, corresponding to low frequencies, prevail throughout the flow field. As was mentioned before, when the forcing frequency is low, the outer maxima dominates the fluctuating streamwise velocity distribution since it corresponds to the large scales.

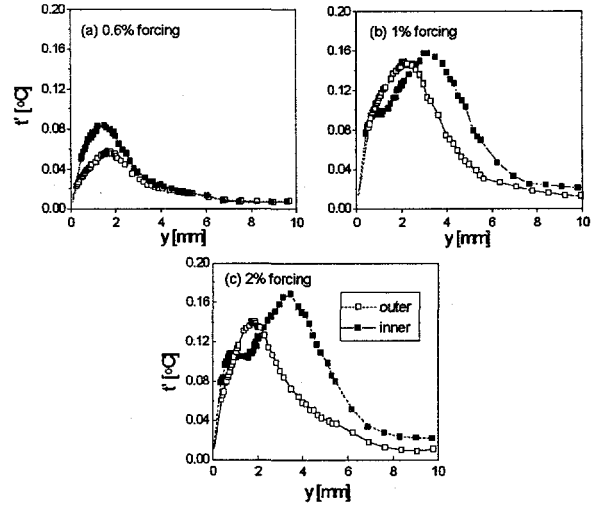


Figure 7: Fluctuating temperature distribution for inner and outer mode forcing for (a) 0.6%, (b) 1% and (c) 2% excitation levels (x/d = 16 and T_w/T_o = 1.03).

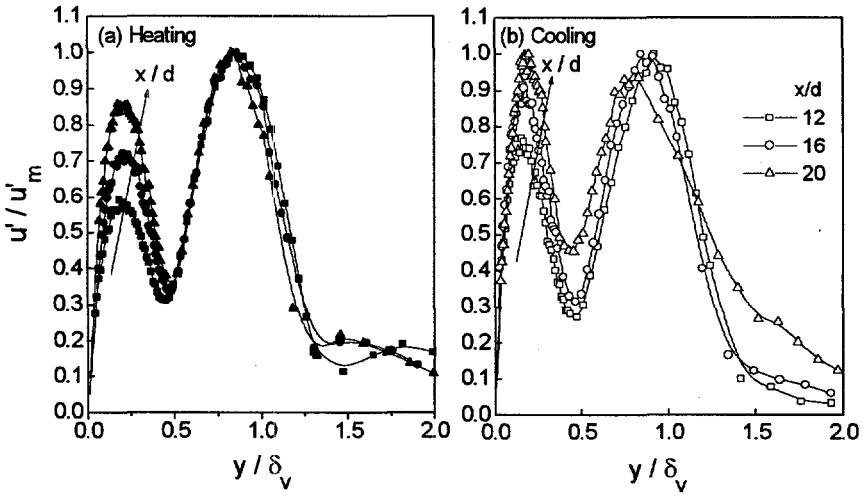


Figure 8: Normalized fluctuating streamwise velocity comparison at three x/d locations and for outer mode forcing at excitation level of 0.6%, (a) wall heating and (b) wall cooling.

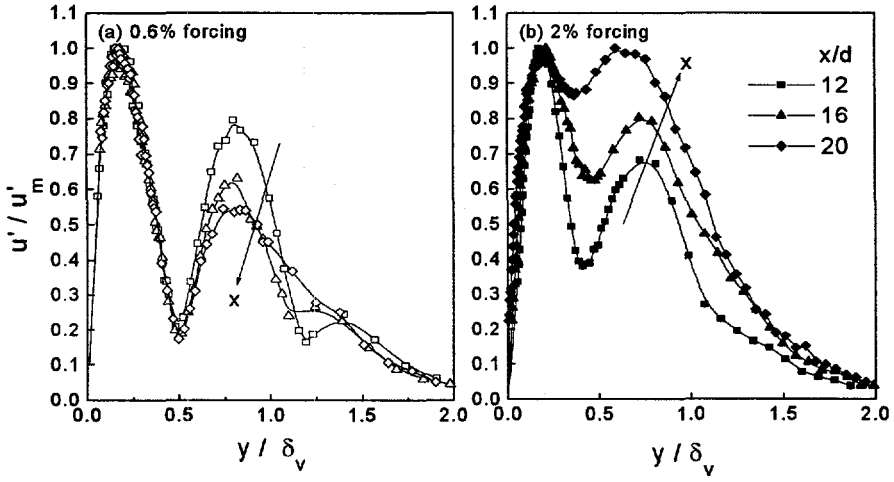


Figure 9: Normalized fluctuating streamwise velocity comparison at three x/d locations and for inner mode forcing, (a) 0.6% and (b) 2% excitation levels and $T_w/T_o = 1.03$.

NUMERICAL SIMULATIONS

Numerical Methods

In parallel to the experimental investigations, Direct Numerical Simulations (DNS) were conducted based on the complete Navier-Stokes equations. For these simulations, two computer codes were employed that were developed for investigating transition in boundary layers, both for compressible and incompressible flows. Both codes, compressible and incompressible, were modified and adapted to the wall jet geometry. In the compressible Navier-Stokes code, the complete Navier-Stokes equations, the energy equation, and the continuity equation are solved in conservative formulation. The fluid is modeled as a calorically perfect gas and Sutherlands Law is used to calculate the viscosity from the temperature. The

time integration was performed using a 4th-order Runge-Kutta method. For the spatial discretization a 4th order upwind-downwind "splitting" scheme was used. For the results shown here, an equidistant grid in the x- and y-direction was used. In the spanwise direction, a pseudo-spectral approach was chosen. In the incompressible Navier-Stokes code the Navier-Stokes equations are solved in the vorticity-velocity formulation in conjunction with the energy equation. A fourth-order polynomial is used for the temperature dependence of viscosity and thermal conductivity. As for the compressible code, time integration is performed using a 4th-order Runge-Kutta method. The streamwise and wall normal directions are discretized using 4th-order split differences. The solution procedure of the viscous terms in the vorticity transport equations was extensively modified to facilitate the introduction of a subgrid-scale (SGS) turbulence model to be used for Large-Eddy-Simulations. A equidistant grid is used in the streamwise direction, whereas in the wall normal direction, a stretched grid is used where points are clustered near the wall.

Computational Results

Before applying the codes to extensive simulations of our laboratory experiments, the codes were extensively tested and validated to demonstrate their ability and efficiency of obtaining results. In both the compressible and incompressible simulations, disturbances are introduced into the flow field by periodic blowing and suction through a slot in the wall near the inflow boundary (see Fig. 10). For this forcing, the stream wise distribution of the disturbance velocity is such that the net mass flux is zero. Therefore, predominantly vorticity disturbances of a specified frequency are introduced. Using these two entirely different numerical approaches, we can validate our calculations by comparing the results of the compressible and incompressible simulations, even in flow regimes that are outside the scope of Linear Stability Theory and our experiments. Both codes (compressible and incompressible) were extensively tested by comparison calculations with linear stability theory. These calculations showed that the codes are efficient and reliable. For calculating the base flow, the similarity solution for the appropriate nondimensional temperature difference between wall and ambient is used as initial condition. Using pseudo-time stepping, the calculation is continued until a given convergence criterion is satisfied. Using these base flows, numerous unsteady simulations were performed for an isothermal, cooled and heated wall jet. With the compressible code, all calculations were done for a global Mach number of M=0.3, based on the maximum velocity at the inflow and the speed of sound in the ambient fluid. This Mach number allows for reasonably large time steps and ensures that compressibility effects are still negligible. A comparison of the amplitude and phase development over the streamwise coordinate x is shown in Fig. 11. The agreement of the compressible and incompressible simulations is excellent.

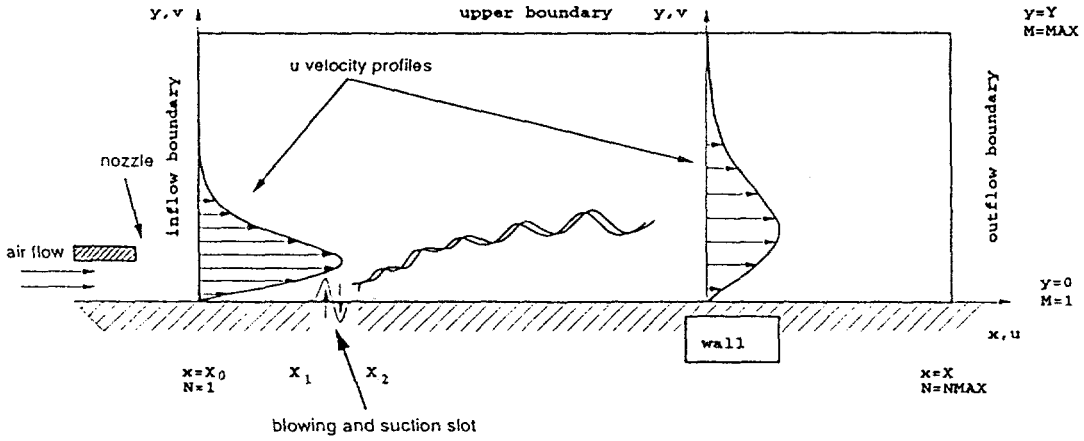


Figure 10. Computational domain for numerical simulations.

In Fig. 11, there is a slight deviation of the numerical results compared to the results obtained from Linear Stability Theory. This difference is due to the fact that Linear Stability Theory can only take local effects into account, whereas our simulations follow the development of the disturbances downstream of the blowing and suction slot. In this particular case, the slot was located upstream of the unstable region of the shear layer mode (outer mode), so that this mode was amplified first. Although it is damped in the region shown, it still has a finite amplitude and it takes some downstream distance until its influence on the velocity maximum is negligible compared to the boundary layer mode. From this point on, the agreement of the results from the numerical simulations and Linear Stability Theory is excellent. Figure 12 shows the eigenfunctions of the streamwise- and wall normal velocity, u and v. Again, the results of the compressible and incompressible simulations show excellent agreement. These results indicate that with our calculations, we are able to capture the relevant physical mechanisms in the flow. Furthermore, we are convinced that the codes are usable for simulations of highly nonlinear flow configurations. For a detailed comparison with the ongoing laboratory experiments, a new set of simulations was started for the case with weak wall heating, using the incompressible code. To account for the effect of the unheated starting length that was found in the experiments, the concept of the artificial Prandtl number (discussed in an earlier report) was utilized and we were able to obtain very good agreement of the experimental and computational mean flow profiles. The linear stability diagram for the case with a heated wall (T_W=32.5 C, T_O=24 C, F=43.5Hz) is shown in Fig. 13. The solid line indicates the Reynolds number/wave number pairs at which the experiments were conducted. Using small amplitude disturbances, the results of experiments and calculations can be compared with Linear Stability Theory. The baseflow u velocity and temperature profiles, as well as the eigenfunctions u' and t', at x/d=20 are shown in Fig. 14. The calculations are in very good agreement with the experiments.

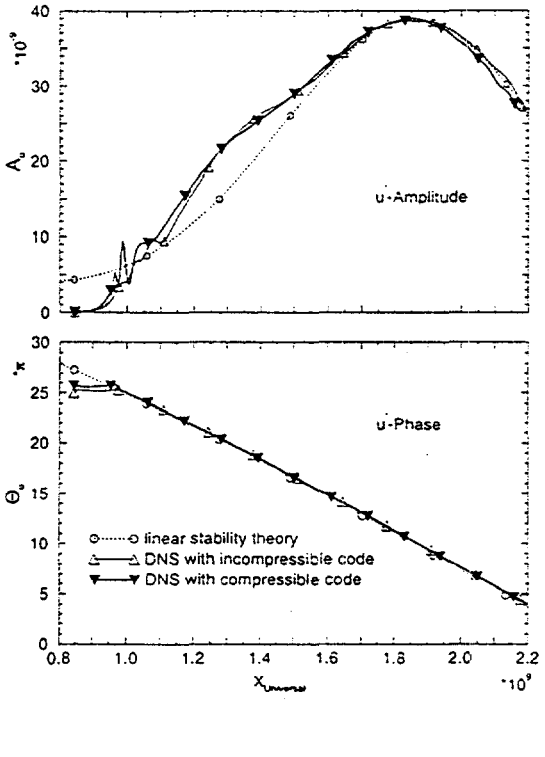


Figure 11. Downstream development of u'-amplitude and phase u'. Comparison of results obtained from compressible and incompressible code with linear stability theory.

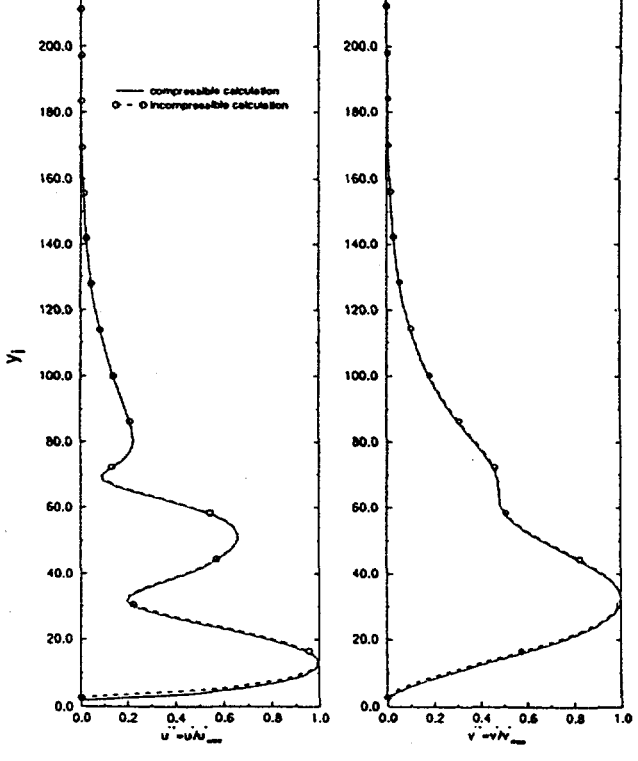


Figure 12. Comparison of amplitude distributions (eigenfunctions) obtained for compressible and incompressible code (a) u'-amplitude, (b) v' amplitude

Using larger forcing amplitudes, large changes in the mean flow profiles due to nonlinear effects can be observed (see Figs. 3 and 5). The shear stress at the wall is reduced and the heat transfer from the wall is increased, in qualitative agreement with the experiments. A quantitative comparison is very difficult because of the high turbulence level (0.5%) in the experimental facility. This "noise" introduces energy into the flow that cannot easily be taken into account in the numerical simulations. For investigations of turbulent wall jets, codes for compressible and incompressible Large Eddy Simulations (LES) are being developed. The incompressible code is currently being used for the calculation of transitional and turbulent flat plate boundary layers with zero and strong adverse pressure gradients. We are confident that the turbulence model used in these simulations can be adapted with only minor modifications for the investigations of turbulent wall jets.

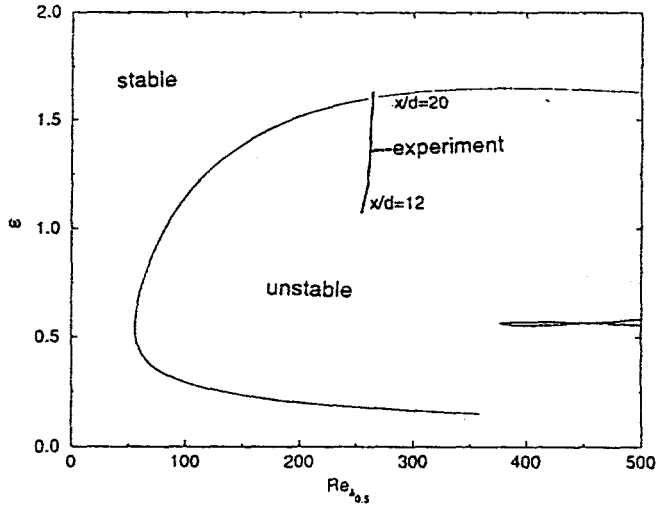


Figure 13. Stability diagram of linear stability theory for wall jet with wall temperature T_w; solid line corresponds to downstream domain of experiments.

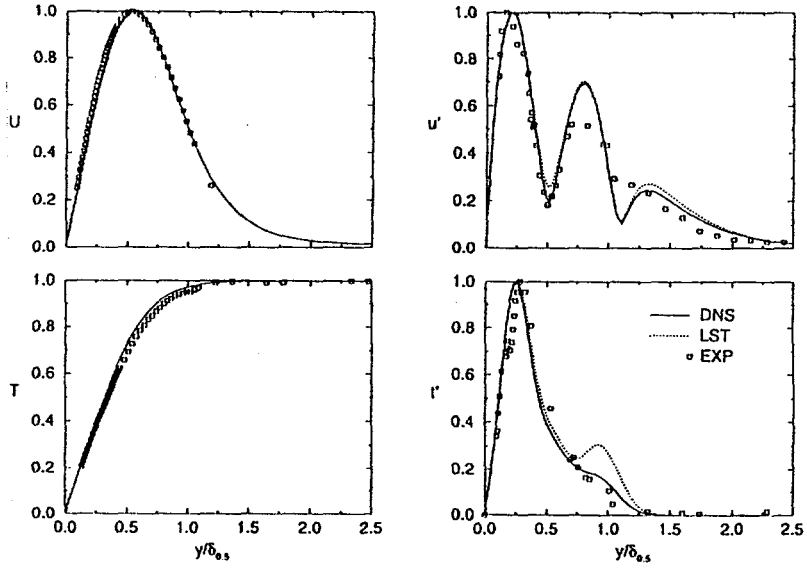


Figure 14. Comparison of experimental measurements, linear stability theory, and Direct Numerical Simulations for Tw=32.5 C, To=24 C, F=43.5 Hz, x/d=20. (a) u-velocity, (b) temperature, (c) amplitude distributions (eigenfunction) for u' fluctuations, (d) amplitude distributions for t' fluctuations.

CONCLUSIONS

In order to better understand the heat and momentum transfer associated with complex unsteady film-cooling problems, the plane, laminar, wall jet was investigated experimentally, theoretically, and numerically for the constant wall temperature boundary condition and with no freestream flow. Measured streamwise mean flow velocity profiles for the unforced case for small temperature differences (6-8 C) were compared to the theoretical predictions. When cooling was applied, the inner velocity gradient was the highest and the maximum velocity was shifted closer to the surface, while heating the surface had the opposite effects. Disturbances were introduced at excitation levels of 0.6%, 1% and 2% and at two forcing frequencies corresponding to the inner and outer modes associated with the wall jet. High levels of forcing (1% and 2%) greatly altered the velocity and temperature profiles. The most significant effects occurred for inner mode forcing where the total skin friction drag (integrated effect) was *reduced* by 45% and the total heat flux was *increased* by 20%. The skin friction and wall heat flux were less susceptible to outer mode excitation but the trends were not altered. The same results occurred for both heating and cooling of the surface. Inner mode forcing had the most significant effects on the velocity and temperature fields. The initial spatial growth rates associated with the inner mode are greater than the outer mode growth rates. We hypothesize that it is the outer row of counter-rotating vortices that receive most of the energy under high levels of forcing and therefore, *decrease* the velocity gradient near the wall. These large vortices enhance thermal mixing and produce the observed *increase* in the temperature gradient near the wall.

ACKNOWLEDGMENTS

This work was supported, in part, by the United States Department of Energy under grant number DE-FG03-93ER14396, the Air Force Office of Scientific Research under grant number F49620-94-1-0131 and NASA Lewis Research Center under the Graduate Student Researchers Program, grant number NGT-70403#1.

REFERENCES

1. Amitay, M. and Cohen, J., 1993, "The Mean Flow of a Laminar Wall Jet Subjected to Blowing or Suction", *Phys. Fluids*, A 5, 2053.
2. Amitay, M. and Cohen, J., 1996, "Transition in a Two-Dimensional Plane Wall Jet Subjected to Blowing or Suction", submitted to *J.Fluid Mech.*
3. Bajura, R. A. and Catalano, M. R., 1975, "Transition in a Two-Dimensional Plane Wall Jet", *J. Fluid Mech.*, Vol. 70, part 4, 773-799.
4. Bajura, R. A. and Szewezyk, A., 1970, "Experimental Investigation of a Laminar Two-Dimensional Plane Wall Jet", *Phys. Fluids*, **13**, 1653.
5. Chun, D. H. and Schwarz, W. H., 1967, "Stability of the Plane Incompressible Viscous Wall Jet Subjected to Small Disturbances", *Phys. Fluids*, **10**, 5, 911-915.
6. Cohen J., Amitay M. and Bayly, B.J., 1992, "Laminar-Turbulent Transition of Wall Jet Flows Subjected to Blowing and Suction", *Phys. Fluids* A 4, 283.
7. Glauert, M.B., 1956, "The Wall Jet", *J.Fluid Mech*, 1, 625-643.
8. Katz, Y., Horev, E. and Wygnanski, I., 1992, "The Forced Turbulent Wall Jet", *J.Fluid Mech*, 242, 577-609.
9. Kline, S. J., and McClintock, F. A., 1953, "Describing Uncertainties in Single Sample Experiments," *Mechanical Engineering*, Vol. 75, Jan., pp. 3-8.
10. Mele, P., Morganti, M., Scibillia, M. F. and Lasek, A., 1986, "Behavior of Wall Jet in Laminar-to-Turbulent Transition", *AIAA Journal*, **24**, 938-939.

11. Mitachi, K & Ishiguro, R., 1974, "Heat Transfer of Wall Jets", *Heat Transfer: Japanese Research*, 27-40.
12. Riley, N., 1958, "Effects of Compressibility on a Laminar Wall Jet", *Jornal of Fluid Mechanics*, **4**, 615-628.
13. Tetervin, N., 1948, "Laminar Flow of a Slightly Viscous Incompressible Fluid that Issues from a Slit and Passes over a Flat Plate", NACA TN, No. 1644.
14. Schwarz W. H. and Caswell B., 1961, "Some Heat Transfer Characteristic of the Two-Dimensional Laminar Incompressible Wall Jet", *Chemical Engineering Science*, **16**, 338-351.
15. Schwarz W. H. and Caswell B., 1961, "Some Heat Transfer Characteristic of the Two-Dimensional Laminar Incompressible Wall Jet", *Chemical Engineering Science*, **16**, 338-351.
16. Zhou, M. D., Heine, C. and Wygnanski, I., 1993, "The Forced Wall Jet in an External Stream", *AIAA* 93-3250.

KINETIC THEORY AND LONG RANGE CORRELATIONS IN MODERATELY DENSE GASES

T. Petrosky and I. Prigogine

Center for Studies in Statistical Mechanics and Complex Systems,
The University of Texas at Austin, Austin, TX 78712 USA.

ABSTRACT

The complex spectral representation of the Liouville operator is applied to moderately dense gases interacting through hard-core potentials in arbitrary d-dimensional spaces. It is shown that Markovian kinetic equations exist for all d. This provides an answer to the long standing question do kinetic equations exist in two dimensional systems. The non-Markovian effects, such as the long-time tails for arbitrary n-mode coupling, are estimated by superposition of the Markovian evolutions in each subspace as introduced in our spectral decomposition. The long-time tail effects invalidate the traditional kinetic theory based on a truncation of BBGKY hierarchy for $d < 4$, as well as the Green–Kubo formalism, as there appear contributions of order t^{-1}, $t^{-1/2}, \cdots$ coming from multiple mode-mode couplings even for $d = 3$.

INTRODUCTION

In spite of much effort the kinetic theory of moderately dense gases faces still a number of difficulties. Does there exist kinetic Markovian equations for two dimensional (2d) systems, such as for example, hard disks? What is the effect of the long time tails on the linear response theory, the so-called Green-Kubo formalism (For an excellent and still up-to-date review of these problems, see ref. [1], as well as the book by Résibois and de Leener [2]. This book will be quoted as *RL*. See also refs. [3-5]). It is well known that reduced distribution functions satisfy master equations which include memory effects (they are "non-Markovian") [6]. Moreover, these equations contain still the full complexity of N-body problem and are mostly formal. We want to show that our recent work on the spectral decomposition of the Liouville operator [7] † provides an answer to the questions we mentioned. Using our spectral decomposition the non-Markovian evolution can be split rigorously into independent Markovian evolutions. In line with earlier work (see *RL Ch.*X) we shall consider the case of "hard spheres" in an arbitrary number d of dimensions. We shall show that kinetic Markovian equations exist for all d. Moreover, we can estimate the long-time tails for arbitrary n-mode coupling and show that they lead to divergences of the Green–Kubo integrals for $d \leq 4$.

† This will be quoted as *I*. For quantum mechanics see ref. [8].

COMPLEX SPECTRAL REPRESENTATIONS

We start with a short summary of our method (see I). The Liouville equation for the N-particle distribution function (d.f.) $\rho(\mathbf{r}^N, \mathbf{v}^N, t)$ is $i\partial\rho/\partial t = L\,\rho$. The Liouvillian L consists of a free Liouvillian L_0 and an interaction part δL, i.e., $L = L_0 + \delta L = \sum_{a=1}^{N} L_0^{(a)} + \sum_{b>a=1}^{N} \delta L^{(ab)}$ where $L_0^{(a)} = -i\mathbf{v}_a \cdot (\partial/\partial \mathbf{r}_a)$. Our method deals with the class of ensembles corresponding to the thermodynamic limit (i.e., the number of particle $N \to \infty$, and the volume $V \to \infty$, with the number density $n = N/V$ finite). This class includes the canonical distribution. It describes "persistent" interactions. The eigenfunctions of L_0 are plane waves ($\sim \exp[i\sum_j \mathbf{k}_j \cdot \mathbf{r}_j]$). The distribution function ρ therefore can be decomposed into contributions from different degrees of correlations, $\rho = P^{(0)}\rho + P^{(1)}\rho + P^{(2)}\rho + \cdots$, where $P^{(\nu)}$ is the projection operator which retains the ν-th degree of correlations (see I). For example, $P^{(0)}$ retains the "vacuum component" $\rho_0(\mathbf{v}^N)$, which is the velocity d.f. with vanishing wavevectors for all particles $\mathbf{k}^N = 0$, $P^{(1)}\rho$ the "inhomogeneity components" $\rho_{\mathbf{k}_a}(\mathbf{v}^N)$ with $\mathbf{k}_a \neq 0$ and $\mathbf{k}^{N-1} = 0$ for particles $a = 1$ to N, $P^{(2)}\rho$ the "binary correlation components" $\rho_{\mathbf{k}_a,\mathbf{k}_b}$ with $\mathbf{k}_a \neq 0$ and $\mathbf{k}_b \neq 0$, and so on. We have $L_0 P^{(\nu)} = P^{(\nu)} L_0$, $P^{(\nu)}P^{(\mu)} = P^{(\nu)}\delta_{\nu,\mu}$ and $\sum_\nu P^{(\nu)} = 1$.

As we have shown (see $(I.45)$), the Fourier coefficients with a smaller number of nonvanishing components in wavevectors $\mathbf{k}^N$ have a higher order dependence in the volume factor V in the large volume limit, for example $\rho_0/\rho_{\mathbf{k}_a} \sim V$. The appearance of these delta-function singularities leads to the usual cluster expansion of the reduced distribution functions. Also, because of these singularities, the d.f. ρ *does not* belong to the Hilbert space, as its Hilbert norm vanishes in the thermodynamic limit, $\langle\rho|\rho\rangle \equiv \int d^N x \rho^*(x^N)\rho(x^N) \to 0$, where $x^N = (\mathbf{r}^N, \mathbf{v}^N)$. As a result, the hermitian operator L acquires *complex* eigenvalues $Z_j^{(\nu)}$ breaking time-symmetry. The time evolution of the system then splits into two semigroups. For $\mathrm{Im}\, Z_j^{(\nu)} \leq 0$ equilibrium is approached in our future that is for $t \to +\infty$, while for $\mathrm{Im}\, Z_j^{(\nu)} \geq 0$ equilibrium had been reached in our past. The domain of the two semigroups do not overlap (see I and ref. [8]). To be self-consistent we choose the semigroup oriented towards our future. Then we have the complex spectral representation,

$$e^{-iLt}|\rho\,(0)\rangle = \sum_\nu \sum_j |F_j^{(\nu)}\rangle e^{-iZ_j^{(\nu)}t} \langle \tilde{F}_j^{(\nu)}|\rho\,(0)\rangle \tag{1}$$

where $F_j^{(\nu)}$ and $\tilde{F}_j^{(\nu)}$ are bi-orthonormal sets of right- and left-eigenstates of the Liouvillian with the complex eigenvalue $Z_j^{(\nu)}$, respectively. This spectral decomposition has to be used in conjunction with observables depending on a finite number of variables (see $(I.67)$).

The eigenstates of L are given by $|F_j^{(\nu)}\rangle = n_{\nu,j}[P^{(\nu)} + Q^{(\nu)}\mathcal{C}^{(\nu)}(Z_j^{(\nu)})]|u_j^{(\nu)}\rangle$ and $\langle\tilde{F}_j^{(\nu)}| = n_{\nu,j}\langle\tilde{v}_j^{(\nu)}|[P^{(\nu)} + \mathcal{D}^{(\nu)}(Z_j^{(\nu)})Q^{(\nu)}]$ (see $(I.98)$), where $Q^{(\nu)} = 1 - P^{(\nu)}$ and $n_{\nu,j}$ is a normalization constant. The "creation operator" $\mathcal{C}^{(\nu)}(z)$ and the "destruction operator" $\mathcal{D}^{(\nu)}(z)$ are defined by $\mathcal{C}^{(\nu)}(z) \equiv G_Q^\nu(z) L P^{(\nu)}$ and $\mathcal{D}^{(\nu)}(z) \equiv P^{(\nu)} L G_Q^\nu(z)$ with the propagator $G_Q^\nu(z) = Q^{(\nu)}[z - Q^{(\nu)}LQ^{(\nu)}]^{-1}$.

The states $|u_j^{(\nu)}\rangle$ and $\langle\tilde{v}_j^{(\nu)}|$ are right- and left-eigenstates of the "collision operators" $\psi^{(\nu)}$, i.e., $\psi^{(\nu)}(Z_j^{(\nu)})|u_j^{(\nu)}\rangle = Z_j^{(\nu)}|u_j^{(\nu)}\rangle$ and $\langle\tilde{v}_j^{(\nu)}|\psi^{(\nu)}(Z_j^{(\nu)}) = \langle\tilde{v}_j^{(\nu)}|Z_j^{(\nu)}$ where $\psi^{(\nu)}$ is given by

$$\psi^{(\nu)}\,(z) = P^{(\nu)} L P^{(\nu)} + P^{(\nu)} L \mathcal{C}^{(\nu)}\,(z)\, P^{(\nu)} \tag{2}$$

The collision operators are dissipative operators, and they are central objects of non-equilibrium statistical mechanics (see I). The operators $\mathcal{C}^{(\nu)}$, $\mathcal{D}^{(\nu)}$, and $\psi^{(\nu)}$ are associated to graphs such as represented for $\nu = 0$ in Fig. 1.

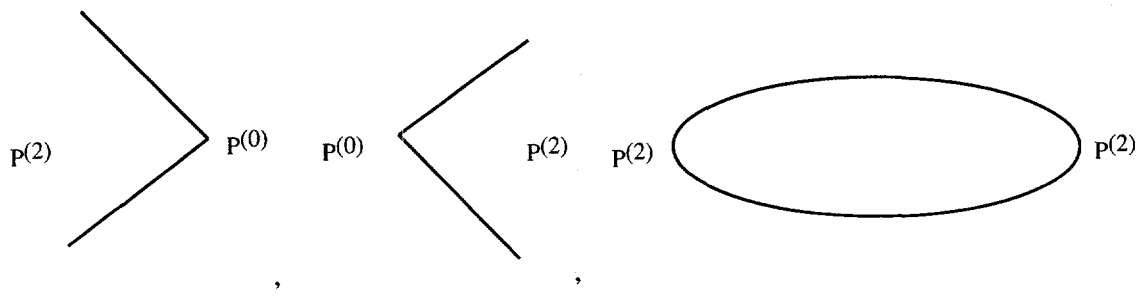


Figure 1: Diagram (a) corresponds to the lowest order contribution of the creation operator $\mathcal{C}^{(0)}$, (b) to the destruction operator $\mathcal{D}^{(0)}$, and (c) to the collision operator $\psi^{(0)}$.

Note that the eigenvalue problem associated to the collision operators $\psi^{(\nu)}(Z_j^{(\nu)})$ is nonlinear as the eigenvalue $Z_j^{(\nu)}$ appears inside the operators. This will play an important role. Also the above relations show that the Liouville operator shares the same eigenvalues with the collision operators. Assuming bi-completeness in the space $P^{(\nu)}$, we may always construct a set of states $\{\langle\tilde{u}_j^{(\nu)}|\}$ bi-othogonal to $\{|u_j^{(\nu)}\rangle\}$, i.e., $\langle\tilde{u}_j^{(\nu)}|u_\beta^{(\mu)}\rangle = \delta_{\nu,\mu}\delta_{\alpha,\beta}$ and $\sum_j |u_j^{(\nu)}\rangle\langle\tilde{u}_j^{(\nu)}| = P^{(\nu)}$. In general, $\langle\tilde{v}_j^{(\nu)}| \neq \langle\tilde{u}_j^{(\nu)}|$ (see I).

In our earlier work, we have repeatedly introduced the concept of "subdynamics." [9,10] The relation of subdynamics to the complex spectral representation can be seen through the projection operators $\Pi^{(\nu)}$ defined by $\Pi^{(\nu)} \equiv \sum_j |F_j^{(\nu)}\rangle\langle\tilde{F}_j^{(\nu)}|$. They satisfy the orthogonality and completeness relations, $L\Pi^{(\nu)} = \Pi^{(\nu)}L$, $\Pi^{(\nu)}\Pi^{(\mu)} = \Pi^{(\nu)}\delta_{\nu,\mu}$, and $\sum_\nu \Pi^{(\nu)} = 1$ (see $(I.139)$). $\Pi^{(\nu)}$ is an extension of $P^{(\nu)}$ to the total Liouvillian L. Because these projection operators commute with the Liouvillian, each component $|\rho^{(\nu)}\rangle \equiv \Pi^{(\nu)}|\rho\rangle$ individually satisfies the Liouville equation. For this reason, the projection operators $\Pi^{(\nu)}$ are associated with *subdynamics*. Then, (1) leads to a *Markovian equation* which is a closed equation for the "privileged component" $P^{(\nu)}\rho^{(\nu)}(t)$ of each subdynamics $\rho^{(\nu)}(t)$,

$$i\frac{\partial}{\partial t}P^{(\nu)}|\rho^{(\nu)}(t)\rangle = \theta^{(\nu)}P^{(\nu)}|\rho^{(\nu)}(t)\rangle \tag{3}$$

where we have introduced the "global" collision operator

$$\theta^{(\nu)} \equiv \sum_j \psi^{(\nu)}\left(Z_j^{(\nu)}\right)|u_j^{(\nu)}\rangle\langle\tilde{u}_j^{(\nu)}| = \sum_j |u_j^{(\nu)}\rangle Z_j^{(\nu)}\langle\tilde{u}_j^{(\nu)}| \tag{4}$$

This result is important, since as mentioned any component $P^{(\nu)}$ of the total $\rho(t)$ obeys a non-Markovian equation with memory effects (see RL). These effects can now be describe by the superposition of Markov processes.

APPLICATION TO MODERATELY DENSE GASES

After this brief summary we now consider the application of our approach to moderately dense systems which consist of N hard spheres ($d = 3$), or hard disks ($d = 2$) with the same mass m. We use the "pseudo-Liouvillian" formalism to take into account the singular nature of the hard sphere interaction (RL *Ch.*X, see also ref. [11]). We also consider situations near equilibrium, then the *linearized* regime of the kinetic equations is applicable (see RL *Ch.*V). Reducing the kinetic equation (3) to one-particle d.f.s, we obtain Markovian equations for the one particle velocity d.f. $f_0^{(0)}(\mathbf{v}_1)$ (which corresponds to the vacuum of correlation $\nu = 0$) as well

as for the inhomogeneous component $f_{\mathbf{q}}^{(1)}(\mathbf{v}_1)$ (for which $\nu = 1$):

$$\frac{\partial}{\partial t} f_{\mathbf{q}}^{(\nu)}(\mathbf{v}_1, t) = L_{\mathbf{q}}^{(\nu)} f_{\mathbf{q}}^{(\nu)}(\mathbf{v}_1, t) \tag{5}$$

where $L_{\mathbf{q}}^{(\nu)}$ are the reduced one particle collision operators corresponding to (4). More explicitly

$$L_{\mathbf{q}}^{(\nu)} = \sum_j \left[-i\mathbf{q}\cdot\mathbf{v}_1 + nK_1 + \delta\mathcal{L}_{\mathbf{q}}^{\nu}\left(\xi_j^{\mathbf{q}}\right)\right] |\chi_j^{\mathbf{q}}\rangle\rangle\langle\langle\tilde{\chi}_j^{\mathbf{q}}| = \sum_j \xi_j^{\mathbf{q}} |\chi_j^{\mathbf{q}}\rangle\rangle\langle\langle\tilde{\chi}_j^{\mathbf{q}}| \tag{6}$$

where $\xi = -iz$ and the states $|\chi_j^{\mathbf{q}}\rangle\rangle$ and $\langle\langle\tilde{\chi}_j^{\mathbf{q}}|$ in (6) are the right- and left-eigenstates of the reduced collision operator $L_{\mathbf{q}}^{(\nu)}$ which shares the same complex eigenvalue $Z_j = i\xi_j$ with the Liouvillian. There are three processes included in (6), the first coming from L_0, the second being the linearlized Boltzmann operator nK_1, and the third coming from mode–mode coupling. To distinguish the reduced one-particle states $|\chi_j^{\mathbf{q}}\rangle\rangle$ from the N-particle states $|u_j^{(\nu)}\rangle$, we have used the double-ket notation for the reduced states. The inner product of the reduced states is defined as $\langle\langle g|f\rangle\rangle \equiv \int d\mathbf{v}_1 \varphi^{eq}(v_1)^{-1} g^*(\mathbf{v}_1) f(\mathbf{v}_1)$, where φ^{eq} is the normalized Maxwellian $\varphi^{eq}(v) = (\beta m/2\pi)^{d/2}\exp(-\beta m v^2/2)$.

The operator K_a is the *linearized Boltzmann collision operator/* defined by (see *RL Chs.* V and X)[‡]

$$K_a\Phi(\mathbf{v}_a) = \int d\mathbf{v}_b \langle 0|t^{(ab)}|0\rangle \left(1 + \hat{P}_{ab}\right) \varphi^{eq}(v_b)\,\Phi(\mathbf{v}_a) \tag{7}$$

where $|0\rangle \equiv |\mathbf{k}^N = 0\rangle$, and $\hat{P}_{ab}$ is an exchange operator of the velocities $\mathbf{v}_a$ and $\mathbf{v}_b$. The "binary collision operator" $t^{(ab)}$ is defined in the pseudo-Liouvillian formalism by

$$t^{(ab)} = -iV\delta L^{(ab)} \equiv a_0^{d-1} \int_{\hat{\mathbf{s}}\cdot\mathbf{v}_{ab}>0} d\hat{\mathbf{s}}\,(\hat{\mathbf{s}}\cdot\mathbf{v}_{ab}) \left[\delta(\mathbf{r}_{ab} - a_0\hat{\mathbf{s}})\,\hat{b}_{\mathbf{v}}^{(ab)} - \delta(\mathbf{r}_{ab} + a_0\hat{\mathbf{s}})\right] \tag{8}$$

Here, a_0 is a diameter of the hard spheres, $\hat{\mathbf{s}}$ is the unit vector, the operator $\hat{b}_{\mathbf{v}}^{(ab)}$ replaces $\mathbf{v}_a$ and $\mathbf{v}_b$ by their pre-collisional values $\mathbf{v}_a'$ and $\mathbf{v}_b'$, and $\mathbf{v}_{ab} \equiv \mathbf{v}_a - \mathbf{v}_b$.

Retaining first only two-mode coupling, the operator $\delta\mathcal{L}_{\mathbf{q}}^{\nu}$ in (6) becomes

$$\delta\mathcal{L}_{\mathbf{q}}^{\nu}(\xi) = \frac{in}{(2\pi)^d} \int d\mathbf{k} \int d\mathbf{v}_b \langle a_{\mathbf{q}}|t^{(ab)}|a_{\mathbf{q}-\mathbf{k}}, b_{\mathbf{k}}\rangle\langle a_{\mathbf{q}-\mathbf{k}}, b_{\mathbf{k}}|g_2(\xi)\,t^{(ab)}|a_{\mathbf{q}}\rangle \left(1 + \hat{P}_{ab}\right)\varphi^{eq}(v_b) \tag{9}$$

where $|a_{\mathbf{k}}\rangle \equiv |\mathbf{k}_a = \mathbf{k}, \mathbf{k}^{N-1} = 0\rangle$, $|a_{\mathbf{k}}, b_{\mathbf{k}'}\rangle \equiv |\mathbf{k}_a = \mathbf{k}, \mathbf{k}_b = \mathbf{k}', \mathbf{k}^{N-2} = 0\rangle$, and the reduced two-particle propagator is defined by $g_2(\xi) \equiv [\xi + iL_0^{(ab)} - nK_a - nK_b + t^{(ab)}]^{-1}$ with $L_0^{(ab)} \equiv L_0^{(a)} + L_0^{(b)}$. The propagator has to be evaluated as an analytic continued function from the right-half plane of ξ. To obtain (9) we have approximated the N-particle propagator G_Q^{ν} in (2) by retaining only the binary correlation subspace, i.e., $G_Q^{\nu} \approx G_2 \equiv P^{(2)}[z - P^{(2)}LP^{(2)}]^{-1}$.

One can expand the propagator g_2 in (9) in a geometrical series of $t^{(ab)}$ using the propagator $g_2'(\xi) \equiv [\xi + iL_0^{(ab)} - nK_a - nK_b]^{-1}$. Retaining only the lowest order term in the geometrical series (i.e., replacing g_2 by g_2' in (9)) we obtain as an approximation of $\delta\mathcal{L}_{\mathbf{q}}^{\nu}(\xi)$, the operator $\delta\mathcal{L}'^{\nu}_{\mathbf{q}}(\xi)$. This new operator $\delta\mathcal{L}'^{\nu}_{\mathbf{q}}(\xi)$ is well known (see *RL Ch.* X). For $\xi = \epsilon \to 0+$ it corresponds to the so-called "ring operator" whose graphical representation is given in Fig. 2.

[‡] In *RL* it denoted by $C_a^{(1)}$.

Figure 2: Graphical representation of the simplest "ring process," where the dots represent the linearized Boltzmann operator.

In Fig. 2 the propagator are renormalized by the linearized Boltzmann operators (c.f. *RL*(X.113)). From the ring diagrams with $\xi = +\epsilon$, one then obtain the kinetic equation for the one particle velocity d.f. $\varphi(\mathbf{v}_1)$: (c.f. *RL* (X.73))

$$\frac{\partial}{\partial t}\varphi(\mathbf{v}_1,t) = \left[nK_1 + \delta\mathcal{L}_0'^0(+\epsilon)\right]\varphi(\mathbf{v}_1,t) \tag{10}$$

In contrast, in our approach the ring operator $\delta\mathcal{L}'^0(+\epsilon)$ is replaced by $\delta\mathcal{L}_0'^0(\xi_j^0)$. Here, ξ_j^0 is of order γ, instead of $\xi = +\epsilon$, where $\gamma^{-1} = t_r = (na^{d-1}v_0)^{-1}$ is the mean free time for a moderate dense system, and $v_0 = (m\beta)^{-1/2}$ is the "thermal velocity." This difference is essential, because our operator is finite while the ring operator $\delta\mathcal{L}_0'^0(+\epsilon)$ diverges for $d = 2$. This can be shown by separating the contribution of (9) into two parts, the "hydrodynamic part" and the "nonhydrodynamic part," $\delta\mathcal{L}_0'^0 = \delta\mathcal{L}_h'^0 + \delta\mathcal{L}_{nh}'^0$. Here, $\delta\mathcal{L}_h'^0$ is defined as the contribution of the integration over $\mathbf{k}$ restricted to the domain $|\mathbf{k}| < k_0$, while $\delta\mathcal{L}_{nh}'^0$ deals with the domain $|\mathbf{k}| > k_0$, where k_0^{-1} is a constant wave number of the order of the mean free path.

For the hydrodynamic part, the real part of the eigenvalues $\lambda_\alpha^{\mathbf{k}}$ of the linearlized Boltzmann collision operator nK_1 are given as $\mathrm{Re}\,\lambda_\alpha^{\mathbf{k}} = -k^2 D_\alpha$, where D_α are "diffusion coefficients" associated to summable invariants $\alpha \in (1, 2, \cdots, d+2)$ (see *RL*). The integration over the hydrodynamic modes then leads to a contribution of the order

$$\delta\mathcal{L}_h'^0(\xi_\alpha^0) \sim n\left(a_0^{d-1}v_0\right)^2 \int_0^{k_0} dk \frac{k^{d-1}}{k^2 D + \gamma} \tag{11}$$

where we have approximated $\xi_\alpha^0 \sim \gamma$, $D_\alpha \sim D \equiv v_0(a_0^{d-1}n)^{-1}$, and $t^{(12)} \sim a_0^{d-1}v_0$. This integral exists for all dimensions. For example, for $d = 2$ we have $\delta\mathcal{L}_h'^0 \sim (\tau_0/t_r)\ln[1 + (k_0 l_m)^2]$, where we have introduced the notations $l_m \equiv \sqrt{D/\gamma} = (a_0^{d-1}n)^{-1} \sim k_0^{-1}$ for the mean free path, and $\tau_0 \equiv t_c/t_r = a_0^d n$ with $t_c = a_0/v_0$. It is easy to show that the geometrical series of the propagator in g_2 converges for moderately dense systems [12]. Furthermore it is well-known that the non-hydrodynamic part of the ring diagram gives a finite correction to the Boltzmann collision operator (see refs. [13,14] and also *RL*). Therefore, Markovian kinetic equations exist for all dimensions. This is in contrast to (10), as (11) is the *diverging* integral for $d = 2$ if γ is replaced by $\epsilon \to 0+$.

For the inhomogeneous component $\mathcal{L}_{\mathbf{q}}^{(1)}$ the eigenvalue equation for the right-eigenstate is given by

$$\left[-i\mathbf{q}\cdot\mathbf{v}_1 + nK_1 + \delta\mathcal{L}_{\mathbf{q}}^1(\xi_j^{\mathbf{q}})\right]|\chi_j^{\mathbf{q}}\rangle\rangle = \xi_j^{\mathbf{q}}|\chi_j^{\mathbf{q}}\rangle\rangle \tag{12}$$

In the hydrodynamic case with $|\mathbf{q}| < k_0$ for $d = 3$, this equation reduces to the same eigenvalue equation as introduced by Ernst and Dorfman to discuss the nonanalytic dispersion relations

in classical fluids. [15] We can repeat this calculation for $d = 2$. We then obtain the following nonanalytic solution for the class of eigenvalues $\xi_\alpha^{\mathbf{q}}$ which vanish in the limit $q \to 0$ for the hydrodynamic modes $\alpha = 1, \cdots, 4$ [12]

$$\xi_\alpha^{\mathbf{q}} = -iqc\sigma_\alpha - q^2 D_\alpha + q^2 \Delta_\alpha \ln\left(\frac{q}{k_0}\right), \qquad \text{for } d = 2 \tag{13}$$

where c is the sound velocity, $\sigma_1 = -\sigma_2 = 1$ for the sound-modes, and $\sigma_3 = \sigma_4 = 0$ for the share-mode and the heat-mode, $D_\alpha > 0$, $\Delta_\alpha > 0$ and $\Delta_\alpha / D_\alpha \sim \tau_0$. In contrast to 3d systems (where the nonanalyticity is given as $\sim q^{5/2}$ [15,16]) the nonanalytic term of order $q^2 \ln q$ in (13) dominates in the real part of the eigenvalues for extremely small values of q satisfying $q/k_0 < \exp(-1/\tau_0)$.

As in the homogeneous case, the original ring diagram $\delta\mathcal{L}_{\mathbf{q}}^1(+\epsilon)$ diverges for the 2d case. Our spectral analysis removes the divergence and leads to a well-defined Markovian kinetic equation (5) in each subspace. Moreover, we have shown that (12) has a well-defined limit for $q \to 0$, that is, the right-hand side of (12) vanishes in this limit [12]. Hence we have a consistent description of the approach to equilibrium.

The qualitative behavior of the eigenvalues of the reduced collision operators for higher-order correlation subspace ($\nu \geq 2$) is the same as in the inhomogenous subspace with $\nu = 1$ in (12). The spectral analysis leads to well-defined Markovian kinetic equations in every subspace. Moreover, we have again a singular contribution proportional to $q^2 \ln q$ for the 2d case in addition to the diffusion process $\sim q^2$ in each subspace [12].

So far, we have considered the time evolution of reduced d.f.s $f_{\mathbf{q}}^{(\nu)}$ that correspond to the privileged components of the distribution functions $P^{(\nu)}\rho^{(\nu)}$ in (3). They obey Markovian kinetic equations. In actual experiments, however, we follow the evolution of the distribution functions obtained by the reduction of $P^{(\nu)}\rho$, such as $\varphi(\mathbf{v}_1)$ in (10), and not of $P^{(\nu)}\rho^{(\nu)}$. In our spectral representation, the long-time tail effects in $\varphi(\mathbf{v}_1)$ is obtained by superposing the Markovian evolution occurring in each subspace. Let us consider the initial condition $|\rho(0)\rangle = P^{(0)}|\rho(0)\rangle$. Then, (1) gives us

$$\begin{aligned}\varphi(\mathbf{v}_1, t) = L^{Nd/2} \sum_\nu \sum_j \int d\mathbf{v}_2 \cdots \int d\mathbf{v}_N \langle 0| \left(P^{(\nu)} + \mathcal{C}^{(\nu)}\left(Z_j^{(\nu)}\right)\right) |u_j^{(\nu)}\rangle \\ \times n_{\nu,j}^2 e^{-iZ_j^{(\nu)}t} \langle \tilde{v}_j^{(\nu)}| \left(P^{(\nu)} + \mathcal{D}^{(\nu)}\left(Z_j^{(\nu)}\right)\right) |0\rangle\langle 0|\rho(0)\rangle\end{aligned} \tag{14}$$

Let us analyze more closely the creation operator $\langle 0|\mathcal{C}^{(\nu)}(z)$ which leads from ν to 0. It corresponds to various dynamical processes (involving vertices such as $P^{(2)}\delta L P^{(3)}, \cdots$ and N-particle propagators $G'_\mu(z) \equiv P^{(\mu)}[z - P^{(\mu)}L'P^{(\mu)}]^{-1}$, where the prime on L' denotes that the interactions $t^{(ab)}$ in the corresponding reduced propagators $g_\mu(\xi)$ defined in (9) is excluded, i.e., $G'_\mu(z)$ corresponds to $g'_\mu(\xi)$ introduced just after (9)). For example, for $\nu = 3$ we have

$$\begin{aligned}\langle 0|\mathcal{C}^{(3)} = \langle 0|G'_0\delta L G'_2 \Big[1 + \delta L G'_2 + \delta L G'_2 \delta L G'_2 + \delta L G'_2 \delta L G'_3 + \delta L G'_3 \delta L G'_3 \\ + \delta L G'_3 \delta L G'_4 + \delta L G'_3 \delta L G'_2 + \delta L G'_0 \delta L G'_2 + \cdots \Big] \delta L P^{(3)}\end{aligned} \tag{15}$$

and a similar expansion for $\mathcal{D}^{(3)}(z)$. As a result, formula (14) corresponds to a variety of dynamical processes, some of which are represented in Fig. 3.

Note that the same diagram may contribute to different subdynamics. For example, diagram (a) gives the contribution $P^{(0)}\exp(-iZ_j^{(0)}t)P^{(0)}$ to the $\Pi^{(0)}$ subspace, as well as $G'_0\delta L \exp(-iZ_j^{(2)}t)\,\delta L G'_0$ to $\Pi^{(2)}$.

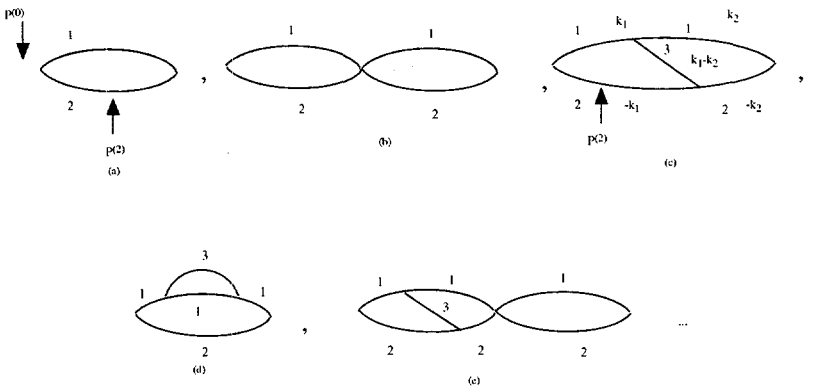


Figure 3: Graphical representation of each term in (14) with the expansion (15). The lines are corresponding to the propagators G'_μ with $\mu \neq 0$ and the vertices to δL. The label of each line corresponds to the index of the velocity, e.g., the line labeled by 1 corresponds to $\mathbf{v}_1$. In diagram (a), $\Pi^{(\nu)}$ indicates the location where the time dependence $\exp(-iZ_j^{(\nu)}t)$ is evaluated. In diagram (c), the wavevectors in the propagators are explicitely written.

Let us sketch the estimation of the long-time tail effect of each contribution in (14) with an example of the $\Pi^{(2)}$ subspace contribution associated to the intermediate wavevector $\mathbf{k}_1$ and $-\mathbf{k}_1$ in diagram (c) (see ref. [12] for detailed calculations). This contribution corresponds to the term which consists of $G'_0\delta L \exp(-iZ_j^{(2)}t)\delta L G'_3 \delta L G'_2 \delta L G'_0$. By integration, the N-particle propagators G'_μ reduce to g'_μ; for example, $G'_3(Z_j^{(2)}) \Rightarrow g'_3(\xi_j^{(2)}) = [\xi_j^{(2)} + i\mathbf{k}_2 \cdot \mathbf{v}_1 - i\mathbf{k}_2 \cdot \mathbf{v}_2 + i(\mathbf{k}_1 - \mathbf{k}_2) \cdot \mathbf{v}_3 - nK_1 - nK_2 - nK_3]^{-1}$, and so on. The long-time tails come from the domain of integration over *small* wavevectors $k_i < k_0$ corresponding to the hydrodynamic modes. In this domain the wave-number dependence of the eigenvalues $Z_j^{(2)}$ are given by the same type of k_i dependence as in (13) for the 2d case (and a similar dependence for 3d case, except that there the singular part is replaced by $k_i^{5/2}$). For a large time scale (with the restriction $t/t_r \ll \exp(1/\tau_0)$ for 2d case), we can show that the singular part of the eigenvalues give negligible correction to the diffusive part $k_j^2 D_\alpha$ [12]. Then, substituting the eigenvalue into $\xi_j^{(2)}$, as well as replacing $i(\mathbf{k}_1 - \mathbf{k}_2) \cdot \mathbf{v}_3 - nK_3$ by its hydrodynamical eigenvalues $-|\mathbf{k}_1 - \mathbf{k}_2|^2 D_\beta$, we obtain $g'_3 \sim (ak_1^2 + b\mathbf{k}_1 \cdot \mathbf{k}_2 + ck_2^2)^{-1}$, where a, b and c are constants of the order of D given in (11), and we have estimated only the share-mode and the heat-mode, since they give the slowest decaying process. The propagator g'_3 has a singularity at $k_1 = k_2 = 0$. We have a similar estimation of the wave-number dependence for g'_μ for $\mu \neq 0$. However, we have a different expression for g'_0, since $g'_0 \sim (k_1^2 D_\alpha + \gamma)^{-1} \sim \gamma^{-1}$ for $k_1 < k_0$. Hence, g'_0 has not a singularity at $k_1 = 0$. The reduction of the vertex δL leads to $t^{(ab)}$, and in the hydrodynamic domain, one can neglect the wave-number dependence (see RL (X.100)), then $t^{(ab)} \sim a_0^{d-1} v_0$ as in (11). Combining these order of magnitude estimations, the time dependence of diagram (c) in Fig. 3 is

$$\int_0^{k_0} dk_1 k_1^{d-1} \int d\hat{\mathbf{k}}_1 \int_0^{k_0} dk_2 k_2^{d-1} \int d\hat{\mathbf{k}}_2 \frac{1}{\gamma} e^{-k_1^2 Dt} \frac{1}{ak_1^2 + b\mathbf{k}_1 \cdot \mathbf{k}_2 + ck_2^2} \frac{1}{dk_1^2 + ek_2^2} \frac{1}{\gamma} \sim t^{2-d} \qquad (16)$$

where $\hat{\mathbf{k}}_i$ are unit vectors, d and e are also constants of the order of D, and we have changed the dummy variables as $x_i = k_i\sqrt{Dt}$, and replaced the upper limit of the integrals by infinity by neglecting exponentially small terms.

One can easily apply this estimation to an arbitrary term in (14). The result is summarized as follows: We first draw the corresponding graph for each term of (14) with the expansion of

type (15). Then we count the number of lines in each intermediate state. Let us denote the maximum number of lines by η. For fixed η, the slowest decaying process is given by the process in which the number μ of lines first increase (starting with $\mu = 0$) and reaching the maximum value $\mu = \eta$, then decrease reaching $\mu = 0$. For example, diagrams (c) to (e) corresponds to the case $\eta = 3$. Among them diagrams (c) and (d) gives the slowest damping, while (e) corresponds to a more rapid decay. There is therefore a close connection between the topology of the diagrams and the long time contributions.

Denoting the contribution of the slowest process by $\varphi^{\eta}(\mathbf{v}_1, t)$ for given η, we have obtained

$$\varphi^{\eta}(\mathbf{v}_1, t) \sim \tau_0^{(d-1)(\eta-1)} (t/t_r)^{(4-d)(\eta-1)/2-2} \tag{17}$$

The classical results for mode-mode coupling are of course recovered as $\varphi^2_{d=3} \sim t^{-3/2}$ and $\varphi^2_{d=2} \sim t^{-1}$, but three mode coupling leads to slower process as $\varphi^3_{d=3} \sim t^{-1}$.

The difference between our results and the results for the velocity autocorrelation functions obtained by Dorfman and Cohen [17], as well as by Résibois and Pomeau [18,19], comes from the fact that we avoid the use of any recipe based on the so-called the "most diverging contributions" in a density expansion (see the comment at p271 in *RL*).*

CONCLUDING REMARKS

Let us discuss consequences of our result (17).

1) For $d < 4$, higher order correlations lead to diverging contribution of (17) for large time, so that there is a radius of convergence t_0. For times $t < t_0$, one can use the expansion (15). The value of t_0 can be found by taking the ratio of (17) between η and $\eta + 1$. Then we have $t_0/t_r = \tau_0^{2(d-1)/(d-4)}$, which gives us $t_0/t_r = \tau_0^{-1}$ for $d = 2$ and $t_0/t_r = \tau_0^{-4}$ for $d = 3$. For $d = 2$, this time scale is much less than $t_r \exp(1/\tau_0)$ for moderate dense systems, which is consistent with the approximation we introduced when we have neglected the singular part $k_i^2 \ln k_i$ of the eigenvalue of the Liouvillian in the hydrodynamic modes. For extremely long time scale $t/t_r > \exp(1/\tau_0)$ when we cannot neglect the logarithmic term in the eigenvalue, t exceeds the radius of convergence t_0. Hence, for this time scale the expansion in terms of the ring processes is not applicable to 2d systems.

2) There is a critical dimension at $d = 4$. Higher order correlations lead to slower decay process for $d < 4$, while they lead to quicker decay process for $d > 4$.

3) For $d \to \infty$, the long-time tail effects vanish, and the evolution of the one-particle velocity d.f. $\varphi(\mathbf{v}_1, t)$ is governed by a Markovian kinetic equation in the $\Pi^{(0)}$ subspace.

Let us conclude with some general remarks. The traditional approach of the kinetic theory based on the BBGKY hierarchy relays upon a *truncation* of the hierarchy at a certain order of correlations. Because the higher order correlations become more important for asymptotic times and for $d < 4$, this truncation is incorrect. There are also difficulties due to truncation in the non-hydrodynamic contributions. With the renormalization of the ring processes binary correlations leads n^m contribution in the density expansion, but so lead also higher order correlations up to $(m+1)$-th order of correlations.

The long-time tail effects described by (17) invalidate the Green–Kubo formalism for $d < 4$. This was well known for $d = 2$, but it is also true for $d = 3$ as there appear contributions of t^{-1}, $t^{-1/2}, \cdots$ coming from multiple mode-mode couplings. Still the linear response formalism remains a valuable tool when used for times where the Markovian approximation to transport

---

* We have obtained a different estimation of their propagator $X_k(t)$ from the one postulated in refs. [18-20], i.e., we have obtained $\|X_k(v_1; t)\| \geq t^{1/2} \exp(-bk^2 t)$ for higher mode processes [12], which in contrast with their assumption of the existence of an upper bound of the propagator.

theory is valid [21]. Also it is rigorous for $d \to \infty$. There appears an amusing analogy with the mean field approach in equilibrium statistical mechanics.

In conclusion our approach based on the spectral decomposition of the Liouville operator which avoids carefully non-dynamical assumptions appears to be of special interest when going beyond the limit of dilute gases.

ACKNOWLEDGMENTS

We want to express our thanks to Professor H. H. Hasegawa for interesting remarks. We acknowledge the U.S. Department of Energy Grant No. DE-FG03-94ER14465.

References

1. Y. POMEAU and P. RÉSIBOIS, *Phys. Report*, 19C, 63 (1975).
2. P. RÉSIBOIS and M. de LEENER, *Classical Kinetic Theory of Fluids*, John Wiley and Sons, New York (1977).
3. J. R. DORFMAN and T. KIRKPATRICK, in *Lecture Notes in Physics: Systems Far From Equilibrium –Proceedings, Sitges*, 263, Springer-Verlag, Berlin, Heidelberg, New York (1980).
4. J. A. McLENNAN, *Introduction to Non-equilibrium Statistical Mechanics* Prentice Hall, New Jersey (1989).
5. M. MARESCHAL, *Adv. Chem. Phys.* 100, 31, John Wiley and Sons, Inc. (1997).
6. I. PRIGOGINE, *Non-Equilibrium Statistical Mechanics* , Wiley Interscience, New York, (1962).
7. T. PETROSKY and I. PRIGOGINE, *Chaos, Solitons & Flactals*, 7, 441 (1996).
8. T. PETROSKY and I. PRIGOGINE, *Adv. Chem. Phys.* 99, 1 John Wiley and Sons, Inc. (1997).
9. I. PRIGOGINE, C. GEORGE, and F. HENIN, *Physica*, 45, 418 (1969).
10. T. PETROSKY and H. HASEGAWA, *Physica*, 160A, 351 (1989).
11. M. H. ERNST, J. R. DORFMAN, W. R. HOEGY and J. VAN LEEUWEN, *Physica*, 45, 127 (1969).
12. T. PETROSKY, unpublised.
13. J. VAN LEEUWEN and A. WEYLAND, Physica 36, 457 (1967).
14. Y. KAN and J. R. DORFMAN: *Phys. Rev. A*, 16, 2447 (1977).
15. M. H. ERNST and J. R. DORFMAN, *Physica*, 61, 157 (1972).
16. M. H. ERNST and J. R. DORFMAN, *J. of Stat. Phys.*, 12, 311 (1975).
17. J. R. DORFMAN and E. G. D. COHEN, *Phys. Rev. A*, 6, 776 (1972); 12, 292 (1975).
18. P. RÉSIBOIS, *Physica*, 70, 431 (1973).
19. P. RÉSIBOIS and Y. POMEAU, *Physica*, 72, 493 (1974).
20. M. THEODROSOPULU and P. RÉSIBOIS, *Physica*, 82A, 47 (1976).
21. P. RÉSIBOIS, *J. of Chem. Phys.*, 41, 2979 (1964).

Double Phase Slips and Bound Defect Pairs in Parametrically Driven Waves

Hermann Riecke and Glen D. Granzow
Department of Engineering Sciences and Applied Mathematics
Northwestern University, Evanston, IL 60208, USA

Spatio-temporal chaos in parametrically driven waves is investigated in one and two dimensions using numerical simulations of Ginzburg-Landau equations. A regime is identified in which in one dimension the dynamics are due to double phase slips. In very small systems they are found to arise through a Hopf bifurcation off a mixed mode. In large systems they can lead to a state of localized spatio-temporal chaos, which can be understood within the framework of phase dynamics. In two dimensions the double phase slips are replaced by bound defect pairs. Our simulations indicate the possibility of an unbinding transition of these pairs, which is associated with a transition from ordered to disordered defect chaos.

I. INTRODUCTION

While low-dimensional chaotic dynamics are quite well understood this is not the case for chaotic dynamics of high dimension as, for instance, spatio-temporal chaos. Spatio-temporal chaos arises in systems in which spatial degrees of freedom play an important role and the structures are not only chaotic in time but also in space. These dynamics arise often in pattern-forming systems when all ordered patterns become unstable. Of particular interest are systems in which the chaos is extensive, i.e. systems in which quantities like the attractor dimension and the number of positive Lyapunov exponents grow linearly with the system size, so that one can think of them as a large number of coupled, chaotic entities.

Spatio-temporal chaos is found in many experimental systems. It has been extensively studied in convection where a number of different types have been observed. In the presence of rotation, a classic result is the occurrence of domain chaos. It is due to the Küppers-Lortz instability which renders steady convection rolls unstable to rolls with a different orientation [1]. Since the new rolls are susceptible to the same instability a persistent switching of patches of rolls of different orientations is observerd [2]. Very recently spatio-temporal chaos has also been found (without rotation) in regimes in which the convection rolls are in fact linearly stable. Sufficiently large perturbations, however, lead to a state of spiral-defect chaos in which spirals and other types of defects dominate [3].

Another type of spatio-temporal chaos arises in electroconvection in nematic liquic crystals in a traveling wave regime [4]. Due to the axial anisotropy of this system the waves travel only in certain directions relative to the axis of anisotropy. In the regime in question they travel obliquely to that direction and because of the reflection symmetries of the system the dynamics is governed by the competition of waves traveling in 4 directions. The chaotic dynamics arise immediately at the onset of convection and is characterized by defects in the various wave components. Associated with each defect is a suppresion of the corresponding wave amplitude leading to domains in which one or two of the wave components dominate. Some understanding of these dynamics and their possible origin has been obtained within a set of coupled Ginzburg-Landau equations [5]. Spatio-temporal chaos of traveling waves in an isotropic system arise in convection in binary mixtures [6].

A very rich class of pattern-forming systems are parametrically driven waves, the standard realization being surface waves on a fluid that are excited by a vertical shaking of the container at twice the frequency of the waves. Depending on the fluid parameters and the driving frequency spatially periodic and quasi-periodic patterns of various kinds have been found as well as transitions to spatio-temporal chaos [7]. Strikingly, very similar phenomena can be found even if the fluid is replaced by a granular medium like sand [8].

In the present communication we present theoretical results for the dynamics of parametrically driven waves in one and two dimensions within the framework of Ginzburg-Landau equations. In the one-dimensional analysis we address the important question how to characterize the behavior of a spatially and

temporally chaotic state on length scales that are much larger than the typical wavelength. The response of regular periodic patterns to long-wave perturbations is well understood; in some sense it is a dissipative analogue to sound waves in crystals. For steady patterns the response is typically diffusive whereas for oscillatory patterns it is propagative. Both can be described by phase equations (or equations for the local wavenumber). Here we describe a chaotic state for which the same type of description is possible, i.e. although the dynamics on short scales is chaotic in space and time, the large-scale behavior of that state is diffusive. This striking behavior is due to the fact that the chaotic state is driven by *double phase slips* as described below. We show that the homogeneously chaotic state can become diffusively unstable on large scales and separates into arrays of chaotic and non-chaotic domains very similar to phase separation in mixtures. This provides a mechanism that can lead to the *localization* of the chaotic dynamics in space.

Experimentally, localized spatio-temporal chaos has been observed in Taylor vortex flow [9], Rayleigh-Bénard convection [10], and parametrically excited surface waves [11]. So far, the localization mechanism in these systems is, however, only poorly understood.

In the second part we present ongoing work on the dynamics that arise in two dimensions in the regime in which double-phase slips occur in one dimension. As discussed below it leads to 'fluctuating bound defect pairs' . We present evidence that indicates a transition from an ordered state of defect chaos to a disordered one. This transition appears to be associated with an 'unbinding' of the defect pairs.

II. GINZBURG-LANDAU EQUATIONS FOR PARAMETRICALLY DRIVEN WAVES

To obtain a tractable model for parametrically driven waves we consider systems that exhibit a supercritical Hopf bifurcation to traveling waves, i.e. when the control parameter is increased beyond a certain threshold value the basic state becomes unstable to small-amplitude traveling waves. This is, for instance, the case in electroconvection of nematic liquid crystals [4][1]. Just below the Hopf bifurcation the traveling-wave modes are only weakly damped. Therefore a small driving is sufficient to excite standing waves of small amplitude. Consequently, a weakly nonlinear description is possible by expanding about the unforced basic state and treating the forcing and the damping as small perturbations. This leads to two coupled equations for the complex amplitudes of the traveling-wave components, i.e. physical quantities like the vertical fluid velocity u in the midplane of a convection system are given by

$$u(\mathbf{r},t) = \epsilon A(X,Y,T)e^{i(\mathbf{q_c}\cdot\mathbf{r}-\frac{\omega_c}{2}t)} + \epsilon B(X,Y,T)e^{i(\mathbf{q_c}\cdot\mathbf{r}+\frac{\omega_c}{2}t)} + c.c. + o(\epsilon). \tag{1}$$

The complex amplitudes A and B vary on the slow time and space scales, $T = \epsilon^2 t$, $X = \epsilon x$, and $Y = \epsilon y$ with $\mathbf{r} = (x, y)$, and $\epsilon \ll 1$.

Using standard symmetry arguments one obtains for the amplitudes A and B in one dimension the Ginzburg-Landau equations [13]

$$\partial_T A + s\partial_X A = d\partial_X^2 A + aA + bB + cA(|A|^2 + |B|^2) + gA|B|^2, \tag{2}$$

$$\partial_T B - s\partial_X B = d^*\partial_X^2 B + a^* B + bA + c^* B(|A|^2 + |B|^2) + g^* B|A|^2. \tag{3}$$

The coefficients in (2,3) are complex except for s and b, which are real. The real part a_r of the coefficient of the linear term a gives the linear damping of the traveling waves in the absence of the periodic forcing and is proportional to the distance from the Hopf bifurcation. The coefficient of the linear coupling term b gives the amplitude of the periodic forcing as can be seen from the fact that it breaks the continuous time-translation

---

[1] Convection in binary mixtures also exhibits a Hopf bifurcation to traveling waves. It is, however, subcritical [6] and the waves appear right away with finite amplitude. An analysis of the parametric forcing of that system is therefore more complicated [12].

symmetry $t \to t + \Delta t$ that implies the transformation $A \to Ae^{i\Delta t\omega_e/2}$ $B \to Be^{-i\Delta t\omega_e/2}$. The imaginary part a_i of the coefficient of the linear term a gives the difference between the frequency of the unforced waves and half the forcing frequency ω_e.

The same Ginzburg-Landau equations are obtained also for systems that do not exhibit a Hopf bifurcation, if they have weakly damped traveling wave modes. This is the case for surface waves on fluids with small viscosity. Then again only a small driving is necessary to excite the surface waves and one can perform the same kind of expansions. In that case $a_r < 0$ represents the damping of the waves and the group velocity parameter s is in general complex, indicating that the dissipation depends already linearly on the wavenumber. However, all dissipative terms are small.

In addition to the trivial solution $A = B = 0$ (2,3) possess three types of simple solutions: $|A| = |B| = const$, $|A| \neq |B|$ (both constant), and $|A| = |B|$ (both time-periodic). We are in particular interested in the first type, which corresponds in the physical system to standing waves which are phase-locked to the parametric forcing, i.e. they are excited by the forcing. With increasing a_r they become unstable to solutions of the second type, which correspond to traveling waves as they exist also in the absence of the periodic forcing. Solutions of the third type correspond to standing waves that are not phase-locked to the forcing. For $g_r < 0$ they are unstable to the traveling waves.

The response of the phase-locked standing waves to long-wave perturbations can be described using a phase equation [14],

$$\partial_T \phi = D(q)\partial_X^2 \phi \qquad \text{with } q = \partial_X \phi, \tag{4}$$

which due to the spatial reflection symmetry of the waves is a (nonlinear) diffusion equation. The diffusion coefficient is not necessarily positive and its sign-change indicates an instability of the waves, the Eckhaus instability. For the one-dimensional case the diffusion coefficient was given in [14]. The resulting stability limits as they are relevant for the first part of the paper are shown in fig.1a. The neutral curve, above which the basic state is unstable to standing-wave perturbations, is given by the dashed line. The solid line gives the Eckhaus instability of the phase-locked standing waves. In the second part of the paper we will consider a case in which traveling waves appear in the absence of forcing, i.e. $a_r > 0$. The corresponding stability regions are shown in fig.1b. The neutral curve is given by the dashed line, neutral curve for the appearance of traveling waves by a dotted line. Their Eckhaus instability is denoted by a solid line. In addition to the Eckhaus instability also a parity-breaking instability arises in which the standing waves become unstable to traveling waves. It is indicated by a dashed-dotted line. Over some range of parameters the parity-breaking instability is preempted by a mode that arises first at finite modulation wavenumber (open squares). It emerges from the parity-breaking instability. The standing waves are stable only inside the region marked by the solid lines and the squares.

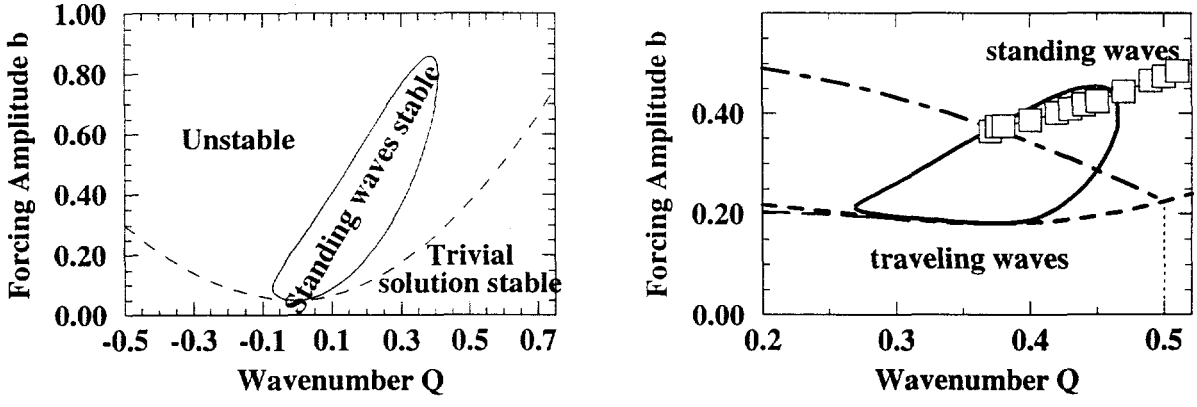


FIG. 1. Linear stability diagram for $c = -1 + 4i$, $d = 1 + 0.5i$, $s = 0.2$, $g = -1 - 12i$. a) $a = -0.05$, b) $a_r = 0.25$.

III. DYNAMICS IN A SMALL SYSTEM

The central new feature of the standing waves to be discussed in this paper is the appearance of 'double phase slips' [15,16]. Usually the Eckhaus instability leads to a single phase slip which changes the total phase in the system and through which the wavenumber of the pattern jumps from outside the stable band to inside the band. Such an event is shown in fig.2a. While this occurs also in this system when the Eckhaus instability is crossed for weak forcing, it is not the case for larger forcing: for $b \approx 0.4$ and above the same perturbation leads to a double phase slip, which consists of two consecutive phase slips that undo each other as shown in fig.2b. After the double phase slip the pattern has the same wave number as before and can undergo the same instability again leading to persistent dynamics. It can be periodic or irregular as discussed in sec.IV.

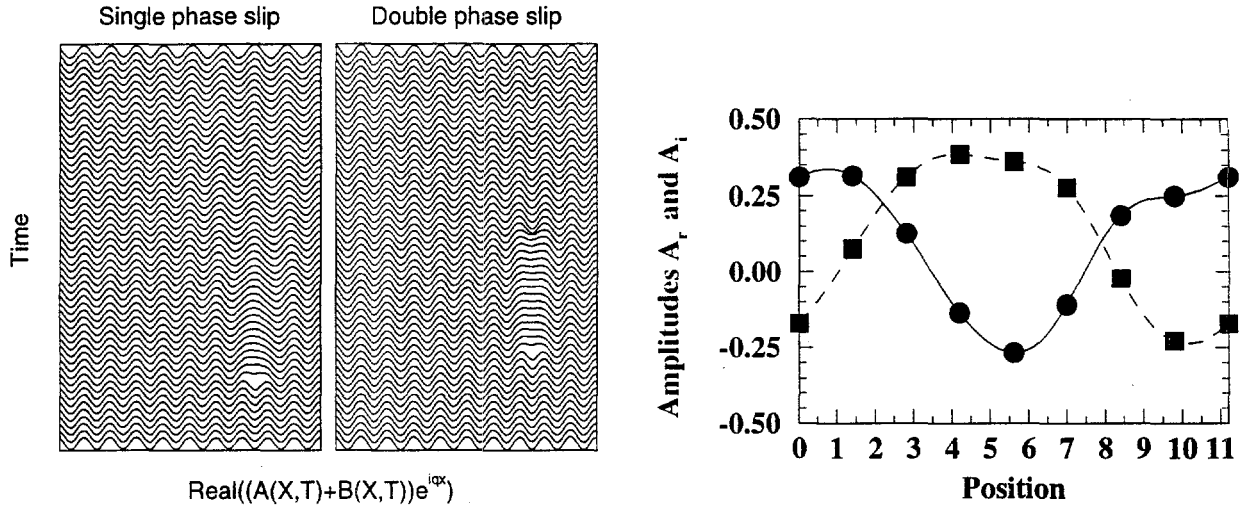


FIG. 2. Space-time diagrams showing a) single phase slip for small values of the forcing amplitude (e.g. $b = 0.1$), and b) double phase slip observed for larger values of the forcing amplitude (e.g. $b = 0.6$). Other parameters as in fig.1a.

FIG. 3. Mixed-mode solution for $L = 11.2$

To study the origin of the double phase slip we consider a minimal system in which only the Fourier modes 0 and 1 are important and simulate (2,3) numerically with a pseudospectral code keeping only the Fourier modes -4 to +4. Changing the length of the system at fixed $b = 1.0$ we find the phase portraits shown in fig.4a and fig.4b. In each run the same pattern very close to the solution with one wavelength in the system is chosen as initial condition (solid diamond). For small length (large wavenumber) this initial condition rapidly evolves to the homogeneous solution (short dashes and open triangle).

When the length is increased the homogeneous solution becomes unstable and a new stable fixed point corresponding to a mixed mode involving Fourier modes 0 as well as 1 arises (open circle). Consequently the solution converges to that mixed mode. Fig.3 shows the real and imaginary parts of the amplitude A of the mixed mode for $L = 11.2$. With increasing L the mixed-mode fixed point moves to the left and the trajectory turns around. For $L = 11.2$ the mixed-mode fixed point, which was a stable node for smaller L, becomes a stable spiral point (open square). When L is increased to $L = 11.3$ the spiral point becomes unstable and generates a stable limit cycle as is shown in fig.4b. Closer inspection of that phase portrait shows that the trajectory follows a three-dimensional path: while it starts *outside* the limit cycle it intersects itself a number of times and eventually approaches the limit cycle from *inside*. This indicates that a model that is to capture the mixed mode, its limit cycle, and the two fixed points corresponding to the homogeneous solution and that with one wavelength has to be at least three-dimensional. For yet larger L, a second mixed-mode

fixed point appears, which is associated with the Eckhaus instability of the solution with one wavelength. It becomes important when the limit cycle grows and eventually becomes homoclinic to this additional fixed point around $L = 12$.

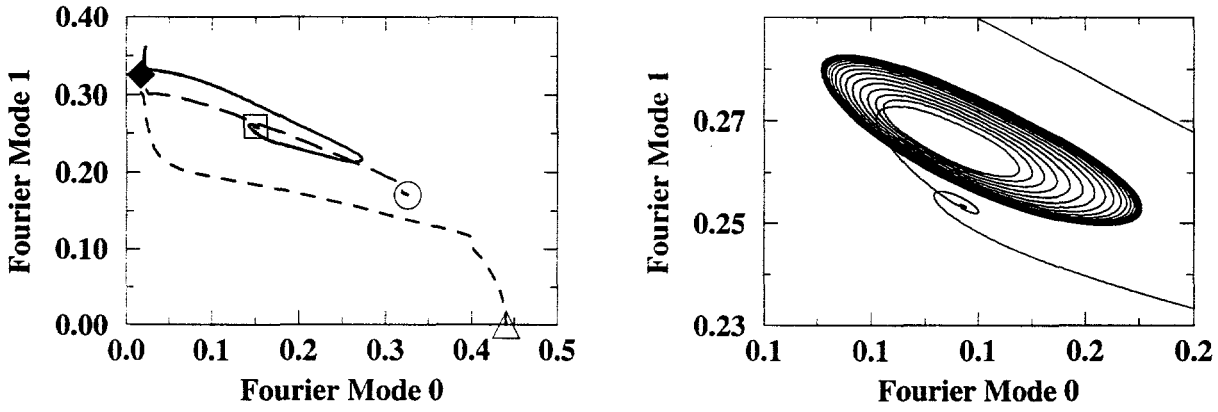


FIG. 4. Phase space projected onto the magnitude of the Fourier modes 0 and 1 for increasing values of the system length L. a) $L = 8$ (short dashes), $L = 10$ (long dashes), and $L = 11.2$ (solid line). Initial condition is marked by a solid diamond. the final fixed point by open symbols. b) $L = 11.3$.

IV. DYNAMICS IN A LARGE SYSTEM

While in the small system discussed in sec.III only simple dynamics was found, complex dynamics arises if the double phase slips can occur at more than one location. Here we describe results for large systems where a surprisingly simple description of the large-scale behavior is possible. Our search for such a description was motivated by states of *localized* spatio-temporal chaos [15,16]. An example is shown in fig.5 where each of the double phase slips is marked as a dot. An initial perturbation triggers a double phase slip. In its vicinity more double phase slips arise and the chaotic activity starts to spread. However, by $t = 50,000$ the width of the chaotic domain stops growing and a stable state is reached, in which the chaotic activity is confined to part of the homogeneous system. At first sight, this result is very surprising; one might have expected that the chaotic activity would always spread through the whole system.

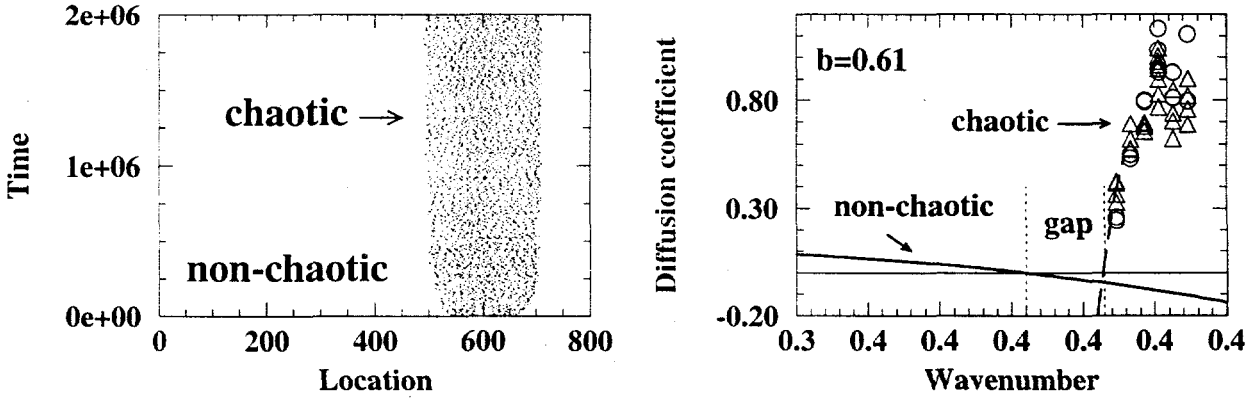


FIG. 5. Localized spatio-temporal chaos for $b = 0.6$ (other parameters as in fig.1a); the averaged effective wavenumber is $< q >= 0.377$).

FIG. 6. Diffusion coefficient of the non-chaotic and the chaotic pattern for $b = 0.6$ (other parameters as in fig.1a).

The mechanism for the localization is due to the fact that double phase slips conserve the total phase, i.e. the wavelength of the pattern is the same before and after the double phase slip. Thus, a time-averaged pattern has a well-defined wavenumber $\hat{q}$ and a well-defined phase $\hat{\phi}$. Since the phase is conserved it is expected to satisfy a slow evolution equation, which on symmetry grounds has the form of a diffusion equation

$$\partial_T \hat{\phi} = \hat{D}(\hat{q}) \partial_X^2 \hat{\phi}. \tag{5}$$

Through a detailed numerical study of the response of the extended chaotic state to a time-periodic perturbation we have explicitly demonstrated the diffusive behavior of the chaotic state on large scales [15,16]. This allowed us also to determine the effective diffusion coefficient $\hat{D}(\hat{q})$ as a function of $\hat{q}$. The result is shown in fig.6. The solid line shows the analytic result for the usual diffusion coefficient $D(q)$ (cf. (4)) which becomes negative at the Eckhaus stability limit. For larger wavenumbers the pattern is unstable to phase slips and undergoes persistent double phase slips. However, a state in which this chaotic activity is homogeneously distributed over the system is not stable to long-wave perturbations as long as the effective diffusion coefficient $\hat{D}$ is negative. Thus, there is a stability gap in wavenumber for which neither the regular nor the chaotic state are diffusively stable (see fig.6). If the initial wavenumber is chosen in that range the pattern has to break up into domains with large wavenumber, which are chaotic, and domains with smaller wavenumber, which are not chaotic. It cannot go into the chaotic (or the non-chaotic) everywhere in space since the total phase, i.e. the integral over the wavenumber, is conserved in the process. Thus, if the wavenumber increases in some part it has to decrease in another part of the system. This separation into domains is very much the same as the phase separation found in equilibrium (fluid) mixtures when they are quenched into the miscibility gap.

V. DYNAMICS IN TWO DIMENSIONS

The wavelength-changing process that occurs in one dimension *via* a phase slip involves in two dimensional patterns the creation (or annihilation) of a pair of defects (dislocations). It is therefore tempting to speculate that a double phase slip will correspond to the creation and annihilation of a 'bound defect pair', i.e. two defects will be created together and will annihilate each other soon thereafter. This would be in contrast to the dynamics observed usually (e.g. in the single complex Ginzburg-Landau equation) where the defects that are created together are not strongly correlated [17,18]. If such a regime of fluctuating bound defect pairs exists one may expect also a transition in which the defects become unbound. In thermodynamic equilibrium such unbinding transitions have found great interest in the context of two-dimensional melting [19] and of vortex unbinding in thin-film superconductors [20]. We are currently investigating the possibility of such a transition in this non-equilibrium system.

We have obtained preliminary results that indicate that such a transition may exist for parameters corresponding to the stability diagram shown in fig.1b. Since no double phase slips seem to occur in the single Ginzburg-Landau equation describing traveling waves we consider a regime in which there exists a transition from parametrically forced standing waves to traveling waves, i.e. a parity-breaking instability. This is the case for $a_r > 0$ as shown in fig.1b. For large b the standing waves are unstable at all wavenumbers as in the case discussed above and double phase slips occur. With decreasing b the parity-breaking instability is approached and we expect that the double phase slips may become replaced by single phase slips in its vicinity.

Fig.7a,b show space-time diagrams for the y-location of defects for simulations for $b = 2$ and $a_r = -0.05$, and $b = 0.5$ and $a_r = 0.25$. For these simulations (2,3) have been extended to two dimensions by replacing the second derivative by a Laplacian. For large b (fig.7a) the defects behave as expected: after their creation they move apart, turn around and annihilate each other again, forming a loop in the space-time diagram. Even in this regime there are loops containing more than one defect pair. Fig.8 shows the statistics of loops

of different sizes (solid symbols): although larger loops occur, their relative frequency decays very rapidly (exponentially) with size. For smaller b the space-time diagram of the defect location is considerably more complicated and offers no simple picture of the dynamics. The corresponding statistics of loop sizes is shown in fig.8 with open symbols. Still most loops contain only one defect pair; this reflects essentially the fact that even if the two defects that are created together moved in a completely uncorrelated fashion they would still be annihilated most likely by their 'partner' since it is closest initially. However, fig.8 shows that large loops are now considerably more frequent. In fact, a power-law decay is more consistent with the data than an exponential decay (cf. fig.8b). It should be mentioned, that the two-dimensional correlation functions of the patterns themselves show also a drastic difference: while for $b = 2$ the pattern is strongly correlated and consists of quite ordered stripes, the correlation function decays rapidly and almost isotropically for $b = 0.5$ [21]. Thus, the two regimes could also be called ordered and disordered defect chaos.

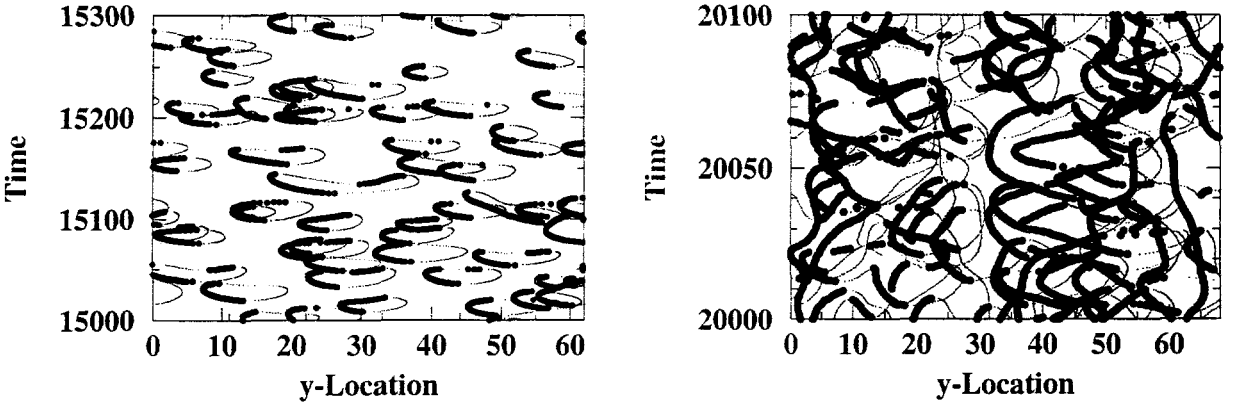


FIG. 7. y-component of the trajectory of defects. a) $b = 2$ (other parameters as in fig.1a); $b = 0.5$ (other parameters as in fig.1b).

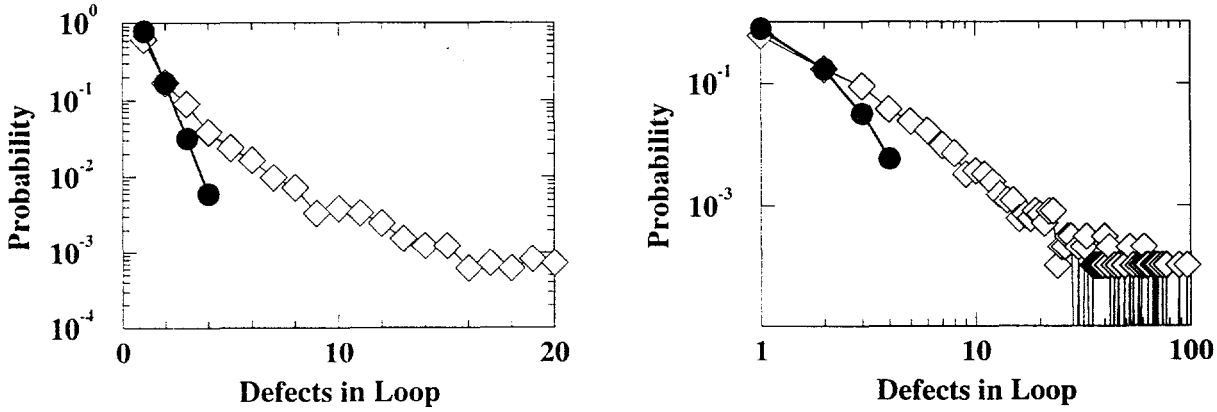


FIG. 8. Relative frequency of defect loops as a function of the number of defect pairs being part of them. open symbols: $b = 2$ (other parameters as in fig.1a); solid symbols: $b = 0.5$ (other parameters as in fig.1b). The double-logarithmic plot indicates that for $b = 0.5$ the distribution is better approximated by a power law than by an exponential.

Current and future work is directed to identify whether the change between the two regimes shown in this paper is smooth or whether it involves a true transition. As diagnostics we are not only using the loop-size distribution, but also the defect life-time, the spatial extent of the loops, the distances travelled by each defect, as well as correlation functions.

This work was supported by the United States Department of Energy through grant DE-FG02-

92ER14303. It also made use of the resources of the Cornell Theory Center, which receives major funding from NSF and New York State with additional support from the Advanced Research Projects Agency, the National Center for Research Resources at the National Institutes of Health, IBM Corporation and members of the Corporate Research Institute.

---

[1] G. Küppers and D. Lortz. Transition from laminar convection to thermal turbulence in a rotating fluid layer. *J. Fluid Mech.*, 35:609, 1969.

[2] Y. Hu, R.E. Ecke, and G. Ahlers. Time and length scales in rotating Rayleigh-Bénard convection. *Phys. Rev. Lett.*, 74:5040, 1995.

[3] S.W. Morris, E. Bodenschatz, D.S. Cannell, and G. Ahlers. The spatio-temporal structure of spiral-defect chaos. *Physica D*, 97:164, 1996.

[4] M. Dennin, G. Ahlers, and D.S. Cannell. Spatiotemporal chaos in electroconvection. *Science*, 272:388, 1996.

[5] H. Riecke and L. Kramer. On the stability of standing waves and traveling rectangles. *in preparation.*

[6] A. La Porta and C.M. Surko. Phase defects as a measure of disorder in traveling-wave convection. *Phys. Rev. Lett.*, 77:2678, 1996.

[7] A. Kudrolli and J.P. Gollub. Patterns and spatiotemporal chaos in parametrically forced surface waves: a systematic survey at large aspect ratio. *Physica D*, 97:133, 1996.

[8] F. Melo, P.B. Umbanhowar, and H.L. Swinney. Hexagons, kinks, and disorder in oscillated granular layers. *Phys. Rev. Lett.*, 75:3838, 1995.

[9] B.W. Baxter and C.D. Andereck. Formation of dynamical domains in a circular Couette system. *Phys. Rev. Lett.*, 57:3046, 1986.

[10] S. Ciliberto and M.A. Rubio. Local oscillations, traveling waves, and chaos in rayleigh-bénard convection. *Phys. Rev. Lett.*, 58:2652, 1987.

[11] A. Kudrolli and J.P. Gollub. Localized spatiotemporal chaos in surface waves. *Phys. Rev. E*, 54:1052, 1996.

[12] H. Riecke, J.D. Crawford, and E. Knobloch. Temporal modulation of a subcritical bifurcation to travelling waves. In P. Coullet and P. Huerre, editors, *The Geometry of Nonequilibrium*, pages 61–64. Plenum Press, 1991.

[13] H. Riecke, J.D. Crawford, and E. Knobloch. Time-modulated oscillatory convection. *Phys. Rev. Lett.*, 61:1942, 1988.

[14] H. Riecke. Stable wave-number kinks in parametrically excited standing waves. *Europhys. Lett.*, 11:213, 1990.

[15] G.D. Granzow and H. Riecke. Phase diffusion in localized spatio-temporal amplitude chaos. *Phys. Rev. Lett.*, 77:2451, 1996.

[16] G.D. Granzow and H. Riecke. Double phase slips and spatio-temporal chaos in a model for parametrically excited standing waves. *SIAM J. Appl. Math.*, submitted.

[17] L. Gil, J. Lega, and J.L. Meunier. Statistical properties of defect-mediated turbulence. *Phys. Rev. A*, 41:1138–1141, 1990.

[18] B.W. Roberts, E. Bodenschatz, and J.P. Sethna. A bound on the decay of defect–defect correlation functions in two-dimensional complex order parameter equations. *Physica D*, 99:252, 1996.

[19] D.R. Nelson. *Defect-mediated phase transitions*, volume 7 of *Phase Transitions and Critical Phenomena*, chapter 1. Academic, New York, 1983.

[20] P. Minnhagen. The two-dimensional Coulomb gas, vortex unbinding, and superfluid-superconducting films. *Rev. Mod. Phys.*, 59:1001, 1987.

[21] G.D. Granzow and H. Riecke. Ordered and disordered defect chaos. *Physica A*, submitted.

ABSOLUTE INSTABILITY FROM LINEAR CONVERSION OF COUNTER-PROPAGATING POSITIVE AND NEGATIVE ENERGY WAVES

A.N. KAUFMAN, A.J. BRIZARD, J.J. MOREHEAD
Lawrence Berkeley National Laboratory
Berkeley, California 94720

E.R. TRACY
College of William and Mary
Williamsburg, VA 23185

ABSTRACT

The resonant interaction of a negative-energy wave with a positive-energy wave gives rise to a linear instability. Whereas a single crossing of rays in a nonuniform medium leads to a *convectively* saturated instability, we show that a double crossing can yield an *absolute* instability.

A negative-energy wave has the property that the system in the *presence* of the wave has *less* energy than in its *absense*. If a positive-energy wave (a) interacts *resonantly* with a negative-energy wave (b), *both* waves may grow in time, while conserving total wave energy:

$$0 = dW/dt = dW_a/dt + dW_b/dt \ .$$

For exponential growth at amplitude-rate γ, we have

$$\begin{cases} dW_a/dt = 2\gamma W_a \ , \\ dW_b/dt = -d|W_b|/dt = -2\gamma|W_b| = 2\gamma W_b \ ; \end{cases}$$

thus $W_b(t) = -W_a(t)$.

For waves propagating freely in a weakly *nonuniform* medium (variation in x, say), we represent the waves ($j = a, b$) in eikonal form, with $k_x^a \neq k_x^b, k_y^a = k_y^b, k_z^a = k_z^b, \omega^a = \omega^b$:

$$\psi_j(\mathbf{x}, t) = \tilde{\psi}_j \mathrm{expi}[\int^x k_x^j(x')dx' + k_y y + k_z z - \omega t],$$

valid in the *non-resonant* regions. *Resonance* occurs at regions where $k_x^a \simeq k_x^b$.

In Figure 1, we illustrate a situation that occurs when both waves [represented by their *rays* $k_x^j(x)$] have *caustics* ($dk_x/dx = \infty$). Then there are *two* ray *crossings*, or resonances. At these crossings (where the wave amplitudes $\tilde{\psi}_j$ change rapidly), both waves *increase* their absolute energy as a result of their interaction.

Consider the lower crossing. Let the ratio of transmitted to incident energy flux of wave a be $T > 1$, and of converted flux of b to incident flux be $C < 0$. Then energy conservation states

$$T + C = 1 \,,$$

or

$$|C| = T - 1 \,.$$

In order that amplification of *both* waves occurs, we need $|C| > 1$, or $T > 2$.

The transmission T can be expressed [1] in terms of the coupling η between the two waves, and the Poisson Bracket magnitude B of the two dispersion functions $D_j(x, k_x)$:

$$T = \exp 2\pi |\eta|^2 / B \,,$$

where

$$\begin{aligned} B &\equiv |\{D_a, D_b\}| \\ &\equiv |(\partial D_a/\partial x)(\partial D_b/\partial k_x) - (\partial D_a/\partial k_x)(\partial D_b/\partial x)| \,. \end{aligned}$$

The condition $T > 2$ thus yields a *threshold* coupling strength η_{th} for the instability:

$$|\eta_{th}|^2 = B(\ln 2)/2\pi \,.$$

To evaluate B, we examine the dispersion functions in more detail. In the neighborhood of a caustic, each ray ($j = a, b$) is a parabolic curve: $x(k_x) = x_j + \beta_j k_x^2$, where $x_j = x(k_x = 0)$ is the caustic location, and $\beta_j \equiv dx/dk_x^2$ is the ray curvature. The two rays thus cross at $k_x = \pm(\Delta x/\Delta\beta)^{1/2} \equiv \pm k_x^c$, where $\Delta x \equiv x_a - x_b$ is the caustic separation, and $\Delta\beta \equiv \beta_b - \beta_a$ is the curvature difference. It follows that the dispersion functions have the form

$$D_j(x, k_x) = (D_x^j)(x - x_j - \beta_j k_x^2) \,,$$

where

$$D_x^j \equiv \partial D_j/\partial x = -(\partial D_j/\partial\omega)(\partial\omega/\partial x)_j = (\partial D_j/\partial\omega)\dot{k}_x^j < 0$$

is proportional to the (constant) ray velocity in the k_x direction. We obtain (after straightforward algebra)

$$|B| = dD_x^a D_x^b(\Delta\beta)k_x^c \equiv 2D_x^a D_x^b(\Delta\beta\Delta x)^{1/2} \,. \tag{1}$$

Thus, for given parameters, the threshold coupling is minimized by *minimizing* the caustic separation Δx.

To determine the latter, we impose the requirement of phase matching: the phase change $\Delta\phi$ of ray a (say), after one circuit, must be an integer multiple N of 2π: $\Delta\phi = 2\pi N$. To find $\Delta\phi$, we use eikonal theory (to lowest order in η): $\Delta\phi = -\oint x(k_x)dk_x - \pi/2$, where the first term is the standard phase integral (the area enclosed by the rays), and the second term is the sum of the two (lowest-order) phase shifts at the conversion points [2]. The evaluation of the phase integral is elementary: $\oint x(k_x)dk_x = (4/3)(\Delta\beta)^{-1/2}(\Delta x)^{3/2}$. The minimum separation $\Delta x_{\min}$ is thus obtained by setting $N = 0$:

$$\Delta x_{\min} = (3\pi/8)^{2/3}(\Delta\beta)^{1/3} \,. \tag{2}$$

Inserting Eq.(2) into Eq.(1), we have $|B|_{\min} = (3\pi)^{1/3} D_x^a D_x^b (\Delta\beta)^{2/3}$, yielding the threshold

$$|\eta_{\rm th}|^2 = [(\ln 2)/2\pi](3\pi)^{1/3} D_x^a D_x^b (\Delta\beta)^{2/3} \ . \tag{3}$$

To obtain the positive growth rate γ (when $T > 2$), we track the energy around one circuit, obtaining $\gamma = [\ln(T-1)]/\tau$, where τ is the time interval for a circuit:

$$\begin{aligned} \tau &= \sum_{j=a,b} \int dk_x / |\dot{k}_x^j| \\ &= (|\dot{k}_x^a|^{-1} + |\dot{k}_x^b|^{-1})(3\pi)^{1/3} (\Delta\beta)^{1/3} \ , \end{aligned}$$

and Eq.(2) has been used.

In conclusion, the general double-crossing mode-conversion problem supports a discrete set of modes, labelled by N above. As the coupling between the modes increases, more of these modes are (absolutely) unstable. For couplings below the threshold value, Eq.(3), all modes are stable. This double-crossing instability can occur in physically different types of wave phenomena.

Acknowledgements

This work was supported by the U. S. Department of Energy under contract No. DE-AC03-76SF00098.

References

1. A.J. Brizard, J.J. Morehead, A.N. Kaufman, and E.R. Tracy, "Double-Cross Instability: An Absolute Instability Caused by Counter-Propagating Positive- and Negative-Energy Waves ," *Phys. Rev. Lett.* **77**, 1500 (1996).
2. J.J. Morehead, "Phase Shifts of Mode Conversion," to be published.

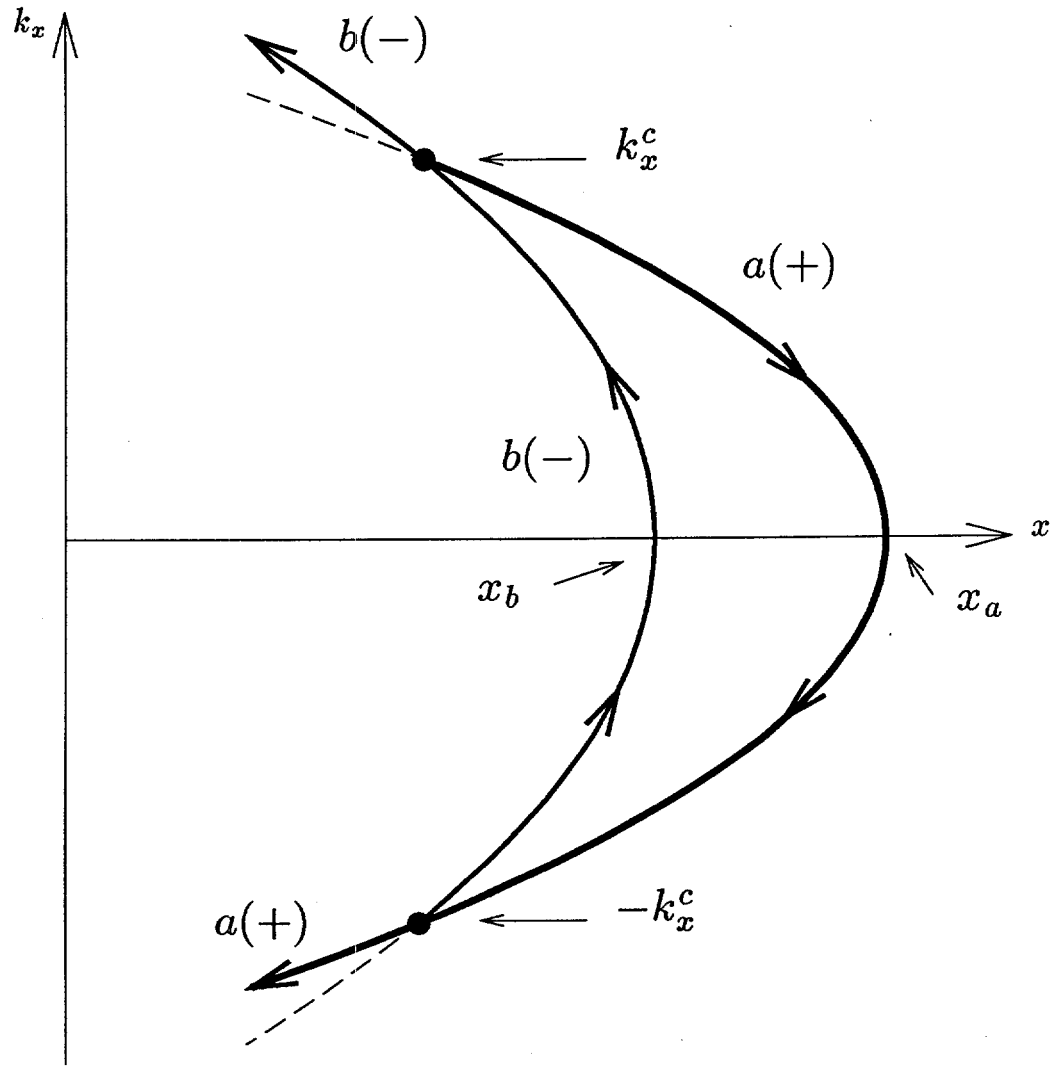


Figure 1. Orbits of the positive-energy wave (a) and negative-energy wave (b)

ENZYME STRUCTURE AND ACTIVITY AT LIQUID-LIQUID INTERFACES

C. J. Beverung, M. J. Tupy, C. J. Radke and H. W. Blanch

Department of Chemical Engineering
University of California, Berkeley, CA 94720

ABSTRACT

Understanding the behavior of proteins interaction at oil/water interfaces is crucial to the design of two-phase bioprocesses (aqueous/organic). An examination of the mechanism of protein adsorption at the oil/water interface was undertaken using tensiometry, transmission electron microscopy (TEM) and a novel total internal reflection fluorescence spectrometer (TIRFS), constructed to monitor adsorption dynamics. Dynamic interfacial tension measurements of protein adsorption show three regimes which can be described by diffusion to the interface, adsorption and denaturation of the adsorbed protein. TEM micrographs show a network of proteins in the adsorbed layer at long times. TIRFS data show that this network formation or protein entanglement in the adsorbed state requires a long period of time to occur. A series of two-monomer random polyamino acids used as model proteins demonstrate many of the adsorption characteristics observed for natural proteins.

INTRODUCTION

Biological approaches for the treatment of hazardous or mixed wastes and for the processing of hydrocarbon mixtures have demonstrated utility [1,2]. Many of these applications require interaction of the reacting phase with a non-aqueous liquid phase. However, many biological catalysts requires an aqueous milieu for activity. In order to exploit

the advantages of these biological mechanisms an efficient means to contact the reacting aqueous medium with the non-aqueous feed is required. Two-phase liquid-emulsion bioreaction systems have been considered for these applications. Although two-phase bioreactors have been developed for various processes, the lack of fundamental understanding presents challenges for design and scaleup.

One aspect of understanding emulsion bioreactor performance is characterizing the oil/water (O/W) interface. The nature of the O/W interface determines emulsion stability, interfacial mass transport and enzyme reactivity. Adsorption of proteins from the aqueous phase at the O/W interface affects all these properties. We have employed interfacial tensiometry, transmission electron microscopy (TEM) and total internal reflection fluorescence spectroscopy (TIRFS) to understand the time scales of adsorption and the affect of the O/W interface on protein conformation.

Understanding protein adsorption is made complex by the nature of the proteins. To facilitate understanding of the adsorption process we have employed poly-amino acid copolymers as model proteins. The dynamic interfacial tension of a series of glutamic acid copolymers with varying hydrophobicities of the second amino acid has been studied.

PROTEIN ADSORPTION

Pendant Drop Tensiometry.

We have applied pendant drop tensiometry to study the dynamics of protein adsorption at a heptane/water interface. This technique involves fitting the Young-Laplace equation, using the interfacial tension as the fitting parameter, to the coordinates of the drop edge as measured by image analysis [3,4]. The pendant drop technique is well-suited to study dynamics because the system is not perturbed by deformation of the interface in order to perform the measurement. This technique is more convenient to apply to O/W systems than DuNouy ring or Wilhelmy plate techniques.

Our recent work has examined the dynamics of adsorption of ovalbumin, β-casein, lysozyme and bovine serum albumin (BSA) at the heptane/buffer interface. Figure 1 shows the concentration dependence of ovalbumin tension at pH 7.1. Low bulk concentrations show a lag period where diffusional effects are significant. As concentration is increased, this lag time is eliminated, and only the long-term effects are seen. At longer times the interface becomes saturated with protein molecules and a steep decrease in the tension is observed. At very long times, on the order of 24 hours, a continuous, steady decrease in interfacial tension is observed. Continuous reduction in tension may be attributed to slow conformational change and aggregation of this globular protein at the O/W interface [5].

Figure 2 shows the dynamic interfacial tension of four proteins studied at a single, low concentration. The same dynamic characteristics discussed above for ovalbumin at low concentrations are observed for lysozyme and BSA. The adsorption of the random coil protein, β-casein, however shows markedly different dynamics. Low concentrations promote short-time diffusional effects. However, at extended periods β-casein approaches a constant

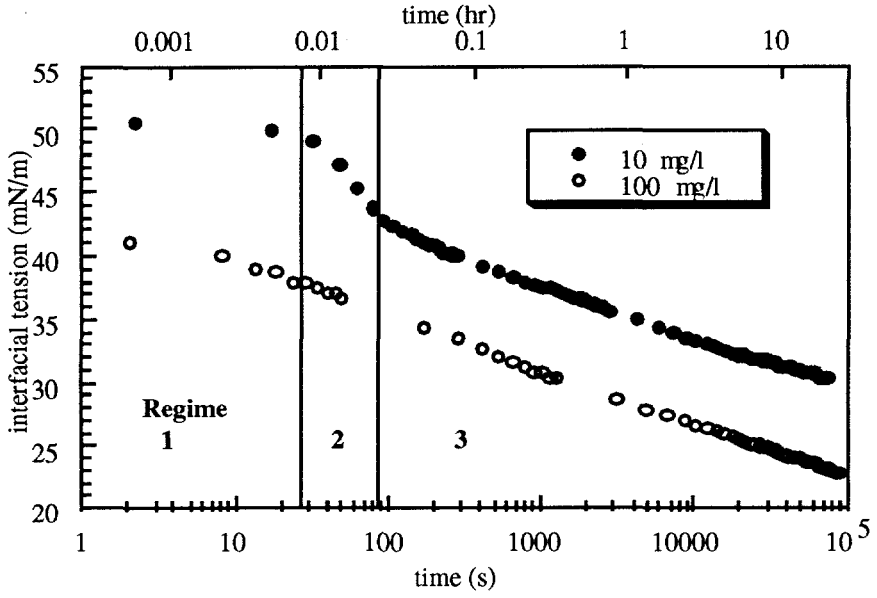


Figure 1. Dynamic Interfacial Tension of Ovalbumin. Tensions of the heptane/water interface were measured at room temperature. The water is a 0.1M sodium phosphate buffer solution at pH 7.1.

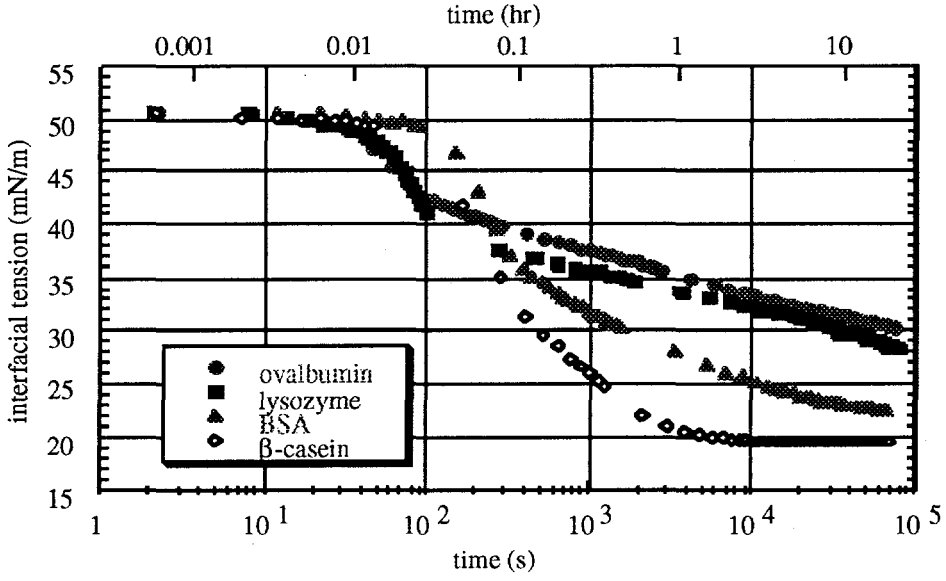


Figure 2. Dynamic Interfacial Tension of Protein Solutions. Tensions of the heptane/water interface were measured at room temperature. The water is a 0.1M sodium phosphate buffer solution at pH 7.1. Protein concentration was 10mg/l for all of the proteins.

interfacial tension of 19 mN/m at a concentration of 10 mg/l. This apparent equilibrium suggests the random-coil nature of this protein enables a more rapid approach to a final adsorption state.

Qualitative information on the adsorbed protein layer has been obtained from a perturbation of the pendant drop after the protein has interacted with the O/W interface for long times (> 18 hours). Withdrawal of the aqueous solution from the drop interior results in

compression of the adsorbed protein layer. Compression of the protein layer leads to collapse of the drop shape and a wrinkling of the gel-like protein film. This phenomenon is macroscopically visible, and suggests a significant interaction between protein molecules present at the interface in order to form an agglomerated network or an entangled polymeric matrix. The strength of this protein film is sufficient to permit it to be harvested in a manner similar to Langmuir-Blodgett techniques.Transmission Electron Microscopy.

Compression of protein layers adsorbed at the O/W interface leads to collapse of the drop shape and a wrinkling of the gel-like protein film. This phenomenon is macroscopically visible, and suggests a significant quantity of interacting protein molecules are present at the interface. TEM of protein layers collected from a planar interface was performed to validate this observation and obtain general structural information of gel layer

For TEM experiments, a glass slide with copper electron microscopy grids is covered with a microporous polymeric formvar film. This film provides structural support for the protein layer, yet allows direct observation through the 0.5 micron holes. The microporous formvar support can be seen in Figure 3 through the protein film and on the right edge of the photograph. The slide was immersed in a protein solution, which was then covered with a layer of heptane. The protein was allowed to adsorb to the O/W interface overnight.

Figure 3 shows a layer of ovalbumin collected in this manner from the heptane/buffer interface. The presence of a protein skin is seen, with significant protein-protein interactions observable in the form of large spherical aggregates. Higher magnification shows an apparent network formation on two length scales. One consists of a homogeneous layer, with strands on the order of the size of individual protein molecules (50 Å). The second is comprised of much larger protein aggregates, with strand thicknesses on the order of hundreds of angstroms and much larger pores. The latter case of aggregations appears to give the protein film its strong mechanical properties and gel-like elastic attributes. Similar results were observed for a film of β-casein adsorbed at the heptane/buffer interface.

The structures seen in the TEM experiments are similar to those seen upon heat-denatured of proteins. The oil phase is an effective medium to induce the order/disorder transition necessary for the denaturation and aggregation of proteins molecules. This effect appears to propagate over length scales beyond single monolayer coverage.

Total Internal Reflection Fluorescence Spectroscopy.

We have employed total internal reflection spectroscopy to investigate the protein conformational changes that were observed from interfacial tension data and TEM images. The interfacial tension at constant temperature is a function of adsorption and the chemical potential of the adsorbed species. Use of extrinsic fluorescent probes and TIRF enable us to follow the adsorption of labeled proteins. By combining the adsorption information and interfacial tension data we gain insight into the energy state of the adsorbed species.

Although more commonly used to investigate the solid/water interface systems [6,7], we have designed and constructed a novel apparatus to measure the fluorescence emitted by adsorbed species at the O/W interface by means of total internal reflection fluorescence. Using light focused at the interface, the evanescent wave produced at the point of total internal

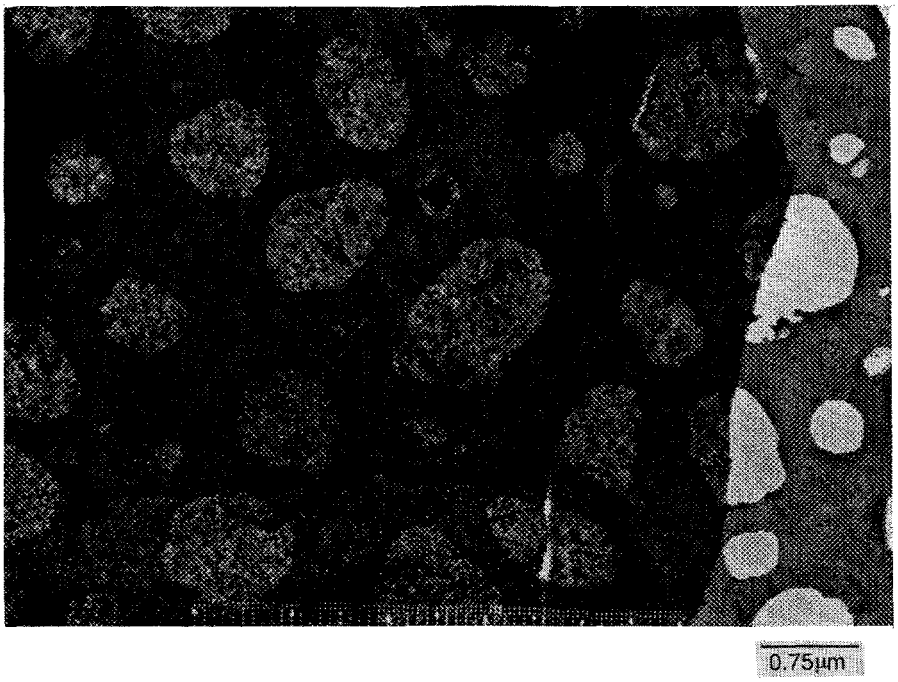


Figure 3. TEM Micrograph of Adsorbed Ovalbumin Film. Film collected from the heptane/buffer interface after 18 hours. The buffer is a pH 7.1, 0.1M sodium phosphate buffer with 100mg/l ovalbumin.

reflection excites fluorophores present in the adsorbed layer. Therefore the fluorescence generated and detected is dominated by those species adsorbed at the interface.

The unique design of our apparatus enables us to follow the dynamic TIRF at O/W interfaces. The centerpiece of the apparatus is the TIRFS cell shown schematically in Figure 4. The cell is designed to ensure diffusive transport of a fluorescent, surface active species from an aqueous bulk solution to the O/W interface. Diffusion of surfactant to the interface occurs in a thin stagnant water layer that is separated by a porous polymeric membrane from a flow channel. The flow channel allows us to introduce the solutions of interest to the TIRFS. By changing the solutions in the flow channel, this design allows the study of the reversibility of adsorption. Within the cell the polymeric membrane is sandwiched by two thin stainless steel sieve plates for support. The oil phase above the interface is sealed in the cell by a Viton O-

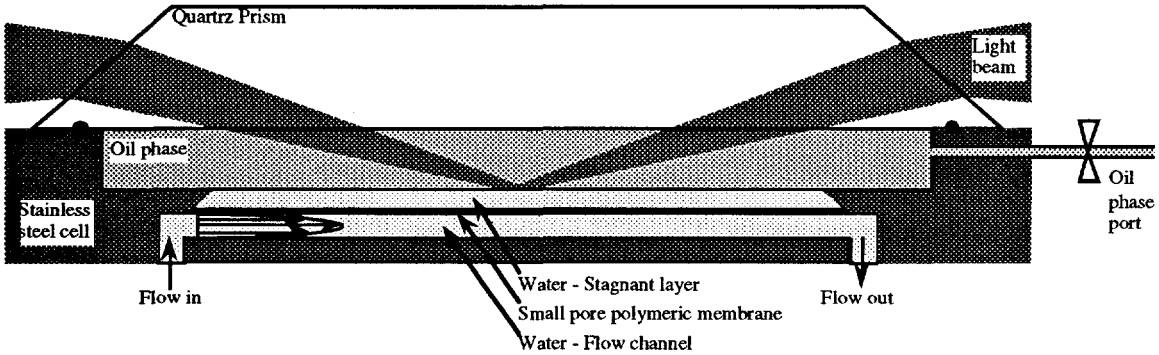


Figure 4. TIRFS Cell Schematic.

ring seal between the cell and a dove quartz prism. Because the oil phase is incompressible and sealed, we are able to eliminate any convective flow through the polymeric membrane thus ensuring pure diffusive transport. To guarantee that the optics stay focused on the interface throughtout the entire experiment the interface is kept flat. This is accomplished by pinning O/W interface at a knife-edge ring and then adjusting the oil phase volume. A flat interface has zero curvature and therefore its shape is not a function of the changing interfacial tension. Once the oil phase is sealed the location of the interface is fixed for the course of the experiment. With this technique, we follow the dynamics of transport of protein to a clean interface and subsequent adsorption from the bulk solution. Use of a monochromator ables the fluorescent spectrum of the adsorbed species to be measured and insight into the chemical environment of the fluorescent probe can be obtained. Studies with a non-adsorbing fluorophore have confirmed the transport mechanism in the TIRFS cell.

Figure 5a shows the loading and washout TIRF for the adsorption of a β-casein/fluorescein conjugate. β-casein dynamic interfacial tension approaches an apparent equilibrium value at long times as discussed earlier. This behavior is unique among the proteins studied and was examined further using the TIRFS to test the adsorption reversibility. Results obtained demonstrate the protein arrival at the interface is consistent with a simple diffusion model. The loading was monitored for a dimensionless time, τ, 1.5 as shown in the figure where τ is defined as:

$$\tau = t \cdot D / h^2 \quad \text{where;}$$

t = time
D = diffusion coefficient
h = thickness of the stagnant layer.

A washout experiment testing reversibility shows that a significant fraction of the adsorbed protein does not desorb over a period of several hours. This result suggests that β-casein is

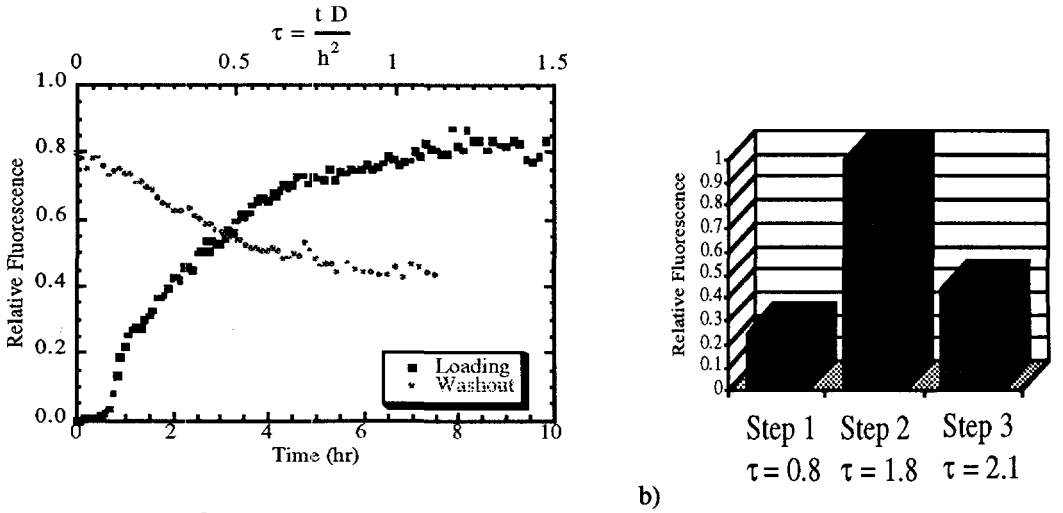


Figure 5. Dynamic TIRFS of β-Casein. Fluorescence excitation at 490nm and emmission measured at 520nm. Experiments were carried out at room temperature. Aqueous solution is pH 7.0, 0.1M sodium phosphate buffer solution with 2mM sodium azide. See text for description of the experiments.

irreversibly adsorbed or at least shows very slow desorption, contrary to the equilibrium picture infers from the interfacial tension data.

A second experiment was designed to investigate β-casein adsorption and possible exchange with the bulk solution. β-casein/fluorescein conjugate labeled with a fluorescein to protein ratio of 1:20 was loaded into the cell for dimensionless time (τ) of 0.8. At this time β-casein/fluorescein conjugate labeled with a fluorescein to protein ratio of 1:4.2 was loaded into the cell for dimensionless time of 1.8. Then washout of the protein with pure buffer solution was carried out for dimensionless time of 2.1. The relative fluorescence at the end of each step is shown in Figure 5b. From this data it is obvious some of the adsorbed non-labeled β-casein was exchanged with labeled β-casein from the second loading step. This observation appears to be contrary to the conclusions reached from the experimental data shown in Figure 5a.

The Protein Adsorption Process.

There is insufficient data at present to make definitive conclusions based on the TIRFS data, but our description of protein adsorption follows. The long time interfacial tension decay shown in Figures 1 and 2 is due to denaturation and entanglement of proteins in the adsorbed state and possible continued incorporation of more proteins from the bulk into the adsorbed layer. The decrease in free energy of the adsorbed layer observed in the tension decrease is due to the entropic gain from greater mobility of the protein polymer chain in the denatured state. β-casein is a disordered protein in bulk solution and apparently gains little additional entropic freedom in the adsorbed state but at long times may become increasingly entangled with other β-casein molecules at the O/W interface. During the TIRFS experiment shown in Figure 5a, the exposure of β-casein to the O/W interface was sufficient to cause significant entanglement. For the experiment shown in Figure 5b, the first loading step with dimensionless time 0.8 was insufficient for significant entanglement between adsorbed protein. The second loading step was of long enough duration to allow for entanglement of β-casein.

ADSORPTION OF POLY-AMINO ACIDS

The goal of this work is to probe the effects of amino acid side chain hydrophobicity, surface charge and bulk conformation on protein adsorption. Synthetic co-polymers of glutamic acid are used as model proteins. However, they are comprised of only two monomers, glutamic acid and a test amino acid. This simplifies the isolation of fundamental driving forces. These systems provide one of the few methods to experimentally attribute adsorption phenomena to particular side chains.

A systematic approach was used here to observe the dynamics of poly-amino acid copolymer adsorption at the heptane/water interface. A series of glutamic acid copolymers with varying hydrophobicities of the second amino acid were studied, with the glutamic acid homopolymer as a control. Dynamic interfacial tension was measured using the pendant drop method. Significant pH effects were observed and correlated with bulk conformational changes observed with circular dichroism (CD).

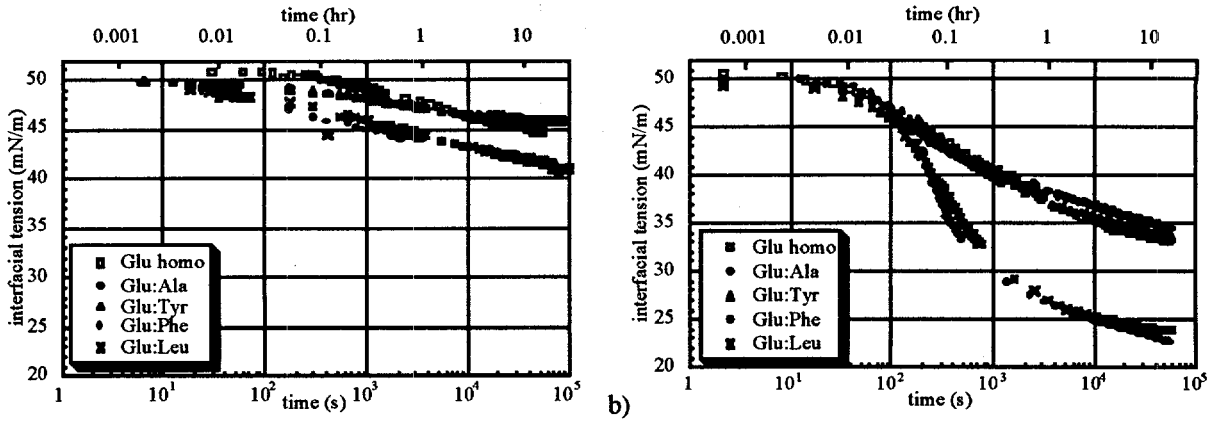


Figure 6. Dynamic Interfacial Tension of Poly-Glutamate Random Co-Polymers. Tension of heptane/water interface at room temperature. Water is a 0.1M sodium phosphate buffer solution with 15mg/l polymer concentration. a) pH 7.1 b) pH 5.3.

The series of copolymers used in this work were glutamic acid with alanine, leucine, phenylalanine and tyrosine as co-monomer. This system was chosen because of the solubility engendered by the charged glutamate side chain as well as the commercial availability of several copolymers of similar mole ratios and degrees of polymerization. Chemical characteristics of these residues encompass a representative range of chain length, aliphaticity and hydrophobicity of the amino acids.

A copolymer concentration of 15 mg/l was used ensure that diffusion effects were observable. Copolymers were examined at pH 7.1 and pH 5.3 to observe charge effects on adsorption. The pKa of glutamic acid is approximately 4.5 for the individual amino acid and 4.7 in proteins. All synthetic poly-amino acids were found to be insoluble at pH between 4.0 and 5.0.

To summarize the dynamic interfacial tension results, Figure 6 shows the combined data for the glutamic acid copolymer systems at pH 7.1 and 5.3. General trends at pH 7.1 indicate minimal adsorption with maximum surface pressures of only 9 mN/m attained for any of the polypeptides. Glutamate residues are primarily charged at this pH, resulting in peptide surface charges greater than -100. The high charge of the polymer is shown to result in a random coil like conformation in the bulk solution by CD spectroscopy. The charged poly-amino acid avoids the non-polar oil phase. A trend attributed to side chain structure is apparent, however. Similar final pressures of 4-5 mN/m are seen for the homopolymer, Glu:Ala and Glu:Tyr, while Glu:Leu and Glu:Phe are able to reach surface pressures of 8-9 mN/m. Even at the high pH, effects of hydrophobicity and structure of the second amino acid are manifested in this small difference in pressure. The polypeptide is able to arrange itself in such a manner that the leucine and phenylalanine groups encounter the heptane phase. The formation of a gel-like adsorbed layer observed for natural proteins was not observed for the poly-amino acids at pH 7.1.

At pH 5.3, the poly-amino acids are closer to the isoelectric point of glutamic acid and hence are less charged. The change in charge lead to α-helix conformation of the poly-amino acids in the bulk solution as measured by CD. The effect on adsorption to the heptane interface is apparent as interfacial tension is reduced much further for all peptides as compared to pH

7.1. The two trends seen at pH 7.1 are enhanced at the lower pH. Glutamate homopolymer, Glu:Ala and Glu:Tyr exhibit similar tension reductions and final surface pressures of ~16-19 mN/m. Glu:Leu and Glu:Phe show significantly different behaviors with tension curves similar to those for globular proteins. Ultimate tension reduction for these systems are ~27 mN/m. The lower electrostatic barrier at pH 5.3 allows the effects of second side chain structure on adsorption to be easily seen. Long chain, hydrophobic residues strongly adsorb (Leu, Phe), provided the chain structure is not terminated with a non-oily moiety (Tyr). Short chains (Ala) and charge or poly (Glu homo) residues also have comparatively low affinities for the oil interface.

CONCLUSIONS

We have applied pendent drop tensiometry, TEM and TIRFS to study protein adsorption. A number of proteins have been examined as well as a series of glutamic acid random co-polymers as model proteins. The dynamic interfacial tension shows three regimes. At early times the proteins show a lag phase. At later times, the proteins diffuse to the interface and adsorb. Once the proteinsadsorb the begin to denature and undergo a very long time process of agglomeration or entanglement. Dynamic TIRFS demonstrates that strong association between adsorbed proteins occurs over a long time periods. At long times, adsorbed proteins form a gel-like skin at the interface that show structuring in TEM micrographs. The poly-amino acid co-polymers demonstate much of the dynamic interfacial tension properties observed for natural proteins.

ACKNOWLEDGEMENT

This work was performed under the auspices of the U. S. Department of Energy

REFERENCES

1. J. TRAMPER, H. C. VAN DER PLAS AN DP. LINKO, eds. *Biocatalysts in Organic Synthesis* (Proceedings of an International Symposium on this Subject), Elsevier, Amsterdam, 1985.

2. C. LAANE, J. TRAMPER AND M. D. LILLY, eds. *Biocatalysts in Organic Media* (Proceedings of an International Symposium on this Subject), Elsevier, Amsterdam, 1987.

3. J. M. ANDREAS, E. A. HAUSER AND W. B. TUCKER, "Boundary Tension by Pendant Drop," *J. Phys. Chem.* 42, 1001 (1938).

4. P. CHENG, D. L. L. BORUVKA, Y. ROTENBERG AND A. W. NEUMANN, "Automation of Axisymmetric Drop Shape Analysis for Measurements of Interfacial Tensions and Contact Angles," *Colloids Surf.* 43, 151 (1990).

5. D. E. GRAHAM AND M. C. PHILLIPS, "Proteins at Liquid Interfaces. I. Kinetics fo Adsorption and Surface Denaturation," *J Colloid Interface Sci.* 70, 415 (1979).

6. T. P. BURGHARDT AND D. AXELROD, "Total Internal Reflection/Fluorescence Photobleaching Recovery Study of Serum Albumin Adsorption Dynamics," *Biophys. J.* 33, 455 (1981).

7. R. D. TILTON, C. R. ROBERTSON AND A. P. GAST, "Lateral Diffusion of Bovine Serum Albumin Adsorbed at the Solid-Liquid Interface," *J. Colloid Interface Sci.* 137, 192 (1990).

APPLICATION OF WELDING SCIENCE TO WELDING ENGINEERING: A LUMPED PARAMETER GAS METAL ARC WELDING DYNAMIC PROCESS MODEL

P. E. Murray, H. B. Smartt, J. A. Johnson

Idaho National Engineering and Environmental Laboratory
Lockheed Martin Idaho Technologies
Idaho Falls, Idaho 83415-2210

ABSTRACT

We develop a model of the depth of penetration of the weld pool in gas metal arc welding (GMAW) which demonstrates interaction between the arc, filler wire and weld pool. This model is motivated by the observations of Essers and Walter [1] which suggest a relationship between droplet momentum and penetration depth. A model of gas metal arc welding [2] was augmented to include an improved model of mass transfer and a simple model of accelerating droplets in a plasma jet to obtain the mass and momentum of impinging droplets. The force of the droplets and depth of penetration is correlated by a dimensionless linear relation used to predict weld pool depth for a range of values of arc power and contact tip to workpiece distance. Model accuracy is examined by comparing theoretical predictions and experimental measurements of the pool depth obtained from bead on plate welds of carbon steel in an argon rich shielding gas. Moreover, theoretical predictions of pool depth are compared to the results obtained from the heat conduction model due to Christensen et al. [3] which suggest that in some cases the momentum of impinging droplets is a better indicator of the depth of the weld pool and the presence of a deep, narrow penetration.

INTRODUCTION

Gas metal arc welding uses a consumable metal electrode which melts in the presence of an electric arc. Droplets are expelled from the electrode and transferred to the workpiece thus producing the weld. Energy transfer from the arc to the workpiece is augmented by the energy of impinging droplets. Transfer of droplets affects many aspects of the welding process including the size, shape and depth of penetration of the weld pool. Penetration is caused by droplets with sufficient momentum to carry energy deep into the pool thus enhancing convective mixing [1]. In constrast, penetration in nonconsumable electrode processes such as gas tungsten arc welding is primarily affected by flow in the weld pool caused by surface tension and buoyancy forces [4].

Theoretical models have been used to predict the size and shape of the cross sectional area of fused metal. Notable examples include the model by Christensen et al. [3] based on Rosenthal's [5] model of a moving point source of energy, and the extension to a moving distributed energy source by Eagar and Tsai [6]. These models assume steady state energy transfer to the workpiece by conduction; the effects of convective heat transfer in the weld pool and the deposition of filler metal are not included. To overcome these limitations, numerical models of heat transfer have been recently developed. These include heat transfer models due to Tekriwal and Mazumder [7] and Pardo and Weckman [8], and fluid flow models due to Tsao and Wu [9] and Kim and Na [10]. The affect of filler metal deposition in these recent models was included

by fixing the mass, velocity and rate of impinging droplets rather than direct simulation of droplet dynamics.

Experiments reported by Essers and Walter [1] established that the momentum of impinging droplets effects the depth of the weld pool. Therefore accurate prediction of the depth of penetration requires accurate prediction of the mass, velocity and rate of impinging droplets. Droplets are expelled from the filler wire by electromagnetic, gravitational, inertial and drag forces acting on a droplet forming at the tip of the wire [11]. The size and rate of droplets expelled from the filler wire determines the mode of metal transfer which affects many aspects of the process [12]. The velocity of droplets striking the pool depends on the arc length and the acceleration caused by the plasma jet. Therefore successful prediction of the depth of penetration of the weld pool requires a model that includes the effects of heat transfer in the arc, mass transfer from the filler wire, plasma jet flow in the arc, heat transfer and fluid flow in the pool, and dynamics of impinging droplets. A model of gas metal arc welding that includes all these effects has not yet been developed.

EXPERIMENTS

Bead on plate welds were made on 3/16 and 3/8 inch thick carbon steel using an automated gas metal arc welding apparatus. A constant voltage power supply was used with a slope approximately equal to a 2.5 V decrease per 100 A increase in current. The shielding gas was a mixture of 98% Argon and 2% O_2 flowing at 50 scfh. The travel speed was fixed at 15 mm/s. The filler wire was carbon steel type AWS ER70S-6 and its diameter was fixed at 1.14 mm. The CTWD, which denotes the contact tip to workpiece distance, was set to 13, 19 and 25 mm. Eighteen separate welds were made by varying the wire feed speed from 120 mm/s to 300 mm/s for each value of CTWD, which produced a variation in arc power from 6 to 10 MW. The samples were prepared by wet chemical etching and examined using a microscope. The depth of the weld pool was measured from the original surface of the workpiece to the bottom of the pool. Therefore the reinforcement height was not included in the measured depth.

PROCESS MODEL

The model of droplet growth and detachment used in this work evolved from a prior model developed by Shaw [13] that simulates water dripping from a faucet. This non-linear, second-order, spring-mass-damper system may also be used, with suitable modifications, to simulate droplet dynamics in GMAW:

$$m\ddot{x} + b\dot{x} + kx = F_{tot},$$

where x is the droplet displacement, $\dot{x}$ is the droplet velocity, and $\ddot{x}$ is the droplet acceleration. The time-varying mass of the droplet is m, b is the damping coefficient, k is the spring constant, and F_{tot} is the sum of the external forces acting on the droplet. The spring constant and damping coefficient represent the surface tension and internal viscous forces of the liquid neck attaching the droplet to the solid electrode.

In addition to the force balance, the model includes information about the electrical properties of the welding circuit. The dynamic response of the electrical circuit is found from Ohm's Law augmented by an arc voltage model as shown by Shepard [14]:

$$\frac{dI}{dt} = \frac{V_{sp} - R_e I - R_s I - V_{arc}}{L_s},$$

where I is the current, V_{sp} is the power supply voltage setpoint, R_e is the total electrode and droplet resistance, R_s and L_s are the source resistance and inductance, and V_{arc} is the voltage drop across the arc.

The resistance of the solid electrode and droplet is given by:

$$R_e = \rho_e(l_e + 0.5(x + r_d)),$$

where ρ is the resistivity of the electrode material and l_e is the stick-out of the electrode. The length of the drop attached to the electrode includes the radius of the droplet, r_d, and the elongation represented by the spring length, x. The voltage drop across the arc is given by Shepard [15]:

$$V_{arc} = V_o + R_a I + E_a(CTWD - l_e),$$

where V_o is the arc voltage constant, R_a is the arc resistance, E_a is the arc length factor, and $CTWD$ is the contact tube-to-workpiece distance.

The electrode melting rate was shown by Lesnewich [16] to be a function of resistive and anodic heating:

$$M_e = C_1(I^2 \rho_e l_e) + C_2(I),$$

where C_1 and C_2 are constants.

The model incorporates two criteria for drop detachment: balance of forces and pinch instability. The first criterion is based on an imbalance of forces acting on the drop. The sum of the external forces on the drop includes four forces that act to produce the dynamics of droplet motion [17]. These are (1) gravity, (2) the Lorentz force due to the interaction of the electric field and its self-induced magnetic field, (3) the aerodynamic drag caused by the arc plasma flowing past the drop, and (4) a force due to electrode melting which adds momentum to the drop. A force due to the surface tension of the liquid metal acts to prevent the drop from detaching. The second detachment criterion is based on pinch instability theory [18]: an electromagnetic pinch force causes a fluid instability and detachment in case the radius of the drop exceeds a critical radius. These criteria are discussed in detail in Reutzel et al. [2].

When either of the detachment criteria are satisfied, a drop detaches and the volume of the detached droplet, Vol_{detach}, is a function of drop velocity and drop volume, Vol_d:

$$Vol_{detach} = Vol_d \frac{1}{2}\left(\frac{1}{1 + e^{-C_s \dot{x}}} + 1\right),$$

where C_s is an empirically determined constant. This relation ensures that an increasing droplet velocity leads to an increasing volume of detachment [13].

Calculation of weld pool depth is based on the assumption that the depth of the weld pool depends on the force of impinging droplets. Therefore we consider the case in which the droplets have sufficient momentum to penetrate the weld pool and cause a deep, narrow penetration at the base of the pool. Otherwise the depth of the pool depends on heat transfer in the workpiece, and in this case models by Christensen et al. [3] and Eagar and Tsai [6] may be used to obtain an approximation to the pool depth.

We introduce a characteristic length scale, $\mu/(\rho V)$, where μ is the viscosity of the pool, ρ is the density of the pool, and V is the velocity of impinging droplets. Using Stokes' Law for the force on a sphere immersed in a fluid, we introduce a characteristic force, μRV, where R is the radius of impinging droplets. Assuming a linear relation between dimensionless force and dimensionless depth, we obtain:

$$\frac{D}{\left(\frac{\mu}{\rho V}\right)} = A\frac{F}{\mu RV} + B,$$

where D is the depth of penetration, F is the force of impinging droplets, and A and B are constants. The force of impinging droplets is given by:

$$F = \dot{M}V,$$

where $\dot{M}$ is the melting rate of the filler wire. We found that the constants $A = 46$ and $B = 9500$ enable us to correlate force and penetration depth for a range of values of arc power and CTWD.

The velocity of impinging droplets is given by:

$$V = \sqrt{V_0^2 + 2a(L - x)},$$

where V_0 is the velocity of the droplet at the time of detachment from the filler wire, L is the arc length, x is the position of the centroid of the droplet at the time of detachment, and a is the acceleration of the droplet in flight, given by:

$$a = g + \frac{3}{8}C_d\frac{\rho_{gas}V_{gas}^{\ 2}}{\rho_{drop}R},$$

where g is the gravitational acceleration, C_d is the coefficient of drag, ρ_{gas} is the density of the plasma jet, V_{gas} is the velocity of the plasma jet, and ρ_{drop} is the density of the expelled droplet. We assume that the velocity of the plasma jet is equal to 100 m/s and the coefficient of drag is equal to 0.44 [12].

To compare our model to previous models, we choose the heat conduction model by Christensen et al. [3]:

$$\frac{D}{\left(\frac{2\alpha}{U}\right)} = \frac{\varsigma}{1+\varsigma}\sqrt{1+2\varsigma},$$

where α is the thermal diffusivity of the workpiece, U is the travel speed, and ς is the radius of the weld pool obtained from:

$$\frac{\eta VIU}{4\pi\alpha^2\Delta H} = \varsigma e^{\frac{\varsigma}{1+\varsigma}},$$

where η is the arc efficiency which is set to 0.80, V is the arc voltage, I is the current, and ΔH is the change in enthalpy associated with raising the workpiece temperature to the melting point.

RESULTS AND SUMMARY

The measured depth of the weld pool as a function of arc power for various values of CTWD is shown in Figures 1-3. These results show that increasing CTWD at constant power leads to an

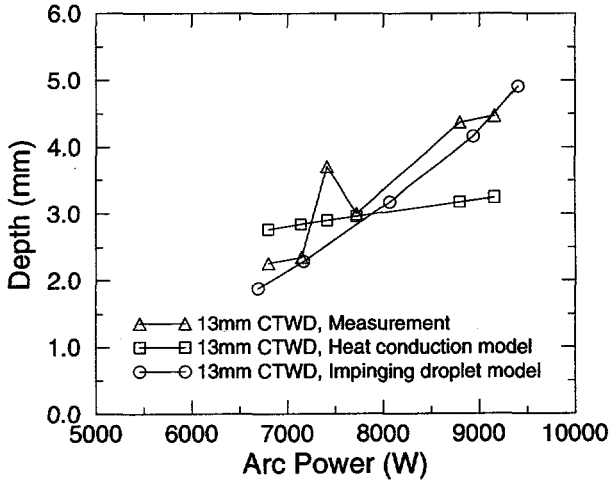


Fig. 1. Measured and simulated depth of the weld pool as a function of arc power for a contact tip to workpiece distance equal to 13mm.

increasing depth of penetration. A larger CTWD is accompanied by an increase in electrical resistance which reduces the current. Therefore the wire feed speed must increase in order to maintain constant power while increasing the CTWD. This leads to an increase in the melting rate of the wire which leads to an increase in the momentum of impinging droplets. Another factor is that a change in CTWD at constant power causes a change in arc length. Since the droplets are accelerated by the plasma jet, a larger arc length increases the momentum of impinging droplets. Therefore the depth of penetration may also increase due to an increase in the arc length.

Comparisons of the measured and simulated depth of the weld pool as a function of arc power for three values of CTWD are shown in Figures 1-3. These results show that in some cases the simulated depth of penetration of impinging droplets is a suitable indicator of the pool depth. In constrast, the simulated depth based on the heat conduction model due to Christensen et al. [3] is not a suitable indicator of the pool depth. Further validation is needed to ensure that the model may be used to correlate droplet momentum and penetration depth for a wider range of welding process variables, including variable travel speed, filler wire size, material type and shielding gas. The results obtained in this study suggest that in some cases the momentum of impinging droplets has a greater effect on the depth of the weld pool than heat transfer from the arc.

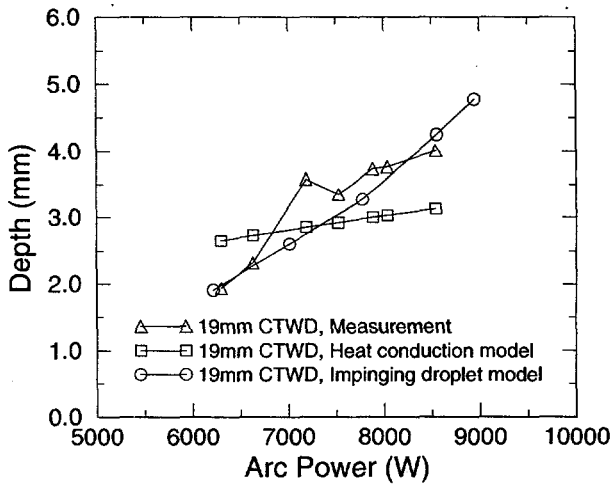


Fig. 2. Measured and simulated depth of the weld pool as a function of arc power for a contact tip to workpiece distance equal to 19mm.

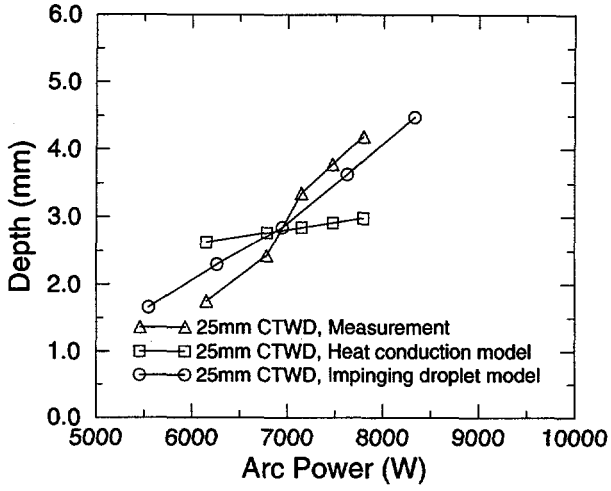


Fig. 3. Measured and simulated depth of the weld pool as a function of arc power for a contact tip to workpiece distance equal to 25mm.

ACKNOWLEDGMENTS

Work supported by the U. S. Department of Energy, Office of Basic Energy Sciences under DOE Idaho operations Contract No. DE-AC07-94ID13223.

REFERENCES

1. W. G. ESSERS. and R. WALTER, "Heat Transfer and Penetration Mechanisms With GMA and Plasma-GMA Welding," *Welding J.*, Vol. 60, pp. 37s-42s, 1981.

2. E. W. REUTZEL, C. J. EINERSON, J. A. JOHNSON, H. B. SMARTT, T. HARMER, and K. L. MOORE, "Derivation and Calibration of a Gas Metal Arc Welding (GMAW) Dynamic Droplet Model,", 4th International Conference on Trends in Welding Research, Gatlinburg, Tennessee, June 5-8, 1995.

3. N. CHRISTENSEN, V. de L.DAVIES, and K.GJERMUNDSEN, "Distribution of Temperatures in Arc Welding," *British Welding J.*, Vol. 2, pp. 54-75, 1965.

4. K. C. MILLS, and B. J. KEENE, "Factors Affecting Variable Weld Penetration," *Int. Mater. Rev.*, Vol. 35, pp. 185-216, 1990.

5. D. ROSENTHAL, "The Theory of Moving Sources of Heat and its Application to Metal Treatments," *Trans. ASME*, Vol. 43, pp. 849-866, 1946.

6. T. W. EAGAR and N.-S. TSAI, "Temperature Fields Produced by Traveling Distributed Heat Sources," *Welding J.*, Vol. 62, pp. 346s-355s, 1983.

7. P. TEKRIWAL and J. MAZUMDER, "Finite Element Analysis of 3 Dimensional Heat Transfer in GMA Welding," *Welding J.*, Vol. 67, pp. 150s-156s, 1988.

8. E. PARDO and D. C. WECKMAN, "Prediction of Weld Pool and Reinforcement Dimensions of GMA Welds Using a Finite Element Model," *Metall. Trans. B*, Vol. 20, pp. 937-947, 1989.

9. K. C. TSAO and C. S. WU, "Fluid Flow and Heat Transfer in GMA Weld Pools," *Welding J.*, Vol. 67, pp. 70s-75s, 1988.

10. J.-W. KIM and S.-J. NA, "A Study on the Effect of Contact Tube to Workpiece Distance on Weld Pool Shape in Gas Metal Arc Welding," *Welding J.*, Vol. , pp. 141s-152s, 1995.

11. J. H. WASZINK and L. H. J. GRAAT, "Experimental Investigation of the Forces Acting on a Drop of Weld Metal," *Welding J.*, Vol. 62, pp. 109s-116s, 1983.

12. Y.-S. KIM, and T. W. EAGAR, "Analysis of Metal Transfer in Gas Metal Arc Welding," *Welding J.*, Vol. 72, pp. 269s-278s, 1993.

13. R. SHAW, "The Dripping Faucet as a Model Chaotic System," The Science Frontier Express, Aerial Press, Inc., 1984.

14. M. E. SHEPARD and G. E. COOK, "A Non-linear Time Domain Simulation of Self Regulation in Gas Metal Arc Welding," 3rd International Conference on Trends in Welding Research, Gatlinburg, Tennessee, June 1-5, 1992.

15. M. E. SHEPARD, Modeling of Self-Regulation in Gas Metal Arc Welding, Ph.D. Dissertation, Vanderbilt University, 1991.

16. A. LESNEWICH, "Control of Melting Rate and Metal Transfer in Gas-Shielded Metal Arc Welding, Part I: Control of Electrode Melting Rate," *Welding J.*, Vol. 37, pp. 343s-353s, 1958.

17. A. D. WATKINS, H. B. SMARTT, and J. A. JOHNSON, "A Dynamic Model of Droplet Growth and Detachment in GMAW," 3rd International Conference on Trends in Welding Research, Gatlinburg, Tennessee, June 1-5, 1992.

18. J. F. Lancaster, The Physics of Welding, Pergamon Press, Oxford, 1984.

STATUS OF RESEARCH AIMED AT PREDICTING STRUCTURAL INTEGRITY

W. G. Reuter

Lockheed Martin Idaho Technologies Company
P.O. Box 1625, Idaho Falls, Idaho, USA 83415-2218

ABSTRACT

Considerable research has been performed throughout the world on measuring the fracture toughness of metals. The existing capability fills the need encountered when selecting materials, thermal-mechanical treatments, welding procedures, etc., but cannot predict the fracture process of structural components containing cracks. The Idaho National Engineering and Environmental Laboratory and the Massachusetts Institute of Technology have been collaborating for a number of years on developing capabilities for using fracture toughness results to predict structural integrity. Because of the high cost of fabricating and testing structural components, these studies have been limited to predicting the fracture process in specimens containing surface cracks. This paper summarizes the present status of the experimental studies of using fracture toughness data to predict crack growth initiation in specimens (structural components) containing surface cracks. These results are limited to homogeneous base materials.

INTRODUCTION

The concern addressed in this paper is to identify the ability and limitations of using a single fracture toughness parameter (K, J, or δ), which is assumed to uniquely quantify the displacement, strain, and stress fields at the crack tip, to predict structural integrity. In predicting the fracture process (crack growth initiation, stable crack growth, and catastrophic failure) for structural components, it is necessary to have some method of measuring the fracture toughness of the component and the ability to relate these measurements to the behavior of the structural component. ASTM test standards exist for measuring plane strain fracture toughness (K_{Ic}) [1] and the critical value of J (J_{Ic}) near the onset of stable crack extension.[2] The critical value for crack tip opening displacement (δ) measured per E1290[3] may be substituted for J. K_{Ic} is limited to linear-elastic behavior whereas J_{Ic} is used for linear-elastic and elastic-plastic conditions. A single fracture parameter is used in many prediction procedures; the most commonly used of these procedures is the failure analysis diagram (FAD). Figure 1 is an example of a FAD. The region of interest, in this paper, is the vertical axis on the left side identified as the Toughness Ratio, which is the applied stress intensity factor (K) divided by the material's fracture toughness (K_{Ic}).

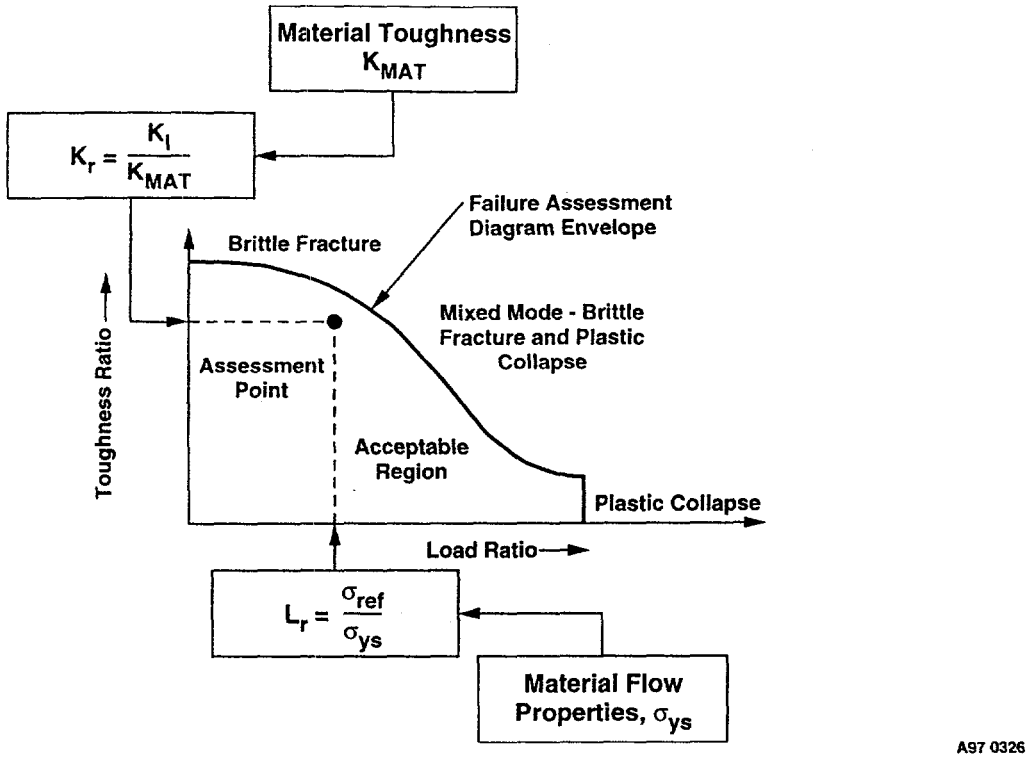


Figure 1. Failure assessment diagram.

Three potentially significant differences exist among a) test specimens used in ASTM E399, E813, and E1290, b) structural components, and c) specimens containing surface cracks (see Table 1). First, the difference in size can often lead to an elastic-plastic or fully plastic condition in the test specimen while the structure may exhibit a linear-elastic behavior. Second, there may be a significant difference between the size of the crack in the test specimen and in the structural component. This difference can lead to a nonconservative estimate of the structural lifetime because of statistical effects. Finally, the longer crack is more likely to intercept an embrittled region (lower fracture toughness) than the shorter crack. Because surface-cracked specimens are more similar to structural components than ASTM test specimens, it is expected that tests of surface-cracked specimens more closely simulate the response of a structure. Therefore, the use of specimens containing surface cracks to simulate the behavior of a structure is a logical step.

The main difference between a through crack and a surface (part-through) crack is that the through crack is frequently treated as a two-dimensional crack problem with the crack driving force reasonably constant along the straight crack front. Crack growth initiation occurs whenever the crack driving force at any location along the crack front equals the fracture toughness. Since the crack driving force is generally constant along the straight crack front, it follows that initiation of crack growth occurs at the same "instant." (It is sometimes observed that the crack growth occurs in a tunnel fashion where the crack growth is retarded at the specimen free surfaces. This situation is of no concern to subject matter being addressed.) The surface crack is a three-dimensional crack (see Figure 2), and the crack driving force varies around the perimeter of the crack front. Currently, the standard approach is to assume that when the crack driving force at any location around the crack perimeter equals the fracture toughness of the material, then crack growth initiation occurs. Another difference is that through-crack test specimens are normally removed from a plate, whereas surface-crack specimens (and structures) can be plates, cylindrical sections, etc.

Table 1. Comparison between specimens and structures.

Configuration	Specimen Size	Crack Configuration	Constraint[a]
Test specimens	Small, flat plates	Straight, through-thickness, crack lengths are generally short	High
Structural components	Large, flat plates or cylindrical sections	Curved, part-through thickness, crack lengths may be large. Often experiences multiaxial loading.	Variable
Surface cracked specimens	Flat plates or cylinders	Curved, crack lengths range from short to medium. Substantial variation in a/t and a/2c. Generally exposed to uniaxial loading.	Variable

[a]Constraint is defined as the ratio of the hydrostatic stress to the equivalent von Mises stress.

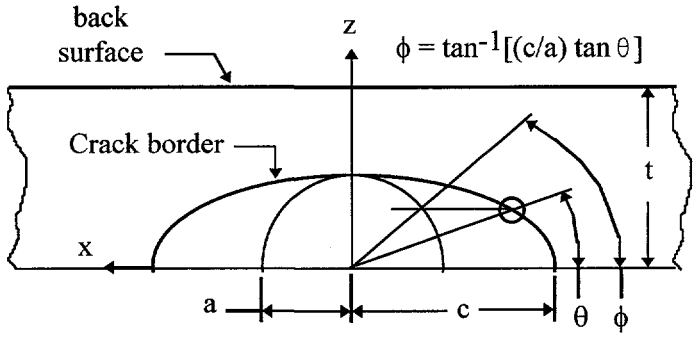


Figure 2. Surface crack geometry.

For test specimens that are configured to obtain a maximum constraint, the measured fracture toughness is generally the lowest value that may be obtained. For structural components that exhibit a lower constraint (and generally a higher fracture toughness) due to shallow cracks or an adjacent free surface, use of the lowest value fracture toughness may be too conservative, i.e., requiring unnecessary work stoppage or repairs. This paper summarizes experimental research that provides answers to many questions of transferability of test data to structural behavior.

APPROACH

Fracture toughness measurements provided in ASTM E 399, E 813, and E 1290 are based on conditions associated with initiation of crack growth. For linear-elastic fracture mechanics (LEFM) conditions, crack growth initiation is often, but not always, synonymous with catastrophic failure. Therefore, an approach based on crack growth initiation appears to be useful for structures that exhibit linear-elastic behavior. An alternative approach is to use a procedure based on crack arrest. ASTM E 1221[4] provides test methods for measuring plane-strain crack-arrest fracture toughness. However, it is very difficult to transfer this concept to a structure because it is necessary to quantify the

compliance of the component as it plays a key role in the crack arrest process. Therefore, crack-arrest fracture toughness will be ignored in this paper.

For elastic-plastic [nonlinear elastic fracture mechanics (NLEFM)] and fully-plastic conditions, it is generally observed that substantial stable crack growth may occur after initiation of crack growth. Therefore, to more accurately predict the fracture process for these conditions, it is necessary to have test results that provide information on stable crack growth. The J-Δa results obtained using ASTM E 1152[5] provide information on stable crack growth. But, from an analytical viewpoint, the wake of a growing crack includes cold-worked, plastically deformed material, where Hutchinson-Rice-Rosengren (HRR) solutions are no longer valid, i.e., no longer unique.[a] The limit of crack growth for J validity has been studied by Xia et al.,[6] who concluded that no approach can be based on a single parameter resistance curve. To evaluate if this is a practical problem, however, requires experimental verification. Dadkhah and Kobayashi[7] and May and Kobayashi[8] performed experiments in which they observed that J no longer provided the HRR fields at the crack tip when Δa exceeded some amount of crack growth. This strongly suggests that J no longer represents the crack tip stress fields when crack growth, Δa, is more than two to three times crack-tip opening displacement (CTOD). Based on these statements, the following discussion is limited to using fracture toughness results obtained per existing ASTM Test Standards to predict conditions for initiation of crack growth in specimens containing surface cracks. These specimens have been fabricated from homogeneous materials.

Because of the complexity of the issues being considered, this paper will examine linear-elastic fracture mechanics and nonlinear elastic fracture mechanics as two separate topics. However, the specimens used to measure fracture toughness were removed from the same piece of material as specimens containing surface cracks.

LEFM CONDITIONS

The ability to predict the maximum flaw size that may be allowed in a structural component is based on knowing the applied stress, the fracture toughness (measured per E399), and having an applicable equation. For these tests, the maximum applied stress intensity factor (K_{max}) was calculated using the failure load, the actual crack size, and the Newman-Raju[9] equations. The ability to predict crack growth initiation was quantified by calculating the ratio of the calculated maximum applied stress intensity (K_{max}) to K_{Ic}, the same ratio used in the FAD diagram.

Results

The following results were obtained from three materials, Ti-15-3, a monolithic SiC, and D6-aC (a high-strength steel). The Ti-15-3 was heat-treated to a yield strength (σ_{ys}) of 1,452 MPa, with resulting plane strain fracture toughness (K_{Ic}) = 41.4 MPa$\sqrt{m}$. The specimens containing surface cracks had a fatigue precrack starter notch fabricated by electric discharge machining (EDM), and the specimens were then load cycled either in tension or in bending to grow the desired fatigue precrack. The fatigue precracks had a crack depth-to-thickness ratio (a/t) ranging from 0.05 to 0.94 and crack depth-to-length ratios (a/2c) ranging from 0.01 to 0.47. The specimens were tested by monotonic loading in either tension or bending. The test results consisted of load versus acoustic emission, load versus crack mouth opening displacement, and load versus displacement (quantified using moiré interferometry) data. The specimens containing surface cracks failed catastrophically with little or no

[a]Personal communication with F. McClintock, September 11, 1996.

stable crack growth. Reuter et al.[10] discussed these results and noted that K_{max}/K_{Ic} ranged from 1.02 to 1.64 for specimens loaded in tension, and from 1.02 to 2.07 for specimens loaded in bending.

Standardized test procedures are not available for measuring the plane strain fracture toughness of ceramics. The procedures provided in ASTM E 399 were used, except that the SiC specimens contained EDM notches as opposed to fatigue precracks. The defects in the specimens containing surface cracks were also made with EDM and had the same notch root radius. It was assumed that comparisons between the EDM-notched specimens would be as valid as comparisons between specimens containing fatigue precracks. The specimens were tested by monotonic loading in bending. The test results were the same as those collected for Ti-15-3. The specimens failed catastrophically with no stable crack growth. Reuter et al.[10] presented these results and noted that K_{max}/K_{Ic} ranged from 0.99 to 1.41 for one series of SiC specimens and from 0.94 (1.00)[b] to 1.93 (1.39) for a second series of SiC specimens.

The tests described above were performed at a single facility and it was desired to broaden the scope of material to include a high-strength steel and multiple test facilities. Therefore, an International Cooperative Test Program was organized to test specimens fabricated from D6-aC, a high-strength steel.[11] The material was heat-treated to σ_{ys} = 1,587 MPa, with resultant K_{Ic} = 54 MPa$\sqrt{m}$. The surface crack configuration had an a/2c ratio ranging from 0.08 to 0.60 (0.56) for tensile loading and from 0.08 to 0.51 for bending loads; the a/t ratio ranged from 0.23 (0.28) to 0.89 (0.83) for tensile loads and from 0.22 to 1.0 (0.84) for bending loads. The fracture toughness results per ASTM E399 were the same as those collected for Ti-15-3 in that they exhibited a nominal elastic behavior. For the surface-cracked specimens, the behavior ranged from general elastic to substantial crack growth prior to attainment of the maximum load. Use of the initial precrack size and shape and the maximum load at failure to calculate K_{max} for comparison with K_{Ic} was inappropriate. Therefore, the test plan was modified to detect the onset of crack growth initiation using d.c. potential drop and acoustic emission. A change of 5% in d.c. potential drop was defined as crack growth initiation. The K_{max}/K_{Ic} ratio ranged from 0.91 (0.98) to 1.61 for tensile loading and from 1.14 to 1.81 for bending loads, see Figure 3.

Discussion

In a vast majority of these 99 tests, $K_{max}/K_{Ic} \geq 1.0$.[c] This illustrates that the use of the Newman-Raju9 equation and the measured K_{Ic} result in conservative estimates of failure for specimens (structures) containing surface cracks. But a number of instances (8 out of 19 metal specimens tested in bending) were observed in which considerable conservatism occurred ($K_{max}/K_{Ic} > 1.50$) when K_{max} occurred at the free surface (ϕ = 0 degrees). These results are acceptable for many applications, but it might be necessary to better understand the conditions controlling fracture. The primary questions of interest are a) What parameter other than the calculated crack driving force (K_{app}) is responsible for initiation of crack growth? and b) Is the local K or modification (average, specific locations, etc.) responsible for initiation of crack growth?

In Table 2, test results of the SiC specimens[10] showed that 10 specimens had K_{max} occurring at the free surface and that the ratio K_{max}/K_{Ic} ranged from 0.99 to 1.41 (1.28), suggesting that the substantial conservatism was not observed in this material. For the Ti-15-3 specimens tested,[10] 10 specimens had K_{max} occurring at the free surface and the ratio K_{max}/K_{Ic} ranged from 1.18 to 2.07

---

[b]Items in parentheses denote the value of the next closest neighbor.

[c]This ratio was greater than 1.0 for 97 of 99 specimens tested.

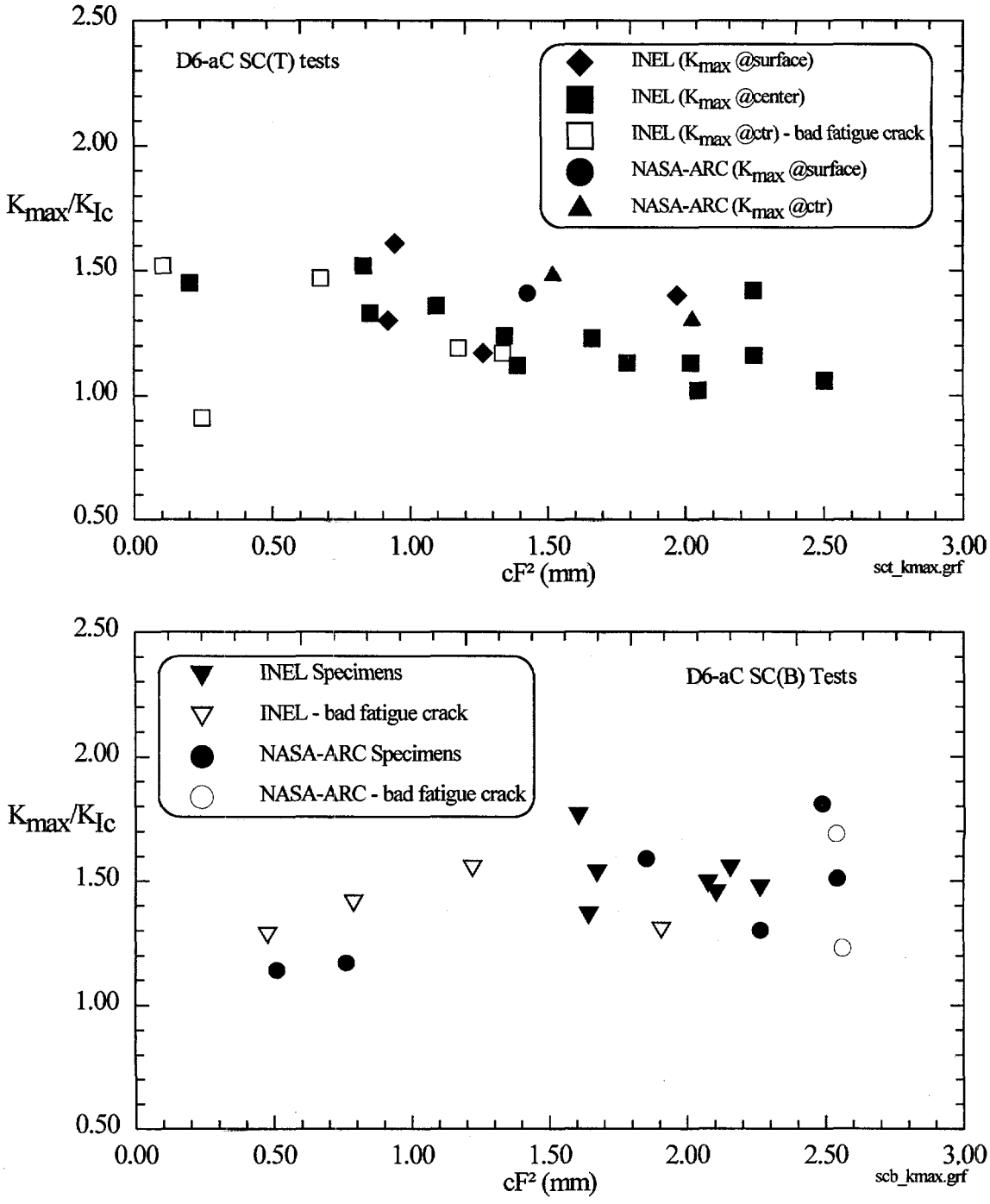


Figure 3. K_{max}/K_{Ic} versus cf^2 for load corresponding to crack growth initiation (d.c. potential drop).

(1.83), suggesting that substantial conservatism was observed in this material. Reuter et al.[10] examined the use of K_{max}, $K_{average}$ (root-mean square of K along the crack front), K_{ave} (average of local K values calculated at all locations along the crack front), and K_{ϕ} (at a specific location). It was concluded that K_{max} was the most conservative single parameter fracture criterion if attainment of K_{Ic} is considered a sufficient condition for fracture. A major concern is how long a crack length segment is required where $K=K_{Ic}$ before crack growth initiation occurs. This has not been answered yet. The use

of $K_{average}$ and $K_{\phi=30°}$ was based on results of Sommers and Aurich[12] for elastic-plastic conditions where it was observed that the maximum crack driving force (J) occurred at $\phi = 30$ degrees. This suggested that the maximum CTOD (δ) also occurred at $\phi = 30$ degrees. Reuter and Lloyd[13] showed that δ was not a maximum at $\phi = 30$ degrees for specimens tested in tension that exhibited elastic-plastic behavior. They observed that the relative magnitude of δ followed the calculated relation for K at applied stresses where crack growth initiation was detected.

As noted earlier, many D6-aC steel specimens containing surface cracks exhibited substantial crack growth. A combination of electric potential change (DCP) and acoustic emission (AE) monitoring was used to detect initiation of crack growth and the associated applied load. The load was reduced after initiation was detected, and cyclic loading was applied to decorate the location and extent of crack growth. Several specimens loaded in tension were examined and it was observed that crack growth initiation occurred at $\phi = 90$ degrees, with the majority of crack growth within ±15 degrees of $\phi = 90$ degrees, and with no crack growth at $\phi = 30$ degrees (see Figure 4). Several of these specimens, loaded in tension, were loaded multiple times, which consisted of (1) fatigue precracking, (2) monotonic loading to obtain stable crack growth, (3) decreased load, (4) cyclic loading to outline the region of the stable crack growth, and (5) unloading. Steps 2, 3, 4, and 5 were repeated until the crack penetrated the back surface (see Figure 4).

For three specimens with thickness (t) = 6.4 mm, K_{init} increased for each cycle of monotonic loading, i.e., with increasing a/t (see Figure 5). This suggests a) a loading history effect, i.e., plastic zone development, b) a constraint effect due to the proximity of the crack tip to the back surface, or c) the Newman-Raju[9] stress intensity distribution limit $a/t < 0.80$ for accurate results has been exceeded. At this time, none of the three possible explanations has been ruled out.

Table 2. Summary of the ratio K_{max}/K_{Ic} for the materials tested.

Material	Test Conditions	Range of K_{max}/K_{Ic}
Ti-6Aal-4V	Tension	1.02 to 1.64
	Bending	1.02 to 2.07
SiC (Two different materials)	Bending	0.99 to 1.41
	Bending	0.94 to 1.93
D6-aC	Tension	0.91 to 1.61, see Fig. 3
	Bending	1.14 to 1.81, see Fig. 3

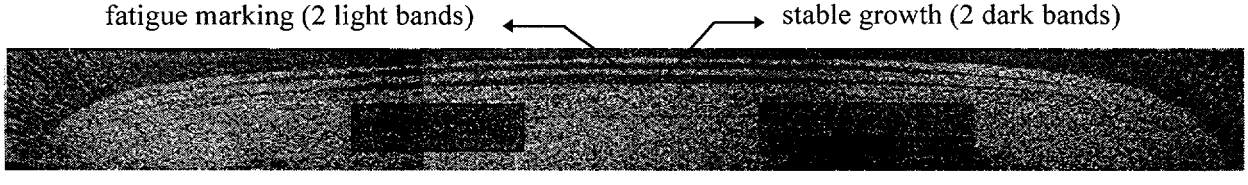


Figure 4. D6-aC Specimen #26 - Fracture surface showing two loading sequences.

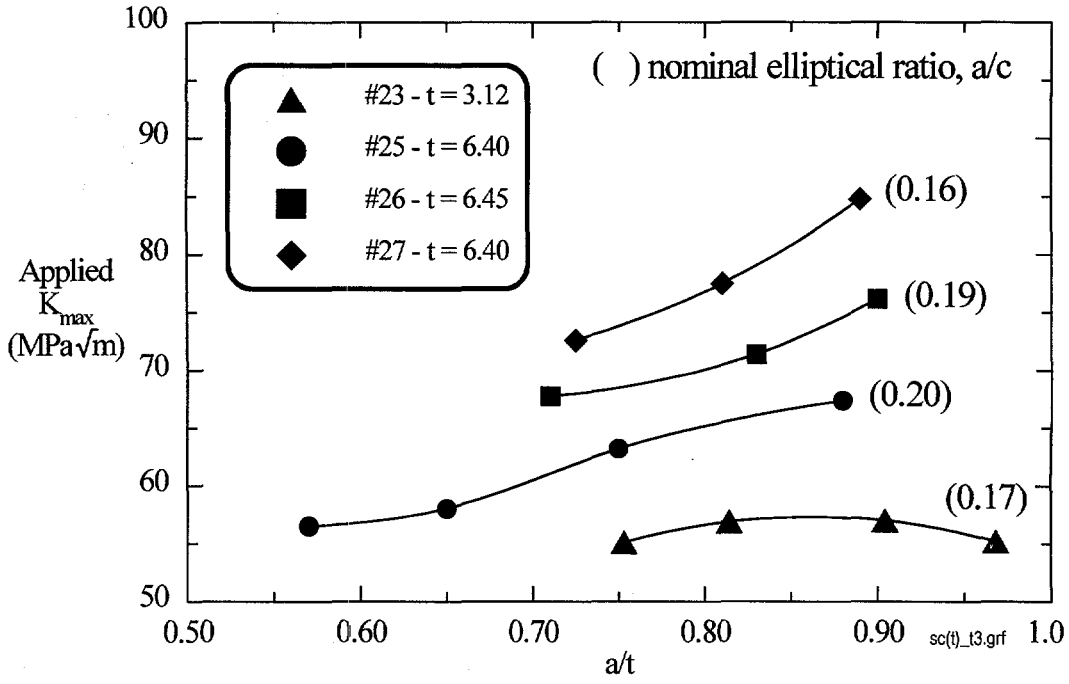


Figure 5. Crack initiation stress intensity factors for varying crack depths.

Chao and Reuter[14] examined several specimens loaded in bending and observed that initiation did not occur at the free surface even though K_{max} was calculated to occur at $\phi = 0$ degrees (see Figure 6). Of 19 specimens tested in bending, the crack growth initiation sites were located for 15 specimens. For these specimens, the crack initiation sites occurred at ϕ ranging from 45 to 9 degrees, with an average of $\phi = 21$ degrees (close to 30 degrees). The reason for $\phi = 21$ degrees is not yet understood.

These results show that the Toughness Ratio used in the FAD (see Figure 1) can be conservative if based on using applicable equations[9] for flat plate specimens (structures) containing surface cracks. (It is necessary to perform a similar series of tests for cylindrical sections.) A potential concern is that the conservatism may be excessive for specimens experiencing bending loads where K_{max} occurs at the free surface. Therefore, it is necessary to perform numerical analyses to develop an understanding of what controls the location of crack-growth initiation.

ELASTIC-PLASTIC CONDITIONS

For conditions in which K is no longer applicable, it is necessary to use either J or δ as the critical fracture toughness parameter. Unfortunately, the boundary separating LEFM from elastic-plastic fracture mechanics is not well defined.

Results

Specimens fabricated from ASTM A710 steel were tested at a temperature (22°C) corresponding to about midway in the ductile-brittle transition region. Reuter et al.[15] performed fracture toughness tests of C(T) (compact tension), SE(B) (single-edge notch bend), M(T) (center-cracked plate), and SC(T) (surface-cracked plates loaded in tension) specimens. Both multiple-specimen techniques and single-specimen techniques were used in the test procedure. For multiple specimens, several replicate specimens were loaded, each to a different value of load or displacement, and the specimens unloaded. Each of these specimens were sectioned, polished, and examined metallographically to measure the

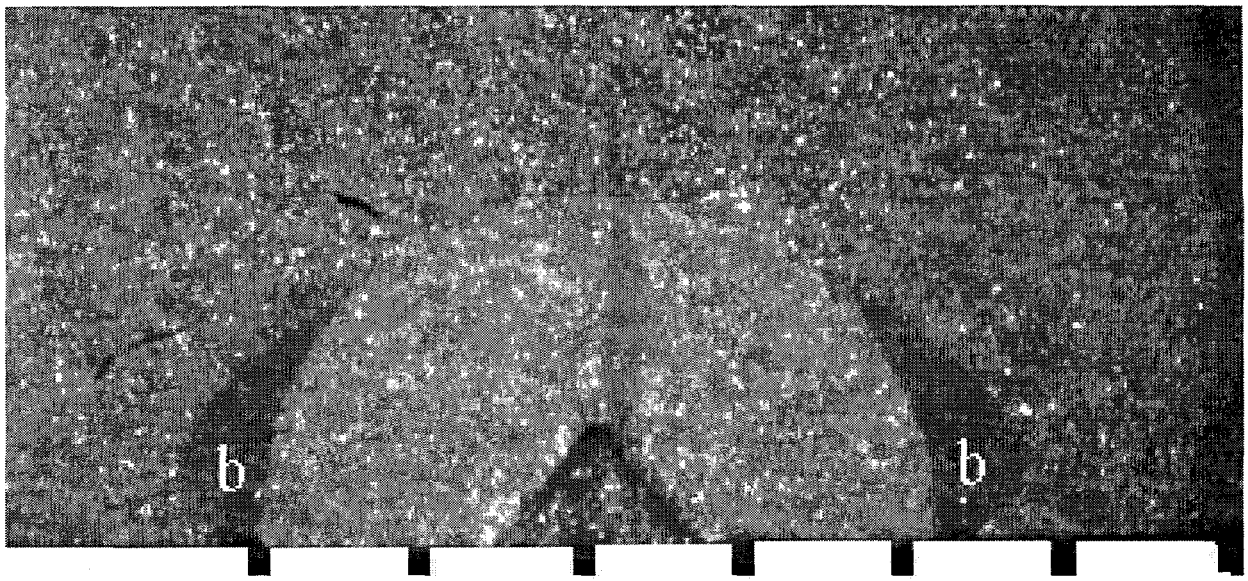


Figure 6. D6-aC Specimen #10 - Fracture surface showing stable growth near surface at "b." The dark region adjacent to the semicircular precrack is where crack growth occurred. (mm scale bars)

extent of crack growth and the corresponding value of δ. These results were used to obtain a plot of δ versus Δa, which was then extrapolated to $\Delta a = 0$ to estimate δ for crack growth initiation. For the single-specimen evaluation, a specimen was loaded until crack growth was detected, the load reduced, and either the specimen experienced cyclic loading to failure or it was loaded in liquid nitrogen to cause cleavage fracture, either of which will decorate the magnitude of stable crack growth. The two fracture surfaces were then examined using microtopography to measure δ at crack growth initiation. See Reuter and Lloyd[13] for additional discussion of the microtopography technique.

In addition, the constraint (hydrostatic stress normalized by dividing by the equivalent stress based on the von Mises yielding criteria) was calculated for each of the above specimen configurations, and a relation was observed between δ (crack growth initiation) and constraint. This relation was expanded and more completely developed in Hancock et al.[16]

In Reuter and Lloyd,[13] a series of tensile tests were performed in which measurements of δ and crack tip opening angle were made as a function of load at several locations around the perimeter of the surface cracks. These specimens were loaded in tension and had a a/2c ratio of 0.1 or 0.5. For specimens with a/2c = 0.1, crack growth initiation occurred at $\phi = 90$ degrees and disappeared at $\phi =$ 0 degrees. As the applied stress (σ) to σ_{ys} ratio approached 0.96, δ_{pl} (plastic component of δ) followed the prediction of Newman and Raju.[9] As $\sigma/\sigma_{ys} > 0.96$, there was a larger increase in δ_{pl} near the free surface than occurred elsewhere.

For a/2c = 0.5, crack growth initiation appeared to occur at $\phi = 90$ degrees and disappeared at $\phi =$ 0 degrees. (Some crack growth was visible around much of the surface crack perimeter, but none at $\phi =$ 0 degrees.) As σ/σ_{ys} approached 1.22, δ_{pl} followed the prediction of the Newman and Raju[9] equation. When $\sigma/\sigma_{ys} > 1.3$, a substantial increase in δ/δ_{el} (δ_{el} is an elastic component of δ) occurs, particularly at $\phi = 90$ degrees.

Discussion

The results in Reuter et al.[15] suggest that the use of δ measured per E1290 may provide too severe conservatism when predicting the condition for crack growth initiation of a structural component containing a shallow crack. If the relation between δ and constraint is known, it is possible to use a more realistic value for δ. The corresponding value of δ for the surface crack specimen was 1.4 times the value obtained in the more highly constrained SE(B) specimen (see Figure 7).

From a metallurgical point of view, the relationship between δ or J and constraint is unique because of the microstructure, trace elements, and thermal mechanical procedures used to develop the material to its final microstructure. Therefore, depending on the consequences of failure, it is strongly suggested that specimens/crack depth that provide the same constraint expected in the structure be tested to verify the magnitude of δ responsible for crack growth initiation.

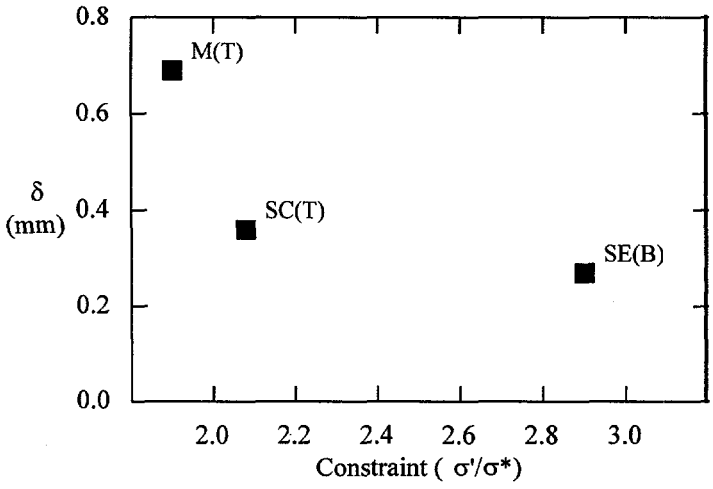


Figure 7. Crack tip opening displacement (at initiation) for different amounts of constraint.

The substantial changes in δ as a function of σ/σ_{ys} observed in Reuter and Lloyd[13] are probably due to loss of constraint. Parks and Wang[17] predict loss of HRR field dominance to occur at about $\sigma/\sigma_{ys} = 0.85$ and 1.04 for a/2c = 0.1 and 0.5, respectively. But constraint was still high at the load at which crack growth initiation occurred, and it was observed that crack growth initiation occurred at $\phi = 90$ degrees.

This suggests that since the surface-cracked specimen (structure?) and the middle-cracked tensile specimen (structure?) will have a constraint equal to or less than that of the test specimen designed per ASTM standards, that J/J_{Ic} or $\delta/\delta_{in} \geq 1.0$. Therefore, the Toughness Ratio in the FAD diagram of Figure 1 will be conservative for flat plate specimens loaded in tension.

CONCLUSIONS

Structural components were not available to determine the ability of using fracture toughness data to predict structural integrity, so specimens containing surface cracks were used instead. This is a logical step as the surface cracks used in the specimens encompass many of the configurations found in structural components but they generally experience only uniaxial loading..

These results cannot be used to evaluate crack size effects associated with differences between the specimen size and the structure size because the surface-cracked specimens were limited in size.

For LEFM conditions, crack growth initiation did not occur when $K_{max} = K_{Ic}$. The use of K_{Ic} and the Newman-Raju equation resulted in conservative predictions of crack growth initiation. This was true for surface-cracked specimens loaded in tension or in bending. For specimens tested in bending, the conservatism may be substantial when K_{max} occurs at the free surface, especially metallic specimens—the magnitude of the conservatism was reduced considerably for monolithic ceramics. These conclusions are limited to conditions in which crack growth initiation is detected. If stable crack growth

occurs, then the use of the maximum applied stress and the original crack size used in conjunction with the Newman-Raju equation to calculate K_{max} has no relevance to K_{Ic}.

At this time, it is not possible to answer if an average K (a K at a specific location or an additional parameter with a local K value) is responsible for initiation of crack growth for surface-cracked specimens. There is no basis to suggest that structural components will behave differently from the surface-cracked specimens tested. (Sufficient tests to evaluate the sensitivity of the fracture initiation process to multiaxial loading have not been performed.)

For NLEFM conditions, it is possible to predict crack growth initiation for specimens containing surface cracks if the relationship between δ and constraint has been quantified. This relation is known to vary as a function of material type and is expected to vary as a function of heat-to-heat variations within a given material. Therefore, once this relation is known for the specific material of interest, it is then necessary to quantify the constraint for the specific crack in the structural component. It is then possible to identify the critical value of δ associated with the specific constraint for predicting structural integrity.

The results for LEFM, where specimens are loaded in tension or in bending, and for elastic-plastic conditions where specimens are loaded in tension, the Toughness Ratio of the FAD, is conservative. The magnitude of the conservatism may become substantial for LEFM conditions where bending loads are encountered and K_{max} occurs at the free surface.

The above conclusions are limited to flat plates in tension or in bending for LEFM and flat plates in tension for elastic-plastic conditions and do not address cylinders.

ACKNOWLEDGMENT

This work was supported in part by the U.S. Department of Energy, Office of Energy Research, Office of Basic Energy Sciences, and Office of Industrial Technology under DOE Idaho Operations Office Contract DE-AC07-94ID13223.

REFERENCES

1 ASTM E399-90. Standard Test Method for Plane-Strain Fracture Toughness of Metallic Materials. *Annual Book of ASTM Standards*, Section 3, Vol. 03.01.

2 ASTM E813-89. Standard Test Method for J_{Ic}, A Measure of Fracture Toughness. Ibid.

3 ASTM E1290-89. Standard Test Method for Crack-Tip Opening Displacement (CTOD) Fracture Toughness Measurement. Ibid.

4 ASTM E1221-88. Standard Test Method for Determining Plane-Strain Crack-Arrest Fracture Toughness, K1a of Ferritic Steels. Ibid.

5 ASTM E1152-87. Standard Test Method for Determining J-R Curves. Ibid.

6 Xia, L., C. F. Shih, and J. W. Hutchinson (1995). A Computational Approach to Ductile Crack Growth Under Large Scale Yielding Conditions. *J. Mech. Phys. Solids 43*, 389-414.

7 Dadkhah, M. S. and A. S. Kobayashi (1994). Two-Parameter Crack Tip Stress Field Associated With Stable Crack Growth in Thin Plate. *Fracture Mechanics, 24th Vol.*, ASTM STP 1207, J. D. Landes, D. E. McCabe, and J. A. M. Boulet (eds.), 48-61.

8 May, G. B. and A. S. Kobayashi (1995). Plane stress stable crack growth and J-integral/HRR field. *Int'l J. Solids and Struct.* 37/617, 857-881.

9 Newman, J. C. and I. S. Raju (1978). Analysis of Surface Cracks in Finite Plates Under Tension or Bending Loads. NASA Technical Paper 1578, National Aeronautics and Space Administration, Washington, D.C.

10 Reuter, W. G., J. C. Newman, Jr., B. D. Macdonald and S. R. Powell (1994). Fracture Criteria for Surface Cracks in Brittle Materials. *Fracture Mechanics: Twenty-Fourth Volume*, ASTM STP 1207, J. D. Landes, D. E. McCabe, and J. A. M. Boulet (eds.), American Society for Testing and Materials, Philadelphia, PA, 1994, 617-635.

11 Reuter, W. G., N. Elfer, D. A. Hull, J. C. Newman, Jr., D. Munz and T. L. Ponontin (1996). Test Results and Preliminary Analysis for International Cooperative Test Program on Specimens Containing Surface Cracks. ASTM-sponsored *28th National Symposium on Fatigue and Fracture, Albany, NY, June 25-27, 1996.*

12 Sommers, A. and D. Aurich (1991). On the Effect of Constraint on Ductile Fracture. *Defect Assessment in Components - Fundamentals and Applications*, J. G. Bland and K.-H. Schwalbe (eds.), ESIS/EGF Publication 9, 141-174.

13 Reuter, W. G. and W. R. Lloyd (1990). Measurement of CTOD and CTOA Around Surface-Crack Perimeters and Relationships Between Elastic and Elastic-Plastic CTOD Values. *Surface-Crack Growth, Experiments, and Structures*, ASTM STP 1060, W. G. Reuter, J. H. Underwood, and J. C. Newman, Jr. (eds.), American Society for Testing and Materials, Philadelphia, PA, 152-176.

14 Chao, Y. J. and W. G. Reuter (1996). Fracture of Surface Cracks. ASTM-sponsored *28th National Symposium on Fatigue and Fracture, Albany, NY, June 25-27, 1996.*

15 Reuter, W. G., S. M. Graham, W. R. Lloyd and R. L. Williamson (1991). Ability of Using Experimental Measurements of δ to Predict Crack Initiation for Structural Components. *Defect Assessments in Components - Fundamentals and Applications*, J. G. Bland and K.-H. Schwalbe (eds.), 1991 Mechanical Engineering Publication, London, 175-188.

16 Hancock, J. W., W. G. Reuter, and D. M. Parks (1993). Constraint and Toughness Parameterized by T. *Constraint Effects in Fracture, ASTM STP 1171*, E. M. Hackett, K.-H. Schwalbe, and R. H. Dodds (eds.), American Society for Testing and Materials, Philadelphia, PA, 21-40.

17 Parks, D. M. and Y. Y. Wang (1988). Elastic-Plastic Analysis of Part-Through Surface Cracks. Proceedings of the Symposium on Analytical, Numerical, and Experimental Aspects of Three-Dimensional Fracture Process, American Society of Mechanical Engineers, Berkeley, CA, June 1988.

Multisensor Based Robotic Manipulation in an Uncalibrated Manufacturing Workcell

Bijoy K. Ghosh, Di Xiao, Ning Xi and Tzyh-Jong Tarn
Campus Box 1040, Washington University
One Brookings Drive
Saint Louis, MO 63130-4899, USA
(314) 935 6039 (Phone)
(314) 935 6121 (Fax)
ghosh@zach.wustl.edu

Abstract

The main problem that we address in this paper is how a robot manipulator is able to track and grasp a part placed arbitrarily on a moving disc conveyor aided by a single CCD camera and fusing information from encoders placed on the conveyor and also from encoders on the robot manipulator. The important assumption that distinguishes our work from what has been previously reported in the literature is that the position and orientation of the camera and the base frame of the robot is apriori assumed to be unknown and is 'visually calibrated' during the operation of the manipulator. Moreover the part placed on the conveyor is assumed to be non-planar, i.e. the feature points observed on the part is assumed to be located arbitrarily in $\mathbb{R}^3$. The novelties of the proposed approach in this paper includes a (i) *multisensor fusion scheme* based on complementary data for the purpose of part localization, and (ii) *self-calibration* between the turntable and the robot manipulator using visual data and feature points on the end-effector. The principle advantages of the proposed scheme are the following. (i) It renders possible to reconfigure a manufacturing workcell without recalibrating the relation between the turntable and the robot. This significantly shortens the setup time of the workcell. (ii) It greatly weakens the requirement on the image processing speed.

1 Introduction

Sensor-guided tracking plays an important role in today's flexible manufacturing systems. Using sensors, a manufacturing system can compensate for changes in environments and uncertainties in its model. Vision becomes an especially useful sensor in Robotics. Use of visual information provides a way of overcoming some difficulties of uncertain models and unknown environment and hence extends the domain of applications of robots without explicit intervention or reprogramming.

Control of robot manipulators with vision in the feedback loop has an exciting history starting probably with the pioneering work of Hill and Park [1], Weiss, Sanderson and Neuman [2]. Subsequent work in the area has focused on "Visual Servoing" wherein the emphasis is to visually locate the position and orientation of an object and to control a robot manipulator to grasp and manipulate the object. If the object is not stationary, then the process of locating the object and repositioning the robot through control must be repeated iteratively until the task has been accomplished. This leads naturally to real time vision based feedback and control problems that have been subsequently studied in [3], [4], [5], [6], [7] and probably many others. As a result, many important tasks, such as Bolt Insertion, Conveyor Belt Picking, Weld Seal Tracking, Part Mating, Road Vehicle Guidance, Juggling, Fruit Picking etc, to name a few, have been accomplished with the aid of computer vision.

However, one has to overcome many difficulties in order to utilize visual information. First, visual data is not always reliable. Vision systems could occasionally fail to generate any useful information but noise due to variation of illumination, overlapping of different workpieces or accidental obstruction of the camera. Secondly, image processing algorithms are always time-consuming. Therefore, direct use of visual information for robot control purpose will lead to a poor accuracy or even, in the case of dynamic visual servoing, can cause stability

problem. In addition, most of previous work in visual servoing area required precise calibration before control schemes were implemented. As is well-known, calibration is also a time-consuming job and makes visual servo control expensive or even impossible in some cases.

We propose in this paper, a new visually guided planning and control of a robot manipulator for precise positioning and tracking in an uncertain and dynamic environment. The proposed scheme is robust against precise position of the camera and utilizes a scheme called *virtual rotation* wherein the observed image is transformed to what it would appear if the camera has been pointed vertically down. Additionally the proposed scheme is also robust against uncertainty in precise position of the robot manipulator.

In this paper, the real time part localization does not entirely depend on the data from the CCD camera. Information from encoders attached to the rotating turntable is also used in a new "Multiple Sensor Integration" scheme, which greatly reduces the requirement on the image processing speed while obtaining a good motion estimation. In addition, the proposed scheme introduces the idea of considering complimentary data (instead of redundant ones as in traditional multisensor fusion schemes) in a fusion process.

This paper also introduces a new 'self-calibration' scheme in order to ascertain the position and orientation of the robot manipulator in the workcell. The scheme utilizes a set of apriori chosen points on the end-effector to compute the coordinate transformation between the base of the manipulator and a certain fixed coordinate attached to the workcell. (iii) Although our approach is feature based, we do not propose to visually estimate structure and motion. The estimation scheme presented in this paper differs considerably from standard results in motion and structure estimation from feature correspondences (see [10]). As a matter of fact, on one hand, the motion of the moving part in a fixed disc frame is obtained by fusing visual information from the camera with readings from an encoder attached to the rotating disc conveyor. On the other hand, the structure information in terms of the orientation of the end-effector, is measured with respect to the base coordinate frame of the robot using the encoders attached to the robot. One of the main issues addressed in this paper is "how to fuse information obtained by the visual and the encoder sensors". The main conclusion is that the procedure of 'multiple sensor integration' already introduced in [8] leads to a unique self-calibration scheme which is reliable and has been tested via experiments (iv) New planning and control schemes in task space are also discussed.

The paper is structured as follows. We emphasize the real-time part localization scheme of a three dimensional part on the turntable in sufficient details. In so doing, we show that even when the height of the part is unknown, the relative position of the part can be computed visually. The self calibration scheme, described in this paper, has already been reported in [8]. Thus the details of this scheme are omitted.

2 Multisensor Based Visual Sensing

The sensor fusion scheme discussed in this paper combines three sources of sensory data to obtain measurements that individual sensors cannot obtain. The scheme is described as follows. From visual information, we recover the relative position and orientation of the part with respect to the reference line joining the disc center and the reference point on the turntable. Since the relative pose of the part does not change over time, from reading the encoder of the turntable the real-time pose of the part with respect to the fixed disc frame can be obtained. Combining the visual information with data from encoders of the robot and the turntable, we also can determine the relation between the fixed disc frame and the base frame of the robot. Finally, we compute the real-time pose of the part in the base frame of the robot.

In this paper, the position and orientation of the part are determined based on feature points. Fundamentally, in our case, the problem of determining the pose of the part can be reduced to that of determining the position of a point in some coordinate frame under the assumptions that the plane of the turntable and XY-plane of the base frame of the robot are parallel and that the part to be manipulated has a simple known shape.

Let $OXYZ$ be any Cartesian coordinate frame with its origin O at the disc center and its Z-axis perpendicular to the turntable. After virtual rotation of the camera, the position of a point p in 3-D space can be represented by

$$X_p = \frac{aX_{ref} - bY_{ref}}{l^2} + \frac{(\bar{a}-a)X_{ref} - (\bar{b}-b)Y_{ref}}{lLf}h; \tag{2.1}$$

$$Y_p = \frac{bX_{ref} + aY_{ref}}{l^2} + \frac{(\bar{b}-b)X_{ref} + (\bar{a}-a)Y_{ref}}{lLf}h, \tag{2.2}$$

with

$$a = \overrightarrow{C_oP_o} \cdot \overrightarrow{C_oR_o}; \qquad b = \overrightarrow{C_oP_o} \times \overrightarrow{C_oR_o}; \tag{2.3}$$

$$\bar{a} = \overrightarrow{C_oO_o} \cdot \overrightarrow{C_oR_o}; \qquad \bar{b} = \overrightarrow{C_oO_o} \times \overrightarrow{C_oR_o}; \tag{2.4}$$

where $\overrightarrow{AB}$ denotes the vector from A to B while $\cdot$ and $\times$ are operators of dot product and cross product, respectively. C_o, R_o and P_o are the transformed images of the disc center, the reference point and the point p via perspective projection, respectively. O_o is the intersection of the optical axis and the transformed image plane. h stands for the distance of the point p from the turntable. $(X_{ref}, Y_{ref}, 0)$ are the coordinates of reference point in the frame $OXYZ$ and (X_p, Y_p, h) the coordinates of the point p in the same frame. L denotes the length of the reference line while l is the length of the transformed image of the reference line; f is the focal length of the camera. Note that a, b, $\overline{a}$, $\overline{b}$ and l can be easily computed in terms of the image coordinates of C_o, R_o and P_o. In our configuration, L and f are known constants. For the fixed camera, l is also a constant. Hence, if the coordinates (X_{ref}, Y_{ref}) of the reference point are given in the frame $OXYZ$, the position of the point p in the same frame is an affine function of h, the distance of the point from the turntable.

2.1 Determination of relative position of a point

Suppose that X-axis of the fixed disc frame $O_dX_dY_dZ_d$ coincides with the initial position of the reference line joining the disc center and the reference point and that X-axis of the attached disc frame $O_aX_aY_aZ_a$ is along with the real-time position of the reference line. Therefore, the two frames are exactly the same initially. The coordinates of the reference point on the turntable in the attached disc frame are $(L, 0, 0)$.

From (2.1) and (2.2), the relative position of a point which has distance of h away from the turntable, with respect to the attached disc frame, can be written as

$$({}^aX_p,\ {}^aY_a,\ {}^aZ_p) = (\frac{aL}{l^2} + \frac{(\overline{a}-a)}{lf}h,\ \frac{bL}{l^2} + \frac{(\overline{b}-b)}{lf}h, h).$$

with

$$\begin{cases} a = (x_r - x_c)(x_p - x_c) + (y_r - y_c)(y_p - y_c); \\ b = (x_p - x_c)(y_r - y_c) - (y_p - y_c)(x_r - x_c); \\ \overline{a} = x_c(x_c - x_r) + y_c(y_c - y_r); \\ \overline{b} = y_c(x_r - x_c) - x_c(y_r - y_c); \end{cases} \tag{2.5}$$

where (x_c, y_c), (x_r, y_r) and (x_p, y_p) are transformed image coordinates of the disc center, the reference point and the point p, respectively.

Especially, for the points on a part of negligible height, we obtain their coordinates in the attached disc frame $(\frac{aL}{l^2}, \frac{bL}{l^2}, 0)$.

Suppose that the distance of a point on the part from the turntable is unknown. we cannot hope to determine the relative position of the point via a single image. However, since the part is moving as the turntable rotates and the relative position of the part with respect to the attached disc frame does not change over time, hence we can also determine the relative position of the part from multiple images taken at different time instants.

Let us consider a point p on the part. Suppose two different images are taken at time t_1 and t_2, respectively. From the image taken at t_i $(i = 1, 2)$, we have the same relative position of the point with respect to the attached disc frame, namely,

$$({}^aX_p,\ {}^aY_p,\ {}^aZ_p) = (\frac{a_iL}{l^2} + \frac{(\overline{a}_i - a_i)}{lf}h,\ \frac{b_iL}{l^2} + \frac{(\overline{b}_i - b_i)}{lf}h, h),$$

where a_i, b_i, $\overline{a}_i$ and $\overline{b}_i$ are computed by means of Eqn.(2.5) with $(x_c(t_i), y_c(t_i))$, $(x_r(t_i), y_r(t_i))$ and $(x_p(t_i), y_p(t_i))$ being transformed image coordinates of the disc center, the reference point and the point p at time t_i, respectively. Therefore, it is seen that

$$\underbrace{\begin{bmatrix} \overline{a}_1 - a_1 + a_2 - \overline{a}_2 \\ \overline{b}_1 - b_1 + b_2 - \overline{b}_2 \end{bmatrix}}_{A} h = \underbrace{\begin{bmatrix} a_2 - a_1 \\ b_2 - b_1 \end{bmatrix}}_{B} \frac{fL}{l}. \tag{2.6}$$

from which one can get a least squire solution, namely $h = (A^TA)^{-1}A^TB$.

As a result, the relative position of the point p can easily be determined. It should be pointed out that so far we have only used the visual information for determination of the relative position of a point on the part with respect to the attached disc frame via one or two images.

2.2 Determination of the relation between the fixed disc frame and the base frame

In order to determine the relation between the fixed disc frame and the base frame of the robot, we need to describe a set of points in both frames since the two frames are related by

$$^bP = {}^bR_d\ {}^dP + {}^bT_d,$$

where ^bP and ^dP are the coordinates of a point in the base frame and the fixed disc frame, respectively. We assume that the plane of the turntable and XY-plane of the base frame are parallel. It is seen that there is one unknown in the rotation matrix ^bR_d. ^bT_d has three unknown elements. For a point q, if we know its coordinates in the both frames, i.e. ^bP_q and ^dP_q, then from the last equality we have three equations for the four unknown variables. Obviously, in order to get a unique solution to the relation, we need to know at least two points in the both frames. Fortunately, from reading encoders of the robot, the coordinates of points on the end-effector with respect to the base frame of the robot can be readily obtained. In what follows, we describe the points on the end-effector in the fixed disc frame with the aid of the single camera.

Suppose that an image was taken at time t. At the very moment the coordinates of the reference point in the fixed disc frame are $(L\cos(\theta(t)), L\sin(\theta(t)), 0)$.

After computing the corresponding a, b, $\overline{a}$ and $\overline{b}$ via the image data, the coordinates of a point on the end-effector in the fixed disc frame can be easily obtained by substituting the results into (2.1) and (2.2), as long as the distance of the point from the turntable is known.

Suppose that the distance between the plane of the turntable and the XY-plane of the base frame is unknown apriori. Now, we need to determine the distance of the point from the turntable in order to describe the point in the fixed disc frame.

In [9] we proposed an algorithm to estimate the distance by considering two basic cases: two points or three points on the end-effector are observed by the camera. Of course, more than three points may be employed to increase robustness against the possible noise in data. To save space, we do not give details here. For those two cases, by combining visual information with information from encoders of the robot, we have a set of quadratic equations of one unknown. The main results we obtained in [9] include the following. In case of two points, there are at most two solutions. In many cases, the unique solution can be determined by taking advantage of physical constraints. In the case where three points on the end-effector are observed, the solution can be uniquely determined. Moreover, in this case, we do not need know the individual correspondence between the points and their images. A by-product is that we can determine the actual individual correspondence as well as the unique solution.

It should be pointed out that similar problem (so-called 3D to 2D problem) has been studied in the computer vision field [10]. As you have seen above, however, the results obtained in [9] are much stronger both regarding the uniqueness of the solution and regarding the correspondence between points and their images because in our cases additional information coming from the robot can be utilized.

2.2.1 Computation of the relation between the frames

Having described at least two points in both the fixed disc frame and the base frame of the robot, we obtain

$$\begin{bmatrix} {}^bx_i \\ {}^by_i \\ {}^bz_i \end{bmatrix} = {}^bR_d \begin{bmatrix} {}^dx_i \\ {}^dy_i \\ {}^dz_i \end{bmatrix} + {}^bT_d, \quad i = 1, 2, \cdots$$

where $({}^bx_i,\ {}^by_i,\ {}^bz_i)$ and $({}^dx_i,\ {}^dy_i,\ {}^dz_i)$ are the coordinates of the i-th point with respect to the base frame and the fixed disc frame, respectively. Recall the assumption that the plane of the turntable and XY-plane of the base frame are parallel, we know that ^bR_d has the following structure

$$^bR_d = \begin{bmatrix} r_{11} & r_{12} & 0 \\ -r_{12} & r_{11} & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{2.7}$$

In this case, we can determine ${}^bR_d$ and ${}^bT_d$ with knowing two points in the two frames. As a matter of fact, it is seen that

$$\begin{bmatrix} {}^bx_2 - {}^bx_1 \\ {}^by_2 - {}^by_1 \\ {}^bz_2 - {}^bz_1 \end{bmatrix} = {}^bR_d \begin{bmatrix} {}^dx_2 - {}^dx_1 \\ {}^dy_2 - {}^dy_1 \\ {}^dz_2 - {}^dz_1 \end{bmatrix} \tag{2.8}$$

Let

$$\begin{bmatrix} {}^be_1 \\ {}^be_2 \\ {}^be_3 \end{bmatrix} = \begin{bmatrix} {}^bx_2 - {}^bx_1 \\ {}^by_2 - {}^by_1 \\ {}^bz_2 - {}^bz_1 \end{bmatrix}, \quad \begin{bmatrix} {}^de_1 \\ {}^de_2 \\ {}^de_3 \end{bmatrix} = \begin{bmatrix} {}^dx_2 - {}^dx_1 \\ {}^dy_2 - {}^dy_1 \\ {}^dz_2 - {}^dz_1 \end{bmatrix}$$

Eqn.(2.8) represents linear equations of r_{11} and r_{12}. As long as the line joining the two points is not parallel to z-axis of the base frame, the equations always have a unique solution. However, such a solution may not satisfy the constraint $r_{11}^2 + r_{12}^2 = 1$ due to the possible noise in observed data and computation errors. In other words, ${}^bR_d$ obtained in this way may not be orthogonal and is therefore meaningless.

Actually, the problem of determining ${}^bR_d$ can be viewed as the optimization problem of determining ${}^bR_d$ with the structure in (2.7) such that

$$\| \begin{bmatrix} {}^be_1 \\ {}^be_2 \\ {}^be_3 \end{bmatrix} - {}^bR_d \begin{bmatrix} {}^de_1 \\ {}^de_2 \\ {}^de_3 \end{bmatrix} \|_2$$

is minimized subject to the constraint $r_{11}^2 + r_{12}^2 = 1$. Solving this optimization problem yields

$$\begin{bmatrix} r_{11} \\ r_{12} \end{bmatrix} = \frac{1}{\sqrt{{}^be_1^2 + {}^be_2^2}\sqrt{{}^de_1^2 + {}^de_2^2}} \begin{bmatrix} {}^be_1\,{}^de_1 + {}^be_2\,{}^de_2 \\ {}^be_1\,{}^de_2 - {}^be_2\,{}^de_1 \end{bmatrix}$$

With knowing ${}^bR_d$, we have

$${}^bT_d = \begin{bmatrix} {}^bx_1 \\ {}^bx_2 \\ {}^bx_3 \end{bmatrix} - {}^bR_d \begin{bmatrix} {}^dx_1 \\ {}^dx_2 \\ {}^dx_3 \end{bmatrix}.$$

Now, we can determine any observed point on the part with respect to the base frame of the robot. Based on the recovered coordinates of the points on the part, we can easily know the position of the centroid of the part and the orientation of the part.

3 Experiments

An experimental system has been set up in the Center for Robotics and Automation at Washington University. It consists of one PUMA 560 manipulator and a turntable with the diameter of 0.9m. On the turntable the center and a reference point are marked so that they can clearly be seen by the vision system. The distance between the center and the reference point is 0.295 m. The computer vision system consists of a CCD camera with image resolution of 256×256 and the Intelledex Vision processor based on a 16 MHz Intel 80386 CPU. The focal length of the camera is 0.0125 m. The vision system interfaces to the host computer, a SGI 4D/340 VGX. Visual measurements are sent to SGI by a parallel interface. The robot is controlled by UMC controller that also interfaces to SGI through memory mapping.

In the experiments, a non-planar part was randomly placed on the turntable. The precise locations of camera and turntable in the base frame of the robot are both unknown, though we assume that the plane of turntable and XY-plane of the base frame of the robot are parallel. The robot has been successfully controlled to pick up the part and then drop it to a prespecified place. In this paper, we focus on the proposed multisensor-based calibration-free schemes for localization of the part in the base frame of the robot. The experimental results regarding robot control are not presented here.

In the experiments, two tasks have been completed. One is to control the robot to pick up a part after the turntable stops rotation. We call it static picking up task. The other one is to require the robot to track and pick up a moving part while the turntable is rotating. We call it tracking and picking up task. We observed two points on the end-effector. Recall that if the line joining the two points is approximately parallel to the turntable then a unique solution for the distance of the end-effector from the turntable can be determined. Therefore, we can determine the position and orientation of the part in the base frame of the robot.

The experimental results for static picking up task are indicated in Fig. 4. The actual and estimated trajectory of the centroid of the part in base frame of the robot in Fig. 4 (a) , while Fig. 4 (b) provides the actual and estimated

orientation of the part with respect to the base frame of the robot. Fig. 4 clearly shows the procedure during which the robot end-effector is tracking, picking up a moving part and drop it to a pre-specified location. In the figure, the part is undergoing a circular motion in a horizontal plane and the robot end-effector moves from some place above the plane, approaches to the part, grasps it and then goes straightforward to the pre-specified place.

In the experiment for the static picking task, we have taken one hundred images. For each image, we can recover the position and orientation of the part in the base frame by using the proposed multisensor fusion scheme. To deal with possible noise in observed data, we choose average of all the results as the final estimation. Note that the relative position and orientation of the part with respect to the turntable are time-invariant. Once the position of the turntable in the base frame and the relative pose of the part with respect to the turntable are determined, one can easily know the trajectory of the part in the base frame via encoder reading. In the experiment for the tracking and picking task, fifty images have continuously been taken. In this case, we take the last computation results of the pose of the part as our estimate and feedback the values to the planner. This is why no "shaking" part can be seen in Fig. 4. Because there are some estimation errors in the position of the center of the turntable, even after completion of the estimation the trajectory error of the part is a periodic function of time due to the rotation of the turntable. It should be pointed out that the estimation errors are within 0.01 (m). In our experiments, if the position errors of the part are within 0.01 (m) and the orientation errors are within 5 degrees, the picking tasks can be successfully accomplished.

The above experiments have verified the proposed estimation scheme. It does not just provide a Calibration-Free Vision for part localization, but also significantly reduces the requirement for processing speed of the vision system. In these experiments, a very primitive vision system with a low speed CPU was used. Nevertheless, it is capable of providing the position and orientation estimations for a fast moving part.

4 Conclusion

In this paper we developed a sensor fusion scheme estimate the position and orientation of parts in an uncalibrated environment in order to manipulate the parts in a typical manufacturing workcell that is composed of a robot manipulator, a rotating turntable and a camera system. Even though the visual computations are performed in low rate, part position and orientation information can still be updated in same rate of feedback loop using an additional encoder sensor. We also demonstrate a practical tracking algorithm which takes into account the fact that the torque that the robot control system can supply is bounded. The proposed algorithm is primarily based on error feedback with an extra error reduction term added in order to force the required torque requirement to remain within acceptable bounds. The main features of our scheme are (i) requirement of speed for image processing is reduced and hence cost of the system is low; (ii) the scheme can handle uncalibrated environment and therefore has certain flexibility, allowing possible reconfiguration of a manufacturing workcell without recalibrating the relation between the turntable and the robot and also (iii) the scheme can be used in various tasks, e.g. peg-in-hole in a horizontal plane and the like. The experimental results clearly demonstrate the advantages of the proposed scheme.

References

[1] J. Hill and W. T. Park , Real time control of a robot with a mobile camera, Proc. of 9th ISIR, Washington D.C., March 1979, pp.233-246.

[2] L. Weiss, A. Sanderson, and C. Neuman, Dynamic sensor based control of robots with visual feedback, *IEEE J. of Robotics and Automation*, **RA-3(5)**:404–417, 1987.

[3] B. H. Yoshimi and P. K. Allen, Alignment using an uncalibrated camera system, *IEEE J. of Robotics and Automation*, **RA11(4)**:516–521, 1995.

[4] E. D. Dickmanns and V. Graefe, Dynamic monocular machine vision, *Machine Vision and Applications*, 1:223–240, 1988.

[5] J. T. Feddema and O. R. Mitchell, Vision-guided servoing with feature-based trajectory generation, *IEEE J. of Robotics and Automation*, **RA-5(5)**:691–700, 1989.

[6] N. Papanikolopoulos, P. K. Khosla, and T. Kanade, Visual tracking of a moving target by a camera mounted on a robot: a combination of control and vision, *IEEE J. of Robotics and Automation*, **RA-9(1)**:14–35, 1993.

[7] P. I. Corke, Visual control of robot manipulators–a review, In K. Hashimoto, editor, *Visual Servoing*, pages 1–32. World Scientific, 1994.

[8] B. K. Ghosh, T. J. Tarn, N. Xi, Z. .Y Yu and Di Xiao, Visually controlled manipulation of parts in a robotic manufacturing workcell, the Proc. of the 1996 IEEE Trans. Robotics and Automation, Minnesota, U.S.A., April, 1996.

[9] Bijoy K.Ghosh, Di Xiao, Ning Xi and Tzyh Jong Tarn, Calibration-free vision based control of a robotic manipulator, The 34th Annual Allerton Conference on Communication, Control and Computing, Allerton, Illinois, USA, Oct.2-4,1996.

[10] T. S. Huang, A. N. Netravali, Motion and structure from feature correspondences: A Review, Proc. of the IEEE, Vol.82, No.2, Feb. 1994, pp.252-268.

[11] R. I. Hartley, Estimation of relative camera positions for uncalibrated cameras, Proc. of 2nd E.C.C.V., Santa Margherita Ligure, Italy, May 1992, pp.579-587.

[12] M. A. Fischler and R. C. Bolles, Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography, Communications of the ACM, Vol.24, NO.6, June 1981, pp.381-355.

[13] Y.Wu, S. S. Iyengar, R. Jain, S.Bose, A new generalized computational framework for finding object orientation using perspective trihedral angle constraint, IEEE Trans. on PAMI, Vol.16, No.10, Oct. 1994, pp.961-975.

[14] K.Kanatani, Group-Theoretical Methods in Image Understanding, Springer-Verlag, 1990.

[15] P.Allen and R.Bajcsy, Object recognition using vision and touching, Proc. of the 9th Joint Conf. on Artificial Intelligence, Los Angeles, Aug. 1985.

[16] A. J. Koivo and N. Houshangi, Real-time vision feedback for servoing robotic manipulator with self-tuning controller. IEEE Trans. on Systems, Man and Cybernetics, Vol.21, No.1, Jan/Feb. 1991, pp.134-142.

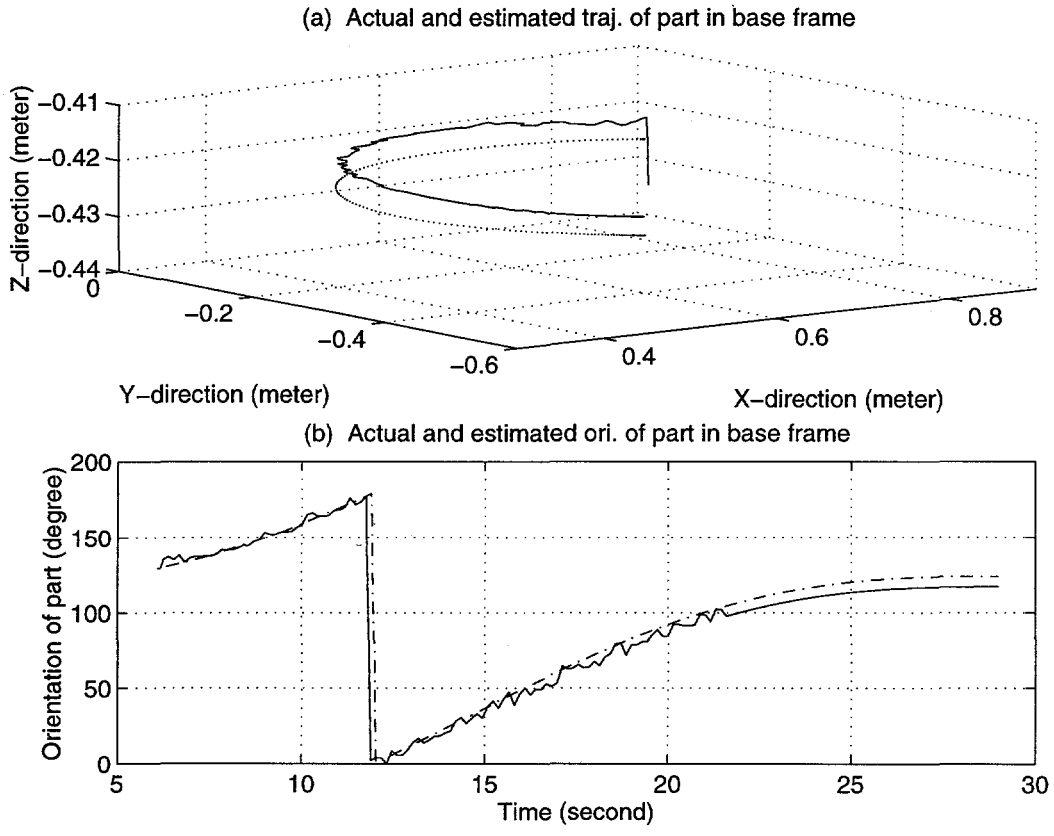


Figure 4.1: The actual and estimated pose of the part in the base frame of the robot for the static picking task.

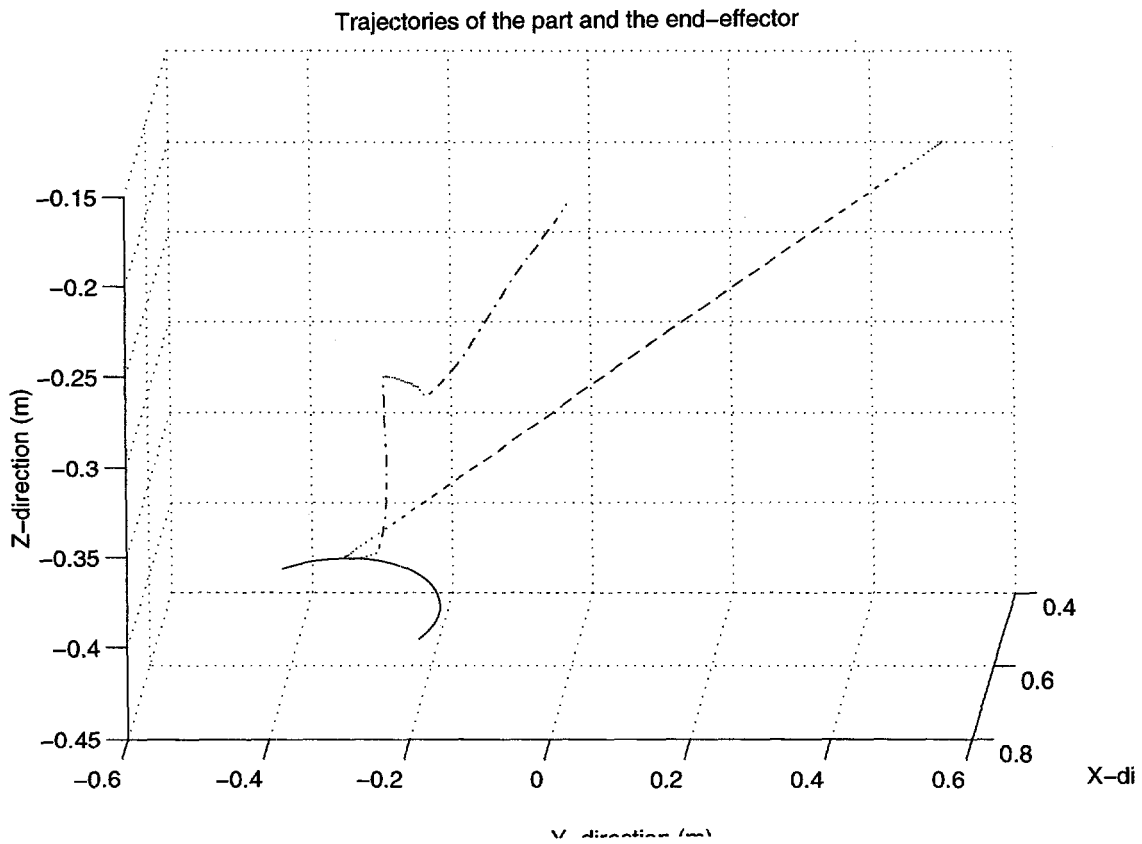


Figure 4.2: The trajectories of the moving part and the end-effector of the robot in the tracking and picking task.

BIFURCATION AND NECK FORMATION AS A PRECURSOR TO DUCTILE FRACTURE DURING HIGH RATE EXTENSION

L. B. Freund and N. J. Sörensen*

Division of Engineering, Brown University, Providence, RI 02912

Abstract

A block of ductile material, typically a segment of a plate or shell, being deformed homogeneously in simple plane strain extension commonly undergoes a bifurcation in deformation mode to nonuniform straining in the advanced stages of plastic flow. The focus here is on the influence of material inertia on the bifurcation process, particularly on the formation of diffuse necks as precursors to dynamic ductile fracture. The issue is considered from two points of view, first within the context of the theory of bifurcation of rate-independent, incrementally linear materials and then in terms of the complete numerical solution of a boundary value problem for an elastic-viscoplastic material. It is found that inertia favors the formation of relatively short wavelength necks as observed in shaped charge break-up and dynamic fragmentation.

INTRODUCTION

During extension of a ductile plate or radial expansion of a ductile shell, deformation is often found to proceed more or less homogeneously within the plastic range until it is interrupted by the formation of localized necks or shear bands. These regions of localized deformation commonly evolve into ductile fractures. The focus here is on the formation of such necks during high rate extension, and on the influence of inertia on the stress level necessary for neck formation.

To this end, a rectangular block of an incompressible elastic–plastic material deforming at high rate under plane strain conditions is considered; see Figure 1. The block can be viewed as a segment of a plate or shell when symmetry is enforced. Opposite ends of the block are subjected to a uniform normal velocity V_0 in the 1–direction. Otherwise, the faces of the block are free of traction. The goal is to establish conditions on loading, geometry and material properties under which the homogeneous deformation can give way to a non–homogeneous deformation.

Up to the instant of bifurcation, the rate of deformation field throughout the material is essentially uniform. Material coordinates X_k referred to an underlying cartesian basis are introduced in the reference configuration for the block, which is taken to occupy the region $-\ell_1 \le X_1 \le \ell_1$, $-\ell_2 \le X_2 \le \ell_2$. The dominant material velocity field referred to these coordinates is

$$v_1^0(X_1, X_2, t) = \frac{X_1 V_0}{\ell_1}, \qquad v_2^0(X_1, X_2, t) = -\frac{X_2 V_0}{\ell_1 \left(1 + V_0 t/\ell_1\right)^2} \tag{1}$$

*Now at Department of Mechanical Engineering, Lund University, Sweden.

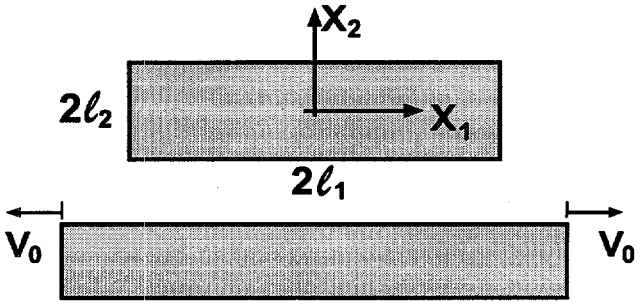


Figure 1: A reference configuration (top) for a block of material, and a subsequent configuration (bottom) achieved by homogeneous plane strain extension.

The instantaneous stretching rate in the 1–direction is V_0/ℓ_1.

It is noteworthy that the deformation (1) cannot be maintained by a homogeneous state of uniaxial stress. The pathlines of material particles are curves in space, which implies a nonuniform distribution of particle acceleration throughout the block. The equations enforcing the balance of linear momentum lead to the conclusion that the stress state is a homogeneous tension in the 1–direction, say σ, plus a hydrostatic pressure which varies in the 2–direction. In terms of components of Cauchy stress σ_{ij}, the distribution has the form $\sigma_{11} = \sigma - p_I$, $\sigma_{22} = -p_I$ and $\sigma_{12} = 0$. In terms of material coordinates in the reference configuration, the pressure p_I has the parabolic distribution

$$p_I(X_2) = \frac{\rho V_0^2}{\ell_1^2} \frac{\ell_2^2 - X_2^2}{(1 + V_0 t/\ell_1)^4} \tag{2}$$

where ρ is the material mass density.

A number of studies of the conditions for bifurcation of the elastic–plastic block under quasistatic imposed extension or compression have been reported. Hill and Hutchinson [1] identified the regimes of behavior and ranges of moduli for which the governing equations are elliptic, parabolic or hyperbolic. They also calculated the spectrum of bifurcation stresses for symmetric and antisymmetric diffuse deformation modes. Young [2] carried out a similar analysis for plane strain compression. Needleman [3] extended the analysis for plane strain tension/compression to solids characterized by a flow rule with the plastic deformation rate not being normal to the flow surface. Recently, Benallal and Tvergaard [4] examined the case of bifurcation of a block for a particular strain gradient plasticity theory. All of these analyses have been based on Hill's [5,6] bifurcation theory for quasistatic deformation. Results on inertial influence on neck formation in tension have been reported in [7–10].

BIFURCATION ANALYSIS

In this section, attention is on the possible onset of a spatially *nonuniform* deformation field as a bifurcation from the velocity field (1) at a configuration which represents some fully developed stage of deformation. A useful simplification of the boundary value problem formulation is obtained by choosing the current configuration, whose state is being interrogated, as the reference configuration and by choosing the time scale so that the system is in this configuration at $t = 0$, as has been done in (1). Thus, the velocity field (1) evaluated at $t = 0$ is that of both the current and reference configurations. For this choice of reference configuration, the nominal stress s_{ij} and the Cauchy stress σ_{ij} are identical (but their *rates* are different, in general). All equations can

be satisfied by continuing homogeneous deformation beyond the reference configuration. However, the goal is to seek conditions for which the governing equations can also be met for a continuing deformation which is *not* homogeneous beyond the reference configuration.

Following Hill and Hutchinson [1], the balance of rate of momentum in terms of rate of nominal stress reduces to

$$(\dot{s}_{11} - \dot{s}_{22})_{,12} + \dot{s}_{21,22} - \dot{s}_{12,11} = \rho\,(v_{1,1} - v_{2,2}) \tag{3}$$

upon elimination of the mean pressure. Likewise, the behavior of an initially isotropic and rate–independent material can be represented by the incrementally linear relationships

$$\dot{s}_{11} - \dot{s}_{22} = \left[2\mu^* - \tfrac{1}{2}\sigma\right](v_{1,1} - v_{2,2}) \tag{4}$$

$$\dot{s}_{12} = \left[\mu + \tfrac{1}{2}\sigma\right] v_{2,1} + \left[\mu - \tfrac{1}{2}\sigma + p_I\right] v_{1,2} \tag{5}$$

$$\dot{s}_{21} = \left[\mu - \tfrac{1}{2}\sigma + p_I\right] v_{2,1} + \left[\mu - \tfrac{1}{2}\sigma\right] v_{1,2} \tag{6}$$

where σ is the uniform deviatoric stress acting in the 1–direction and p_I is the hydrostatic pressure given in (2). This material is characterized by two parameters, the instantaneous tangent modulus $4\mu^*$ for ongoing extension and the instantaneous tangent modulus μ for shearing of principal directions. For elastic–plastic materials the parameters are limited by $0 < \mu/\mu^* < \frac{1}{2}$.

The perturbation fields on the uniform background deformation which must satisfy these equations are subject to the boundary conditions

$$v_1 = 0 \quad \text{at} \quad X_1 = \pm\ell_1 \tag{7}$$

$$\dot{s}_{12} = 0 \quad \text{at} \quad X_1 = \pm\ell_1\,, \qquad \dot{s}_{21} = 0 \quad \text{and} \quad \dot{s}_{22} = 0 \quad \text{at} \quad X_2 = \pm\ell_2 \tag{8}$$

The condition (7) is the kinematic condition on the velocity field, and the conditions (8) are the natural boundary conditions.

The equations governing the linear rate problem are equivalent to the variational statement that the functional

$$\Phi[v_i] = \int_{t_1}^{t_2} \int_{A_X} \left[\tfrac{1}{2}\dot{s}_{ij} v_{j,i} - \tfrac{1}{2}\rho \dot{v}_i \dot{v}_i\right] dX_1\, dX_2\, dt \equiv W[v_i] - K[v_i] \tag{9}$$

is stationary under variations of the instantaneous velocity field, where A_X is the area of the plane occupied by the block in its reference configuration and the times t_1, t_2 are arbitrary. It is important to note that, for purposes of the variational statement, the nominal stress rate depends on velocity only through its linear dependence on *current* velocity gradient.

It is obvious that the kinetic energy–like functional $K[v_i]$ is positive for all admissible velocity fields. In analogy with the linear theory of elasticity, the uniqueness of the incrementally linear rate problem is therefore assured if the stress work–like functional $W[v_i] > 0$. The coefficients in the constitutive equations (4)–(6) are different from those in the quasistatic case because of the pressure term p_I. Thus, the failure of the inequality can be expected to occur under different conditions in the dynamic case, in general, than in the quasistatic case. Hutchinson's [11] sufficiency criterion for a quasistatic bifurcation of a plastic solid with a smooth yield surface under multiaxial stress states would be unaffected by the additional hydrostatic pressure due to inertia because only deviatoric stress measures enter this criteria. Thus, a possible dynamic bifurcation from the uniform velocity field (1) differs from the quasistatic bifurcation because of the different constitutive coefficients or because other types of restrictions arise from the high velocity V_0 in the background motion (1).

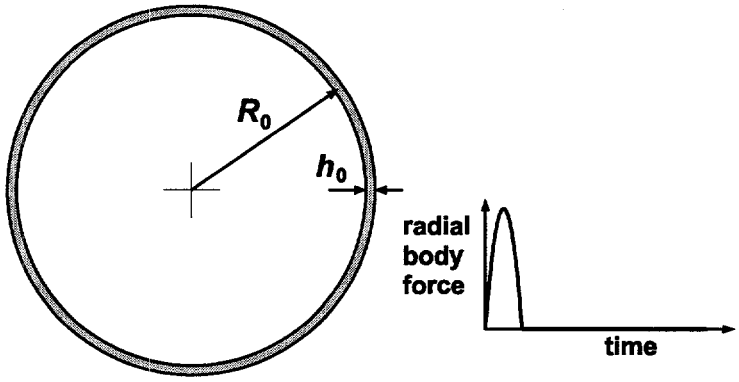


Figure 2: Initial configuration of the ring deformed in plane strain due to action of an impulsive radial body force with time history as indicated.

Obviously, the coefficients in the constitutive equations (4)–(6) reduce to the coefficients for the quasistatic case for rectangular blocks with a large aspect ratio because the effect of lateral inertia as given by (2) vanishes with the square of the aspect ratio of the block. It is noteworthy that the classification into elliptic, parabolic and hyperbolic regimes of the dynamic incremental field equations *without* the hydrostatic pressure due to lateral inertia is identical to the classification reported by Hill and Hutchinson [1] for the quasistatic case. The neck bifurcation modes are thus also relevant for the dynamic extension of a block with a large aspect ratio, as long as these modes are otherwise admitted by the dynamic nature of the problem.

NECK FORMATION IN DYNAMICALLY EXPANDING RINGS

In this section, results of simulation of steel rings expanding radially under plane strain conditions due to dynamic loading are summarized; see Figure 2. For all results reported, the rings have an initial radius of $R_0 = 0.07\,\mathrm{m}$ and an initial thickness-to-radius ratio of $h_0/R_0 = 1/35,\ 1/70$ or $1/140$. The loading is a uniformly distributed radial body force of large amplitude and short duration. The time history of the body force follows a half sine curve as indicated in Figure 2. For all results reported here, the magnitude of the body force variation is $10^8\,\mathrm{N/kg}$ and the duration is $10\,\mu\mathrm{s}$. Momentum imparted to the material by this impulse drives the deformation thereafter due to material inertia. The wall thickness is assumed to have an imperfection in the form of a periodic variation in half–thickness $\Delta h_0(\theta) = \xi \cos 8\theta$ where θ is the angular distance along the ring. The resulting deformation is assumed to be periodic and only a 22.5° segment of the ring is analyzed.

The simulations were based on a dynamic finite element procedure (e.g., see Knoche and Needleman 1993) with an internal variable dilatant viscoplastic material model used to describe the ductile failure by void nucleation, growth and coalescence (cf. Gurson [12], Tvergaard [13]). In the present implementation of the Gurson model, only strain–controlled void nucleation was used. Also, thermal softening by adiabatic heating was taken into account.

The material data used in the simulations correspond to a ductile pressure vessel steel (denoted pvs in the following) which has been characterized by Naus et al. [14]. Some of the relevant data for the pvs steel ring as used here are the strain hardening exponent $N = 0.1$, the strain rate hardening exponent $m = 0.002$, the room temperature yield strength $\sigma_0 = 426\,\mathrm{MPa}$, the material mass density $\rho = 7850\,\mathrm{kg/m^3}$, the initial temperature 20°C, the coefficient of thermal expansion $\beta = 1.1 \times 10^{-5}/°\mathrm{C}$, the fraction of plastic work converted into heat $\chi = 0.9$, the elastic

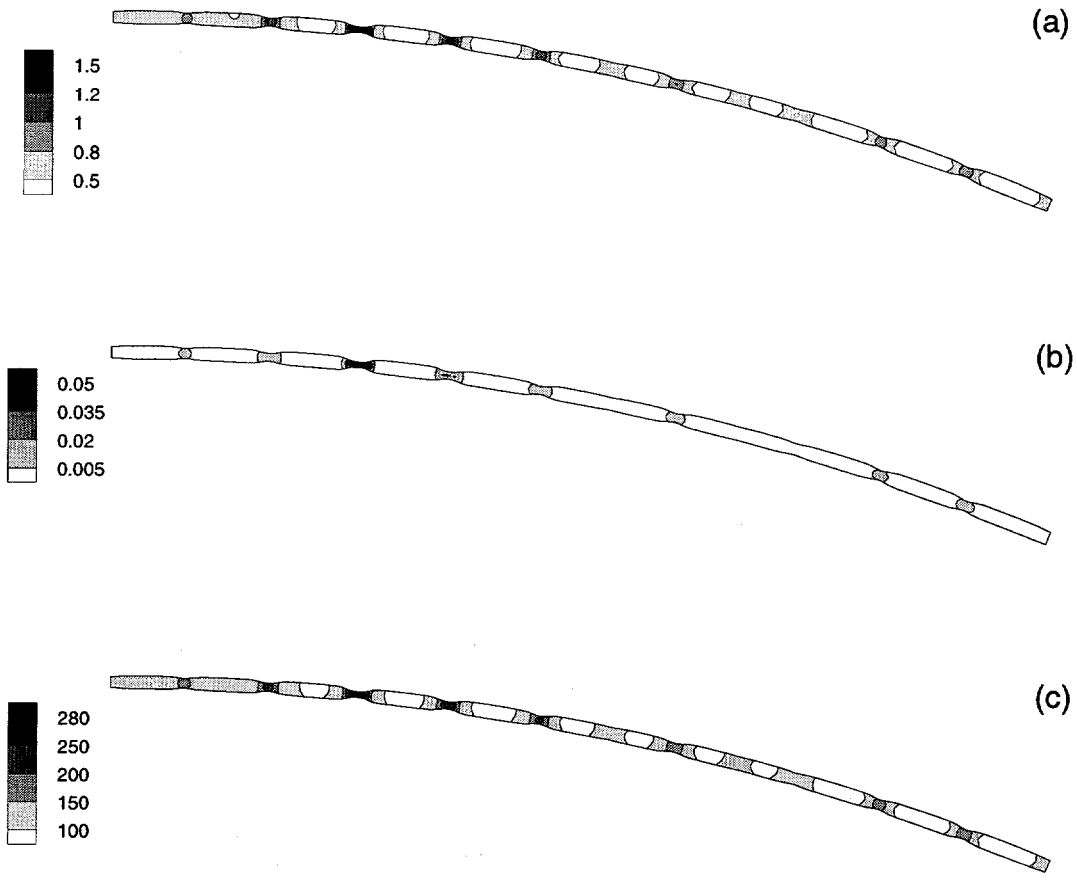


Figure 3: Level curves of (a) effective plastic strain, (b) void volume fraction, and (c) temperature in °C for a pvs–steel ring with $\frac{h_0}{R_0} = \frac{1}{70}$ deformed to $\frac{\Delta R}{R_0} = 0.65$. The ring is initially thinnest at the extreme left end and thickest at the right end in the figure ($\xi = 0.01$).

modulus $E = 206.9\,\text{GPa}$, the Poisson ratio $\nu = 0.3$ and the volume fraction of strain–controlled void nucleating particles $f_v = 0.002$.

A study of dynamic necking in rings has been conducted recently by Han and Tvergaard [8] using a rate–independent plasticity model. The present computations have been carried out along the same lines; for details on boundary conditions the reader is referred to this study. However, the phenomenon of multiple necking observed in [8] for a rate–independent material and found here for a slightly rate–dependent material are very similar, suggesting that strain–rate–hardening effects play a secondary role compared with the imperfection sensitivity of the viscoplastic material in the process of neck formation.

Typically, the average strain rate in these calculations has increased to a value somewhat less than $10^4/\text{s}$ at the time the impact loading on the ring ceases, and it then decays as the rate of expansion of the ring gradually diminishes. For the case of $h_0/R_0 = 1/70$, contour plots of several fields are shown in Figure 3 at a nominal hoop strain of 65%. It is evident that roughly periodic necking has emerged with spacing between necks being much less than the wavelength of the initial imperfection in wall thickness.

In order to visualize the evolution of the situation depicted in Figure 3, an average effective plastic strain rate over a cross–section, normalized by the average strain rate $\dot{R}/R$, provides a useful

instantaneous measure of the deviation from uniform deformation. Results in the form of contours of this normalized strain rate $\dot{e}_s(\theta, t)$ in the θ, t–plane are shown in Figure 4 for three impact simulations representing rings with different thickness–to–radius ratios. Viewed in this way, the darker (lighter) shading shows plastic strain rates which are higher (lower) than the average strain rate. For all cases, the strain rate distribution around the circumference is found to be nearly uniform at the time when the impact loading is completely removed, that is, after 10μs.

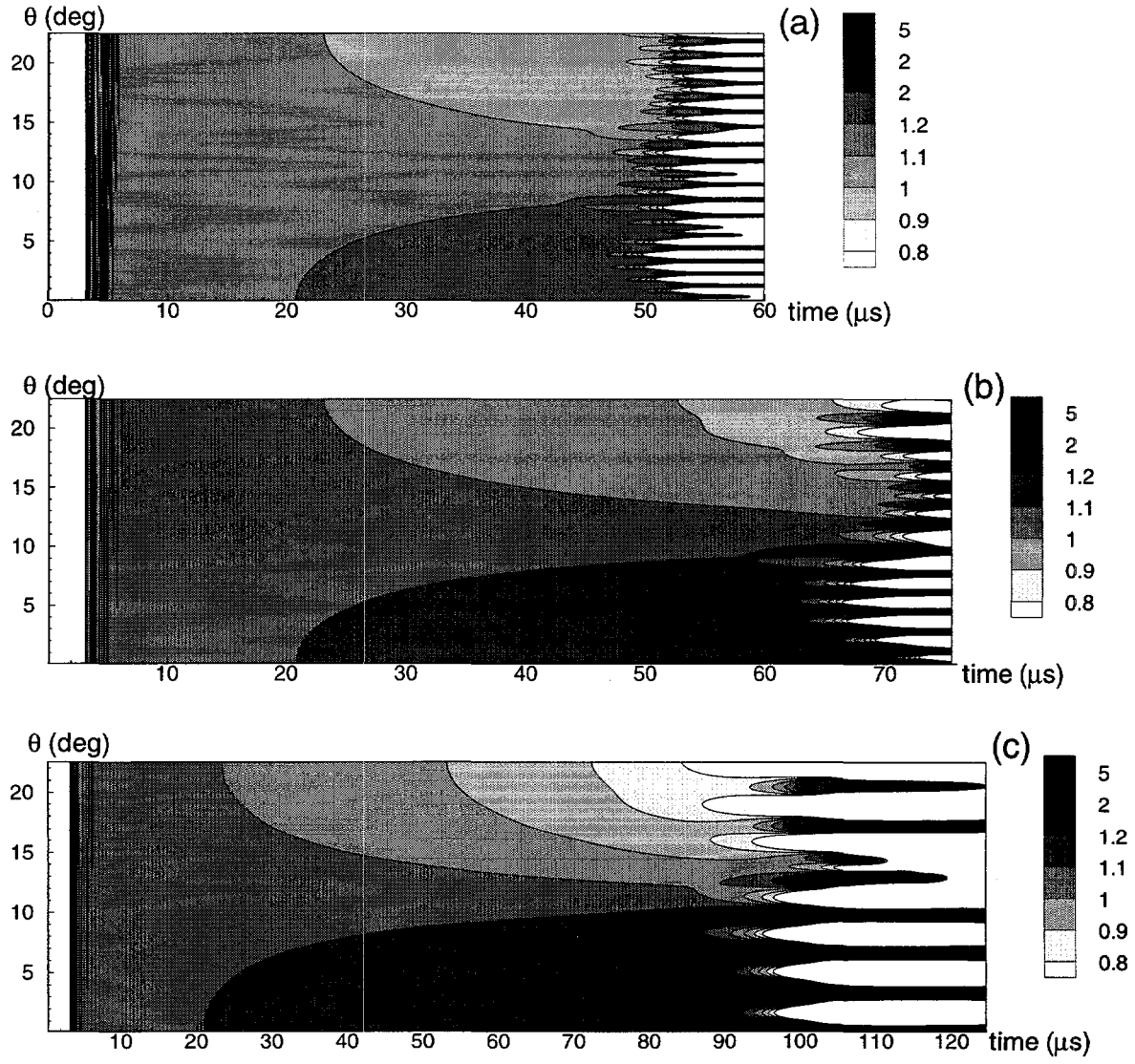


Figure 4: Level curves of normalized cross–sectional effective plastic strain rate $\dot{e}_s$ over the θ, t–plane for rings of the pvs–steel with the initial thickness–to–radius ratio being equal to (a) $\frac{1}{140}$, (b) $\frac{1}{70}$ and (c) $\frac{1}{35}$.

After the applied load is completely removed, inertia continues to drive the deformation which is nearly uniform for the three cases in Figure 4 until the time is approximately two times the impact time. In all cases, a slightly higher strain rate is found at the thinnest point of the ring shortly after $20\,\mu$s and this higher strain rate region grows in size thereafter. Almost simultaneously, a strain rate slightly smaller than average is visible at the thicker part of the ring. Even with higher

resolution in the contour plots than can be represented here, this pattern seems to be common to all cases.

In the context of the bifurcation results for a quasistatic plane strain tension test, this situation can be described qualitatively as slow growth of the long wavelength mode which, in the quasistatic case, would lead to a neck developing symmetrically near 0° and extending to a few degrees to either side of this point. However, in addition to the slowly growing long–wavelength mode, the three cases in Figure 4 show rather different behavior from the quasistatic case in that multiple, closely spaced high/low strain rate zones are seen to develop relatively rapidly for the dynamic cases analyzed here. The order of appearance of these high/low strain rate zones is not the same for the three cases. Furthermore, these short wavelength, highly non–uniform plastic strain rate patterns do not all develop into necks with equal strain intensity. However, this mode appears to be the *critical mode* in the sense that some of these *neck–precursors* do develop into macroscopic necks that are visible in contour plots like those in Figure 3. Some of the neck–precursors corresponding to the critical mode shown Figure 4 die out and the strain levels in the necks in Figure 3 vary accordingly.

SUMMARY

Although the phenomenon of interest has been studied only over a limited range of system parameters, some noteworthy features or trends have emerged. Among these are:

(i) even though the eventual array of necks observed seemed to have little correlation with the original imperfection in wall thickness, it is noteworthy that the necks did not appear at all up to overall strains of unity if no imperfection was present. This suggests that some nonuniformity of the deformation is probably important in precipitating the distribution of necks.

(ii) strain-rate-sensitivity effects seem to play a secondary role compared with the imperfection sensitivity of the viscoplastic material in the process of neck formation.

(iii) the role of the dilatant plasticity model is of minor importance in the process. The constitutive features which essentially control the failure evolution accounted for here are apparently the adiabatic softening and the basic viscoplastic material response.

(iv) for the three cases studied, the spacing of the array of necks varied directly as the initial wall thickness of the ring. At the time of neck formation, the spacing is roughly three to four times the wall thickness.

(v) although most calculations were done for strain hardening parameter $N = 0.1$, a few cases with larger or smaller values were considered. A significant decrease in N results in similar neck formation but at an earlier time, while a signifiant increase in N suppresses formation of necks, particularly in the portion of the wall that is initial thicker.

ACKNOWLEDGMENT

The work leading to the results reported here was supported by the U.S. Department of Energy, Office of Basic Energy Sciences. This support is gratefully acknowledged.

REFERENCES

1. Hill, R. and Hutchinson, J. W. (1975), Bifurcation phenomena in the plane strain tension test, *Proc. Roy. Soc. (London)* **A23**, 239-264.
2. Young, N. J. B. (1976), Bifurcation phenomena in the plane strain compression test, *J. Mech. Phys. Solids* **24**, 77-91.
3. Needleman, A. (1979), Non–normality and bifurcation in plane strain tension and compression, *J. Mech. Phys. Solids* **27**, 231–254.
4. Benallal, A. and Tvergaard, V. (1995), Nonlocal continuum effects on bifurcation in the plane strain tension–compression test, *J. Mech. Phys. Solids* **43**, 741–770.
5. Hill, R. (1958), A general theory of uniqueness and stability in elastic–plastic solids, *J. Mech. Phys. Solids* **6**, 236–249.
6. Hill, R. (1961), Bifurcation and uniqueness in non–linear mechanics of continua, *Problems of Continuum Mechanics* **n**, 155–164.
7. Fressengeas, C. and Molinari, A. (1994), Fragmentation of rapidly stretching sheets, *Europ. J. Mech. A/Solids* **13**,251–268.
8. Han, J. and Tvergaard, V. (1995), Effect of inertia on the necking behavior of ring specimens under rapid radial expansion, *Europ. J. Mech. A/Solids* **14**, 287–307.
9. Needleman, A. (1991), The effect of inertia on neck development, *Topics in Plasticity* **n**, 151–160.
10. Tugcu, P., Neale, K. W. and Lahoud, A. E. (1990), Inertial effects on necking in tension, *Int. J. Solids Struc.* **11**, 1275–1285.
11. Hutchinson, J. W. (1974), Plastic buckling, in *Advances in Applied Mechanics*, edited by C. S. Yih, Academic Press, New York, **14**, 67–144.
12. Gurson, A. L. (1977), Continuum theory of ductile rupture by void nucleation and growth – Part I. Yield criteria and flow rules for porous ductile media, J. Engr. Matl. Tech., **99**, 2–15.
13. Tvergaard, V. (1990), Material failure by void growth to coalescence, in *Advances in Applied Mechanics*, edited by J. W. Hutchinson and T. Wu, Academic Press, New York, **27**, 83–151.
14. Naus, D. J., Keeney–Walker, J., Bass, B. R., Fields, R. J., DeWit, R. and Low III, S. R. (1989), High temperature crack–arrest behavior in 152–mm thick SEN wide plates of quenched and tempered A533 grade B class 1 steel, NUREG/CR–5330 (ORNL/TM–11083), Oak Ridge National Laboratory, National Institute of Standards and Technology.

MODELING OF HIGH HOMOLOGOUS TEMPERATURE DEFORMATION BEHAVIOR FOR STRESS AND LIFE-TIME ANALYSES

Erhard Krempl

Mechanics of Materials Laboratory
Rensselaer Polytechnic Institute
Troy, NY 12180-3590

ABSTRACT

Stress and life-time analyses need realistic and accurate constitutive models for the inelastic deformation behavior of engineering alloys at low and high temperatures. Conventional creep and plasticity models have fundamental difficulties in reproducing high homologous temperature behavior. To improve the modeling capabilities "unified" state variable theories were conceived. They consider all inelastic deformation rate-dependent and do not have separate repositories for creep and plasticity. The viscoplasticity theory based on overstress (VBO), one of the unified theories, is introduced and its properties are delineated. At high homologous temperature where secondary and tertiary creep are observed modeling is primarily accomplished by a static recovery term and a softening isotropic stress. At low temperatures creep is merely a manifestation of rate dependence. The primary creep modeled at low homologous temperature is due to the rate dependence of the flow law. The model is unaltered in the transition from low to high temperature except that the softening of the isotropic stress and the influence of the static recovery term increase with an increase of the temperature.

INTRODUCTION

Design stress analyses of highly loaded components are now performed using finite element programs. For most structural components linear elasticity is appropriate. However, when inelastic material behavior is experienced this model is not adequate. This can happen at low

homologous[1] temperature when the elastic limit is exceeded by an overload event or is allowed to be exceeded to ensure an economic use of the material. At high homologous temperatures a true elastic limit may not exist. Stress levels in the quasi elastic region may lead to significant creep deformation and even to creep rupture. For these and other cases material models that capture the inelastic deformation behavior are needed. Once developed, they need to be implemented in constitutive equation subroutines of finite element programs.

For the inelastic deformation behavior of metals and alloys rate(time)-independent plasticity and combined plasticity and creep are generally used at low and high homologous temperature, respectively. Plasticity and creep models were developed independently and were combined as soon as the finite element programs were ready to accept these nonlinear models and as soon as the technology demanded their use. This was the case in the 1970s when the consumer oriented society started to demand safety and predictability of performance of engineering structures. When the predictions of these models were compared with experimental results major deficiencies were detected in modeling of cyclic loading, of sequences of monotonic loading and of intermittent creep.

"Unified" models were then developed that consider all inelastic deformation as rate-dependent. This approach is in agreement with the notions of materials science where plastic deformation is considered to be a rate process. No separate repository for creep is introduced in the "unified" constitutive equations. Creep and relaxation are manifestations of rate-dependence and of diffusion processes when the boundary conditions in homogeneous states of deformation are zero stress rate and zero total strain rate, respectively.

The conventional plasticity-creep formulation assumes that the low homologous temperature deformation is rate-independent and that rate(time)-dependent deformation commences suddenly when the so-called creep range is reached. Then the model has to be changed from plasticity to combined plasticity and creep.

The separation of creep and plastic deformation models does not have a physical basis. Inelastic deformation is caused by changes in the defect structure such as dislocation motions for monotonic (increasing stress) and creep (constant stress) conditions.

Even if the loading on the boundary of engineering structures is kept constant in time, the constant stress condition generally does not exist inside a structure. The inhomogeneous stress field causes a redistribution of the stresses with time due to the stress distribution and a creep event with constant stress does generally not exist inside the structure.

The combined plasticity-creep formulation can also lead to contradictions. Creep tests performed at stress levels within the "quasi elastic" region of a stress-strain diagram can lead to significant creep deformation and even to creep rupture at high temperature. The idealization of plasticity requires that the yield surface encloses the purely elastic region as defined by the linearity of the stress-strain diagram. So, creep deformation and even creep rupture are taking place inside the yield surface!

"Unified" models were invented to circumvent these conceptual and technological problems.

The purpose of this paper is to introduce a "unified" model, the viscoplasticity theory based on overstress (VBO), and to show some of its qualitative properties for high and low homologous temperature.

---

[1] The dimensionless temperature ratio of test temperature over melting temperature both measured in °K. Engineering structures seldom operate at a homologous temperature greater than 0.6 and operating temperatures below 0.3 are usually considered low homologous temperature.

THE "UNIFIED" STATE VARIABLE THEORIES

General Remarks

"Unified" models including VBO are basically continuum models and assume a representative volume element exists. At this level the contributions of the many possible microstructural mechanisms are only recognized through their aggregate effects. In experiments, the specimen representing a macroscopically homogeneous state of deformation (tensile bar or thin-walled tube) is the representative volume element. It serves as an integrator of all mechanisms. The individual contributions are recognized in " smeared out" form as long as they are influencing the specimen behavior.

In VBO and other state variable theories the modeling of the changing microstructure is accomplished by the so-called state variables and their growth laws. There is no one-to-one correspondence between the state variables and certain micromechanisms. The correspondence is rather diffuse. In VBO the state variables are motivated by experimental results and some qualitative considerations, see Krempl [1]. Generally speaking, state variables for modeling work-hardening in monotonic loading, for cyclic hardening (softening) and for the Bauschinger effect are needed.

Inelastic deformation in metals and alloys is primarily affected by dislocation motion and by other changes in the defect structure. In most of the cases the dislocation density increases with inelastic deformation and further movement of dislocations is impeded by their increasing density. As a consequence work hardening is observed macroscopically. In some cases, when cold-worked metals are subjected to cyclic inelastic loading for example, cyclic softening occurs indicative of an easing in the passage of dislocations and a decrease in dislocation density.

At low homologous temperatures diffusion is negligible and the defect structure acquired during inelastic deformation is stable in the absence of mechanical loading. As the temperature increases diffusion processes become important. Defects can now change by "thermal action." Generally, diffusion tends to counteract the hardening effects of inelastic deformation. Hardening due to inelastic deformation and softening due to diffusion occur while external mechanical loads are applied. In the absence of external loads hardening essentially ceases and the effects of diffusion continue until equilibrium is attained[2]. The defect structure is observed to change with time, "static" recovery is said to occur. Subsequent loading shows that the hold periods with zero external load can lead to a softening of the response[3].

Hardening due to inelastic deformation and softening due to diffusion (static recovery) act simultaneously. Depending on the loading and on the temperature, hardening or softening may be pronounced. At low homologous temperatures hardening is dominant. Hardening and softening can also be in equilibrium as happens when the stress is constant and secondary creep can be observed. As the temperature increases diffusion effects become increasingly important and hardening ceases when the melting temperature is approached.

---

[2] Although there are no external mechanical loads acting on the material, residual internal stresses with zero resultant exist and they influence the change in the microstructure. These internal stresses are self equilibrating and do not enter into a continuum formulation.

[3] It is assumed that the material behavior is "normal" and that effects like strain aging and age hardening do not occur.

At low temperatures where diffusion is negligible the growth laws for the state variables are homogeneous of degree one in the rates so that neither time nor the rate of loading has an influence on hardening. The modeling of the diffusion processes in a continuum approach mostly follows the Orowan-Bailey format, via a static recovery term in the growth laws of the state variables. This term is introduced to counteract the hardening and it is not homogeneous of degree one in the rates. With it rate and time of loading have an influence on the growth of the state variables. When the material is unloaded to zero load and is left at zero load, the recovery terms return the state variables to zero with time. Majors and Krempl [2] have, however, shown that the static recovery term alone cannot model cyclic softening and tertiary creep.

Viscoplasticity Theory Based on Overstress (VBO)

Assuming isotropy, small strains, variable temperature and volume preserving inelastic deformation the deviatoric flow law can be written as

$$\dot{\boldsymbol{e}} = \dot{\boldsymbol{e}}^{el} + \dot{\boldsymbol{e}}^{in} = \frac{d}{dt}\left(\frac{1+\nu}{E}\boldsymbol{s}\right) + \frac{3}{2}F[\Gamma]\frac{\boldsymbol{s}-\boldsymbol{g}}{\Gamma} \qquad (1)$$

where a bold face quantity denotes tensor and a superposed dot or $\frac{d}{dt}$ designates material time derivative; $\boldsymbol{e}$ and $\boldsymbol{s}$ denote the strain and stress deviators, respectively; $\boldsymbol{g}$ is the equilibrium stress deviator, $\boldsymbol{s}$ - $\boldsymbol{g}$ is the overstress (effective stress) deviator and $\Gamma = \sqrt{\frac{3}{2}tr\left[(\boldsymbol{s}-\boldsymbol{g})(\boldsymbol{s}-\boldsymbol{g})\right]}$ is the overstress invariant or effective overstress. The elastic modulus is E and ν is the elastic Poisson's ratio. The function F has the dimension of 1/time and is positive, increasing with $F[0]=0$. Experimental evidence for introducing the overstress dependence of the inelastic strain rate is discussed by Krempl [1].

When all rates are zero then $\boldsymbol{s} = \boldsymbol{g}$ is a solution of Eq. (1); $\boldsymbol{g}$ therefore represents the stress that can be sustained at rest or in equilibrium and this property has given $\boldsymbol{g}$ its name.

From Eq.(1) the effective inelastic strain rate can be obtained as $\sqrt{\frac{2}{3}tr\left(\dot{\boldsymbol{e}}^{in}\dot{\boldsymbol{e}}^{in}\right)} = \dot{p} = F[\Gamma]$.

The deviatoric relations have to be augmented with a statement that the response to a hydrostatic state of stress is elastic

$$tr\dot{\boldsymbol{\epsilon}} = \frac{d}{dt}\left(\frac{1-2\nu}{E}tr\boldsymbol{\sigma}\right) + 3\alpha\dot{T} \qquad (2)$$

where $\boldsymbol{\sigma}$ and $\boldsymbol{\epsilon}$ are the stress and strain tensors, respectively. The coefficient of thermal expansion is α and T -T_0 is the change from the reference temperature T_0. The writing of the first terms on the right hand side of Eqs. (1) and (2) ensure path independence of elastic deformation when the elastic constants are dependent on temperature, see Krempl [3].

The equilibrium stress enables the modeling of nonlinear rate sensitivity and the separation of viscous, work-hardening and rate-independent contributions to the flow stress, see Krempl [3]. Its growth law with a static recovery term is

$$\dot{\boldsymbol{g}} = \frac{\psi[\Gamma]}{E}\dot{\boldsymbol{s}} + \boldsymbol{s}\frac{\partial}{\partial T}\left(\frac{\psi[\Gamma]}{E}\right) + F[\Gamma]\left(\frac{\boldsymbol{s}-\boldsymbol{g}}{\Gamma} - \frac{\boldsymbol{g}-\boldsymbol{f}}{A}\right)\psi[\Gamma] + \left(1-\frac{\psi[\Gamma]}{E}\right)\dot{\boldsymbol{f}} - R[\bar{g}]\boldsymbol{g} \qquad (3)$$

The positive decreasing function $\psi[\Gamma]$ with the dimension of stress controls the transition from the initial elastic behavior to the fully established inelastic flow. It is bounded by $E > \psi[\Gamma] > E_t$

where E_t is the tangent modulus in the inelastic range based on total strain rate. The second term on the right hand side, together with corresponding expressions in the flow law, ensures path independent elastic behavior.

The deviatoric quantity $\boldsymbol{f}$ is another state variable called the kinematic stress. It is the repository for modeling the Bauschinger effect. In addition work-hardening, work-softening or flow at constant stress in monotonic loading is reproduced depending on whether the slope based on inelastic strain rate $\hat{E}_t$ is positive, negative or zero. The slopes in the flow stress region are related by $\hat{E}_t = E_t/(1-E_t/E)$ and are almost equal for most engineering alloys. The growth law for the kinematic stress is

$$\dot{\boldsymbol{f}} = \hat{E}_t F[\Gamma]\frac{\boldsymbol{s}-\boldsymbol{g}}{\Gamma} \tag{4}$$

It is also possible to make $\hat{E}_t$ a function if there is a variable slope in the flow stress region.

The positive, increasing function $R[\bar{g}]$ with the dimension of reciprocal time constitutes the static recovery term. The equilibrium stress invariant is $\bar{g} = \sqrt{\frac{3}{2}tr(\boldsymbol{gg})}$ and $R[0] = 0$. The static recovery term reduces the growth of $\boldsymbol{g}$ and it is zero when the equilibrium stress is zero.

The scalar state variable A is called the isotropic stress. It is primarily the repository for modeling of cyclic hardening, cyclic softening and tertiary creep, see Majors and Krempl [2]. An appropriate growth law has to be specified.

At low homologous temperature A represents the rate-independent contribution to the flow stress, see Fig. 1 of Krempl [3]. As the temperature increases diffusion becomes important and the contribution to the flow stress represented by A depends on the duration of the test. It and with it the equilibrium stress will become rate-dependent.

To account for rate-independent cyclic hardening/softening at low homologous temperature the growth law must be formulated to be homogeneous of degree one in the rates. In this formulation time has no effect. Examples of modeling complex situations such as extra hardening in out-of-phase loading, see Krempl and Choi [4] and Choi and Krempl [5].

As the temperature increases a rate-dependent growth law with softening is appropriate to model tertiary creep, see Majors and Krempl [2], Tachibana and Krempl [6] and Maciucescu et al.[7]. The growth law used by Maciucescu et al. [7] is

$$\dot{A} = -\beta\left(\frac{A-A_2}{A_2}\right)p \tag{5}$$

where β is a constant with dimension stress/time and A_2 is the final value of A. If the initial condition is $A_1 > A_2$, softening will be modeled. This is appropriate for high homologous temperature. Hardening would be represented if $A_1 < A_2$. The quantity p is the inelastic strain path length defined above.

Modeling of Deformation Behavior

The above equations represent the three dimensional version of VBO. To ascertain its modeling capabilities the material constants E, E_t, ν, and α and the functions

$F[\Gamma]$, $\psi[\Gamma]$, and $R[\bar{g}]$[4] must be determined from appropriate tests. Included are monotonic and cyclic loading in displacement (strain) or load (stress) control, creep (stress is constant) and relaxation (displacement is constant). Such tests are usually performed in the uniaxial state of stress or, less frequently, in biaxial stress states. To start the determinations the three dimensional equations must be specialized for the appropriate stress state and the test conditions. Then values for the constants are assumed and the coupled, nonlinear, stiff differential equations are integrated numerically. The results are compared with the experiments. New constants are assumed and integration yields new results. This process is continued until a satisfactory match of simulation and experimental data is achieved. Once the constants have been determined the constitutive equations should be specialized for another set of tests which had not been used in the determination of the constants. Then the model is to be specialized for these tests to affect a prediction. If the comparison with experiment is satisfactory, the model can be used with confidence in stress analyses.

Reduction to the Uniaxial State of Stress

Using the definition of a deviator for the strain rate Eqs. (1) and (2) can be written as

$$\dot{\boldsymbol{\epsilon}} = \frac{d}{dt}\left(\frac{1+\nu}{E}\boldsymbol{s} + \frac{1-2\nu}{3E} tr\boldsymbol{\sigma}\boldsymbol{I}\right) + \alpha\dot{T}\boldsymbol{I} + \frac{3}{2}F[\Gamma]\frac{\boldsymbol{s}-\boldsymbol{g}}{\Gamma} \tag{6}$$

where $\boldsymbol{I}$ is the identity matrix. For the uniaxial state of stress the stress deviator is given by

$$[\boldsymbol{s}] = \frac{\sigma}{3}\begin{bmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \tag{7}$$

where σ is the uniaxial true stress. Assuming the same relation for the equilibrium stress with G denoting the counterpart of σ, the uniaxial component of the strain rate $\dot{\varepsilon}$ is found to be

$$\dot{\varepsilon} = \frac{d}{dt}\left(\frac{\sigma}{E}\right) + \alpha\dot{T} + F[\Gamma]\frac{\sigma - G}{\Gamma} \tag{8}$$

It is easy to convert the growth laws for the state variables to the uniaxial state of stress. They are

$$\dot{g} = \frac{\psi[\Gamma]}{E}\dot{\sigma} + \sigma\frac{\partial}{\partial T}\left(\frac{\psi[\Gamma]}{E}\right) + F[\Gamma]\left(\frac{\sigma - G}{\Gamma} - \frac{G - \bar{f}}{A}\right)\psi[\Gamma] + \left(1 - \frac{\psi[\Gamma]}{E}\right)\bar{f} - R[|G|]G \tag{9}$$

and

$$\dot{\bar{f}} = \hat{E}_t F[\Gamma]\frac{\sigma - G}{\Gamma} \tag{10}$$

where $\bar{f}$ is the counterpart of σ and G and where $\Gamma = |\sigma - G|$. The growth law for A does not change at all but p is now the accumulated inelastic strain.

Eqs. (5) and (8) through (10) are the uniaxial constitutive equations which must be specialized for the test conditions before they can be used.

With modern servo-controlled testing machines load or displacement can be imposed on the specimen. In most of the tests the rates are constant over certain time intervals and can change instantaneously. For example loading up to a creep stress level can be done with a certain stress or

[4] They in turn contain other constants. It is possible to use constants instead of functions thus reducing the number of constants needed, see Maciucescu et al. [7].

strain rate with a change to zero stress rate when the creep test starts. Either engineering or true quantities can be controlled.

It is useful to recall the relations between the true and the engineering quantities (designated by the same symbol but with a ^). With constant density assumed these relations are $\varepsilon = \ln(1+\hat{\varepsilon})$ and $\sigma = \hat{\sigma}(1+\hat{\varepsilon})$. To simulate the described tests the conditions listed in the last column of the table below must be substituted into the set of constitutive equations with an appropriate value for the constant rates.

Type of Control	Condition	Expression to be substituted
Strain control	$\dot{\hat{\varepsilon}}$ = constant	$\dot{\varepsilon} = \dot{\hat{\varepsilon}}/(1+\hat{\varepsilon}) = \dot{\hat{\varepsilon}}/\exp\varepsilon$
True strain control	$\dot{\varepsilon}$ = constant	$\dot{\varepsilon}$
Stress control	$\dot{\hat{\sigma}}$ = constant	$\dot{\sigma} = \dot{\hat{\sigma}}(1+\hat{\varepsilon}) + \hat{\sigma}\dot{\hat{\varepsilon}} = \dot{\hat{\sigma}}\exp\varepsilon + \sigma\dot{\varepsilon}$
True stress control	$\dot{\sigma}$ = constant	$\dot{\sigma}$

In addition the temperature history must be specified,

Note that for relaxation both the engineering and the true strain rates are zero. There is a difference between the " constant load" creep test (engineering stress rate is zero) and the "constant stress" creep test (true stress rate is zero.). It is known that the experimentally observed constant load and constant true stress creep curves can be different. Therefore, a possibility exists to model these differences with VBO.

Other test conditions such as a cyclic test require the specification of the stress or strain and temperature histories.

With these expressions the coupled nonlinear non-autonomous differential equations must be integrated numerically to simulate the specific tests. For a good model the curves obtained from the numerical test should be identical to the experimental one.

Without specification of the test condition the constitutive equations cannot be solved. The response of the model will depend on the temporal evolution of the input conditions.

Isothermal Creep with Constant True Stress

Loading up to the flow stress region is performed with a strain rate $\dot{\varepsilon}^-$. Then the true stress is kept constant. Applying the uniaxial equations without the recovery term immediately before and after the switch to constant stress enables the calculation of the initial creep rate and of the equilibrium stress rate. They are

$$\dot{\varepsilon}^+ = \dot{\varepsilon}^-(1 - E_t/E) \tag{11}$$

and

$$\dot{G}^+ = \dot{\varepsilon}^- E_t(1 - \psi/E) \tag{12}$$

where $^+$ designates the rate at the beginning of the creep test. The equilibrium stress rate before the start of the creep test is $\dot{G}^- = \dot{\varepsilon}^- E_t$ where we have made use of the long-term asymptotic solutions of VBO, see Krempl [3]. It is seen from Eqs. (11) and (12)that the rates at the beginning of the creep test are reduced from the values just before the test. It is also seen that the equilibrium stress continues to increase. Although it is believed that similar relations hold in the presence of the static recovery term it cannot be proven mathematically.

beginning of the creep test are reduced from the values just before the test. It is also seen that the equilibrium stress continues to increase. Although it is believed that similar relations hold in the presence of the static recovery term it cannot be proven mathematically.

Usually creep is classified as primary (creep rate decreases in magnitude), secondary (creep rate magnitude is constant) and tertiary (creep rate magnitude increases)[5]. To ascertain the properties of the model it is best to differentiate the flow law Eq. (8) which results in

$$\ddot{\varepsilon}^{in} = F'\left[\left|\sigma_o - G\right|\right]\left(-\dot{G}\right) = \ddot{\varepsilon}^{creep} \tag{13}$$

The prime designates total derivative and the subscript $_o$ indicates that the stress is constant.

It can be seen that the sign of the second derivative is determined by the sign of the equilibrium stress rate. For primary creep $\dot{G} > 0$ and $\dot{G} < 0$ is required for tertiary creep. The equilibrium stress rate has to be zero for the modeling of secondary creep.

The purpose of the static recovery term and the softening of the isotropic stress is to bring about the changes in the slope of the equilibrium stress to represent secondary and/or tertiary creep at long times.

By substitution of the inelastic strain rates in the growth law for the equilibrium stress it can be shown that all the terms in the growth law for the equilibrium stress, except the static recovery term, are homogeneous of degree one in the rates, see Krempl [3]. Rate and elapsed time have no influence on their contribution to the growth of G. The influence of the static recovery term which reduces the growth of the equilibrium stress rate increases with test duration. Also the decreasing isotropic stress A reduces the equilibrium stress rate. Both are needed to model secondary and tertiary creep. Examples of modeling are given in Tachibana and Krempl [5] and Maciucescu et al. [7].

At low temperatures diffusion has no influence and the static recovery term is set to zero. It is not possible to reproduce tertiary creep as long as the model is required to reproduce work-hardening or inelastic flow at constant stress in a tensile test, see Krempl [3][6]. Only primary or at the most secondary creep can be modeled at low homologous temperature usually referred to as "Cold Creep." "Cold creep", relaxation and loading rate sensitivity are simply manifestations of rate-dependence under different boundary conditions. This type of behavior was found in engineering alloys at room temperature, see Krempl [8] and Kujawski and Krempl [9].

CONCLUSION

The qualitative properties of VBO in modeling creep behavior were described. The transition from low homologous temperature behavior to that at high homologous temperature is affected by the static recovery term and the softening of the isotropic stress which is a scalar state variable of VBO. They become increasingly important as temperature and duration increase. "Cold" as well as high temperature creep including tertiary creep can be modeled with one constitutive model. It is not necessary to switch to a different set of constitutive equations as temperature increases using this "unified" approach.

[5] We refer to magnitude to make the description valid for tension and compression.

[6] Tertiary creep can be modeled only if there is work-softening in monotonic loading.

ACKNOWLEDGEMENT

The author acknowledges the support of the US Department of Energy, Grant DE-FG02-96ER14603.

REFERENCES

1. Krempl, E., "The overstress dependence of the inelastic rate of deformation inferred from transient tests," Materials Science Research International, **1**, 3-10 (1995).
2. Majors, P. S. and E. Krempl, "The isotropic viscoplasticity theory based on overstress applied to the modeling of modified 9Cr-1Mo steel at 538 °C," Materials Science and Engineering, **A186**, 23-34 (1994).
3. Krempl, E., "A small strain viscoplasticity theory based on overstress," in *Unified Constitutive Laws of Plastic Deformation,* A. S. Krausz and K. Krausz, editors, Academic Press, 281-318 (1996).
4. Krempl, E. and S. H. Choi, "Viscoplasticity theory based on overstress: The modeling of ratchetting and cyclic hardening of AISI Type 304 Stainless Steel," Nuclear Eng. and Design, **113**, 401-410 (1992).
5. Choi, S. H. and E. Krempl, "Viscoplasticity theory based on overstress: The modeling of biaxial cyclic hardening using irreversible plastic strain," in *Advances in Multiaxial Fatigue, ASTM STP 1191,* D. L. McDowell and R. Ellis, editors, American Society for Testing and Materials, Philadelphia, PA, 259-272 (1993).
6. Tachibana, Y. and E. Krempl, "Modeling of creep and tensile behavior at high homologous temperature using the viscoplasticity theory based on overstress (VBO): Part I-Creep and tensile behavior," Journal of Engineering Materials and Technology, **117**, 456-461 (1995).
7. Maciucescu, L., T.-L. Sham and E. Krempl, "VBO: A state variable constitutive equation for a solder alloy," to be published in the proceedings of INTERpack'97 to be held June 15-19, 1997, Island of Hawaii.
8. Krempl, E., "An experimental study of room-temperature rate-sensitivity, creep and relaxation of AISI Type 304 Stainless steel," J. Mech. Phys. Solids, **27**, 363-385 (1979).
9. Kujawski, D. and E. Krempl, "The rate(time)-dependent behavior of Ti-7Al-2Cb-1Ta Titanium alloy at room temperature under quasi-static monotonic and cyclic loading," Journal of Applied Mechanics, **103**, 55-63 (1981).

ORIGINS OF ASYMMETRIC STRESS-STRAIN RESPONSE IN PHASE TRANSFORMATIONS

Huseyin Sehitoglu and Ken Gall

Department of Mechanical and Industrial Engineering
University of Illinois, Urbana, IL 61801, USA

ABSTRACT

It has been determined that the transformation stress-strain behavior of CuZnAl and NiTi shape memory alloys is dependent on the applied stress state. The uniaxial compressive stress necessary to macroscopically trigger the transformation is approximately 34% (CuZnAl) and 26% (NiTi) larger than the required uniaxial tensile stress. For three dimensional stress states, the response of either alloy system is dependent on the directions of the dominant principal stresses along with the hydrostatic stress component of the stress state. The stress state effects are dominated by the favored growth and nucleation of more martensite plates in tension versus compression. The effect of different hydrostatic pressure levels between stress states on martensite plates volume change is considered small.

INTRODUCTION

The purpose of this work is to determine the physical origins of the tension-compression asymmetry and the hydrostatic stress effect with novel experiments and measurements in two technologically important materials (CuZnAl and NiTi). Using unique equipment, considerable sensitivity to hydrostatic stress state has been obtained experimentally for the first time. It should be noted that there are currently no studies available in which both effective and hydrostatic stresses were systematically changed. Since shape memory alloys (SMA) can store large amounts of recoverable pseudo-elastic energy, they could be used in many applications where large strains are essential, but permanent deformation and energy loss due to plastic dissipation is undesirable. Additionally, SMA's have an advantage over traditional materials since the large pseudo-elastic mechanical strains can be triggered thermally, electrically, or mechanically.

SMA's owe their unique stress-strain behavior to a reversible thermoelastic martensitic transformation. It is widely accepted that the stress-induced martensitic transformation produces two unique macroscopic stress-strain responses, pseudoelasticity and the shape memory effect [1, 2] Analogous to stress-strain curves in the plastic regime, pseudoelastic and shape memory stress-strain curves demonstrate macroscopic yield points, hardening regions, and mechanical

hysteresis upon unloading. The primary difference between the three curves concerns the mechanism of recoverable strain. Plastically deformed materials recover strains upon reverse loading, pseudoelastic materials recover strains immediately upon unloading, and materials exhibiting the shape memory effect recover strains after being subsequently heated. The existence of one phenomenon over another in any given alloy system is a function of test temperature, material composition, processing technique, and heat treatment.

Despite the wide ranging applicability of SMA's, there is a limited amount of experimental work on the response of SMA's to stress states other than tension [1-3]. This gap in research efforts is intriguing since the dependence of stress-induced martensitic transformations on the applied stress state was observed some 40 years ago in steels [4]. With this in mind, the purpose of the current study is to expose the issues related to the transformation behavior of CuZnAl and NiTi shape memory alloys under different stress states. More precisely, this work will focus on the dependence of the critical transformation stress level on the applied stress state.

EXPERIMENTAL TECHNIQUES

Polycrystalline $Cu_{59.1}Zn_{27.0}Al_{13.8}$ and $Ni_{50.0}Ti_{50.0}$ weight percent alloys were employed for the study. The normal to the habit plane and the twinning direction have the direction cosines (.199, .6804, .705) and (.1817,-.7457,.6411) respectively. CuZnAl demonstrates a small negative change (-0.3%) in volume upon transformation from the parent phase to the martensitic phase. The habit plane normal and transformation direction are given as (-.8889,.404,.215) and (.435,.7543,.4874) respectively for NiTi. This results in a small positive volume change (.19%). The heat treatment in both cases consisted of a solution heat treatment followed by an aging treatment. The treatment was performed to keep the martensite start temperature, M_s, at a reasonable level below room temperature. This assures that the transformation will be stress-induced. On average, the NiTi M_s was about -18 °C, while the CuZnAl M_s was about -10 °C. The tests in this study were conducted at room temperature where the sample is fully austenitic ($T > A_f$). Details of the unique experimental equipment used for the triaxial tests can be found in a recent publication [3] and are also summarized below.

In our work, a servohydraulic test machine fitted with a unique high pressure vessel is used for triaxial testing of NiTi and CuZnAl specimens. The schematic of the test system is provided in Figure 1. As Figure 1 indicates, axial stresses are applied to the specimen by the servohydraulic actuator of the MTS test machine; diametral stresses are applied to the specimen through the introduction of pressurized fluid into the pressure vessel. The axial stress is changed by applying force in the longitudinal direction, and circumferential and radial stresses are related to applied pressure ($=-p$). The ability of the present triaxial testing apparatus to simultaneously ramp the lateral and axial stresses on the specimen represents one of its main advantages over previous triaxial research efforts. In previous works, hydrostatic compression was typically applied first and the uniaxial stress was increased in a secondary operation. The present scheme circumvents any arguments regarding the role of initial hydrostatic compression on the material behavior. A personal computer was used for all test definition, command generation, and data acquisition tasks. More details of the pressure intensifier, load and strain measurements can be found in a recent publication [3].

EXPERIMENTAL RESULTS

The effective stress-strain curves for the CuZnAl and NiTi are shown in Figures 2 and 3 respectively. Stress states #1 and #3 are simple uniaxial tension and compression. Stress state

#2 has the following combination of principal stresses: $\sigma_1 = 2p$, $\sigma_2 = -p$, $\sigma_3 = -p$, while stress state #5 is governed by: $\sigma_1 = -2p$, $\sigma_2 = -p$, $\sigma_3 = -p$. Several other triaxial stress states were studied, and the results are discussed more thoroughly in two recent publications [13, 14]. The stress-strain curves are only shown up to 3% strain because strains much larger that this introduce considerable plastic deformation and non-recoverability [13]. In general, the NiTi has a much higher transformation yield point while the CuZnAl demonstrates a larger post-yield hardening modulus. Although the difference in hardening behavior is not as drastic as Figures 1 and 2 might indicate (Figure 2 has a scale twice as large as Figure 1) the difference is still notable. Both CuZnAl and NiTi show transformation yield points which are much higher in compression than in tension. In addition, the yield point of the zero hydrostatic case lies close to the yield point in pure tension for both materials. The yield point of CuZnAl under triaxial compression lies in between the tension and compression yield points. However, in NiTi, the yield point of the triaxial compression test lies considerably above both the tensile and compressive yield points.

DISCUSSION

For the most part, the stress state effects in CuZnAl and NiTi can be directly linked to micro-mechanical phenomenon. When a particular stress state is imposed on a SMA specimen, transformation strains are accumulated through the nucleation and subsequent growth of several preferred martensite plates (variants) [5]. Figures 4a and 4b show the typical arrangement and number of martensite plates caused by an applied stress in CuZnAl. The first image (a) is a magnified view of the plates in a single grain while the image (b) show the formation of different plates in several grains. The purpose of the two images is simply to demonstrate that two variants usually control the stress-induced transformation and that different grains favor the formation of selected variants. Unfortunately, it is not trivial to compare Figures 2-4 and completely understand the stress-state effects. To link the experimental behavior to the microscopic observations, a micro-mechanical model must be incorporated [1,6].

Although the model will not be extensively discussed here, the predictions of the model are a cornerstone in the understanding of stress-state effects in these alloys. As in a real material the model has the possibility of forming 24 martensite variants per grain. However, consistent with experimental observations, the model predicts that only 2 or 3 of these variants actually control the transformation under an applied stress state [1]. The advantage of the model is that it allows the "observation" of microscopic variables controlling the transformation which are not easily observed experimentally. One of the key predictions of the model is that more variants will activate under an applied tensile stress versus a compressive stress (Figures 5(a) and 5(b)). Clearly, the favored activation of martensite variants between stress states is one cause of stress state effects in these alloys. Simply stated, if a particular stress state has dominant principal stresses in tension, more variants will activate, the transformation will microscopically proceed quicker, and the macroscopic transformation yield stress will be lower.

Balancing the effect of the number of transforming variants is the relationship between the volume change during transformation and the hydrostatic component of the applied stress state. Thus, if the applied stress state has a negative hydrostatic component then the transformation will be triggered at a lower effective stress for CuZnAl. Macroscopically, the difference in the transformation yield point caused by differences in the hydrostatic stress component between stress states is not visible unless the hydrostatic stress difference is substantial. For example, the hydrostatic stress component due to pure compression is slightly more negative than the hydrostatic stress component due to pure tension. However, from Figures 2 and 3 it is clear that transformation in tension is favored over transformation in compression. One would have expected compression to have a lower yield point since its

hydrostatic stress component is compressing in the direction which the transformation wants to proceed. Through Figure 5(a), the model demonstrates that the transformation is indeed "microscopically triggered" at a lower effective stress in compression in the CuZnAl case (note the very small offset in the number of transforming variants curve). However, this small offset is quickly overshadowed when more variants begin contributing to the transformation. In the case of NiTi the habit plane normal and the transformation direction lead to a positive volume change, consequently, the effect of volume change and transforming variants are additive leading to a higher sensitivity of the results to hydrostatic stress (Figure 5(b)).

Although the origin of stress state effects is clearly related to the microscopic aspects of the transformation, there still exists some experimental phenomenon that are not completely accounted for. The current theory of the authors is that texture is playing an intense role in the 3-D transformation behavior. At any rate, research is now in progress to experimentally view microscopic martensite growth *in situ* to better understand the dependence of martensite growth on the stress state.

CONCLUSIONS

(1) The uniaxial compressive stress necessary to macroscopically trigger the transformation is approximately 34% (CuZnAl) and 26% (NiTi) larger than the required uniaxial tensile stress. For three dimensional stress states, the response of either alloy system is dependent on the directions of the dominant principal stresses along with the hydrostatic stress component of the stress state.

(2) Stress state effects in CuZnAl and NiTi alloys are a balance between the number of transforming variants and the hydrostatic pressure (volume change) effect. The variant effects are more pronounced when two stress states have a small difference in hydrostatic stress components and the principal stresses are in different directions. The hydrostatic pressure effects become evident when there are extremely large differences in hydrostatic pressures between stress states.

ACKNOWLEDGMENTS

The research is supported by Department of Energy (DOE DEFG02-93ER14393), Basic Energy Sciences Division, Germantown, Maryland.

REFERENCES

1. K. Gall, H. Sehitoglu, H. J. Maier, and K. Jacobus, *Submitted to Acta. Met.*, (1997).
2. K. Jacobus, H. Sehitoglu and M. Balzer, *Met. Trans.*, **27A**, 3066, (1996).
3. M. Balzer and H. Sehitoglu, *Experimental Mechanics*, **37-1**, (1997).
4. S.A. Kulin, M. Cohen and B.L. Averbach, *J. Metals.* , **4**, 661, (1952).
5. T.A. Schroeder and C.M. Wayman, "Pseudoelastic Effects in Cu-Zn Single Crystals" *Acta Met.* **27**, 405, (1979).
6. E. Patoor, A. Eberhardt and M. Berveiller, Mechanics of Phase Transformations and Shape Memory Alloys, eds. L.C. Brinson and B. Moran, ASME, New York, NY, 23, (1994).

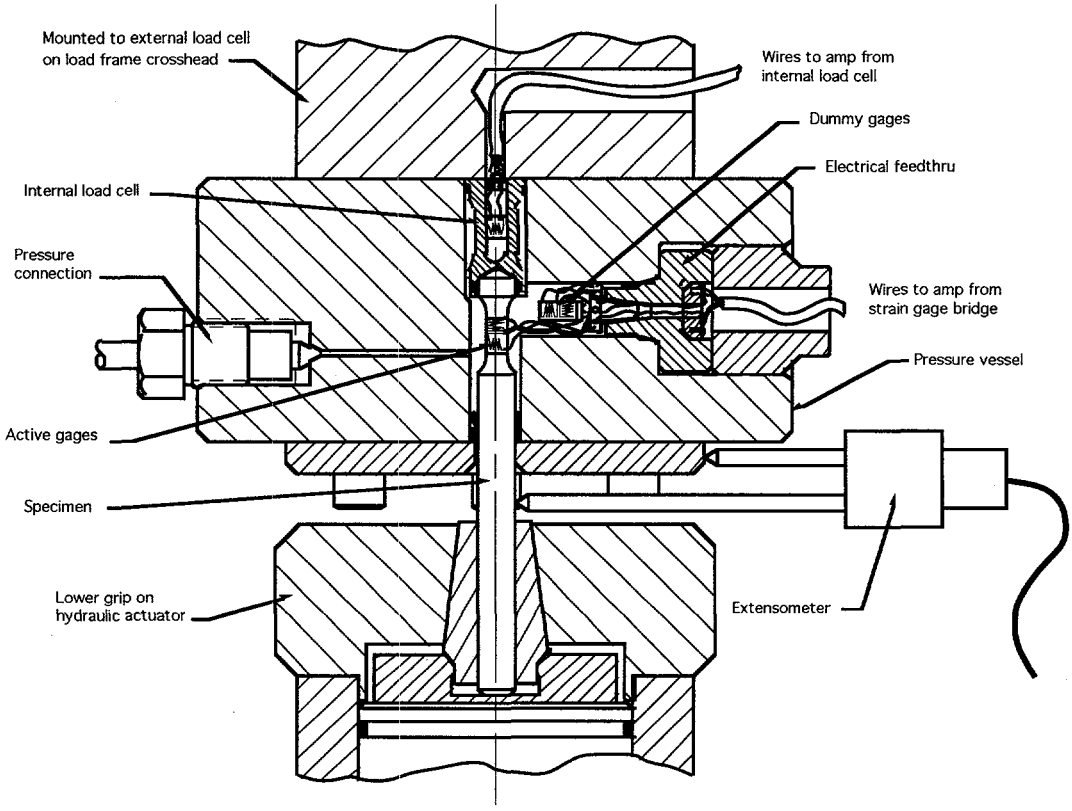


Figure 1. Schematic of the pressure vessel for studying mechanical behavior at high pressures [3].

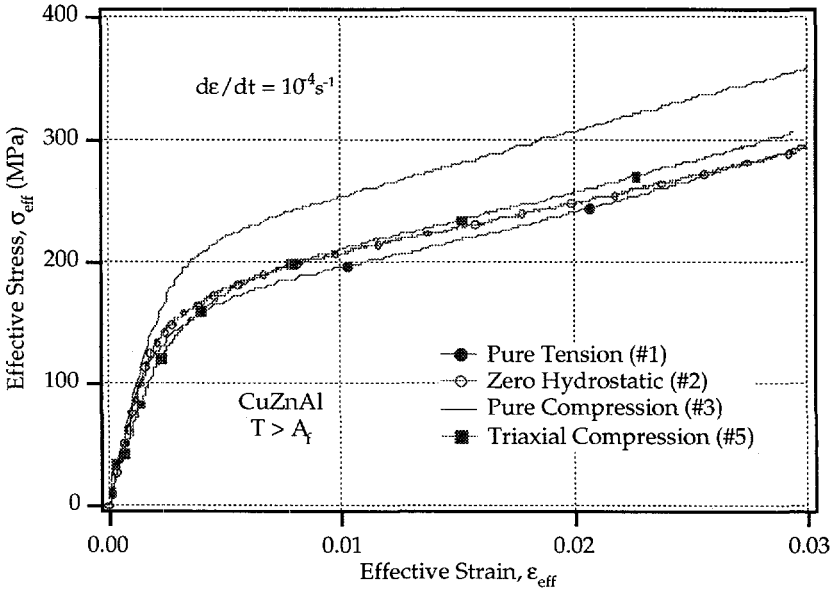


Figure 2. Effective stress-strain plots for a polycrystalline CuZnAl shape memory alloy specimen above the austenite finish temperature, A_f [1].

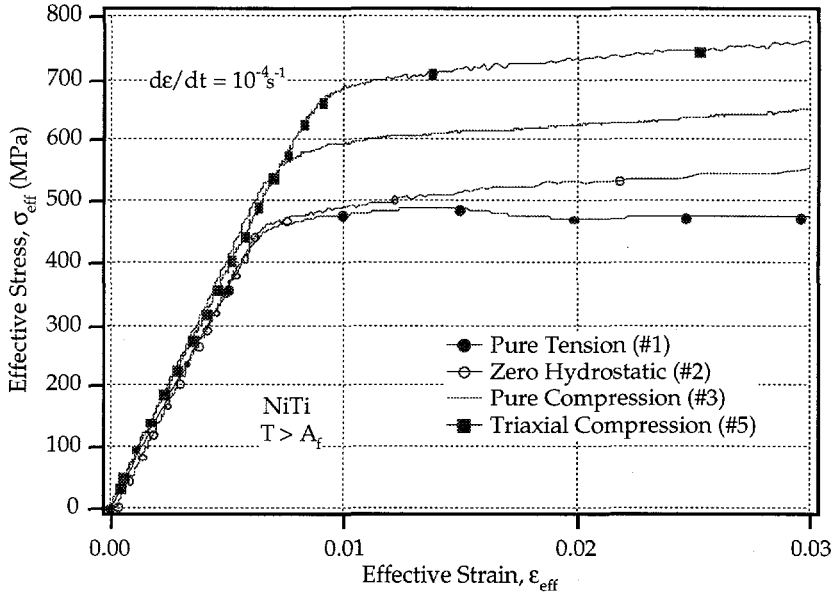


Figure 3: Effective stress-strain plots for a polycrystalline NiTi shape memory alloy specimen above the austenite finish temperature, A_f [2].

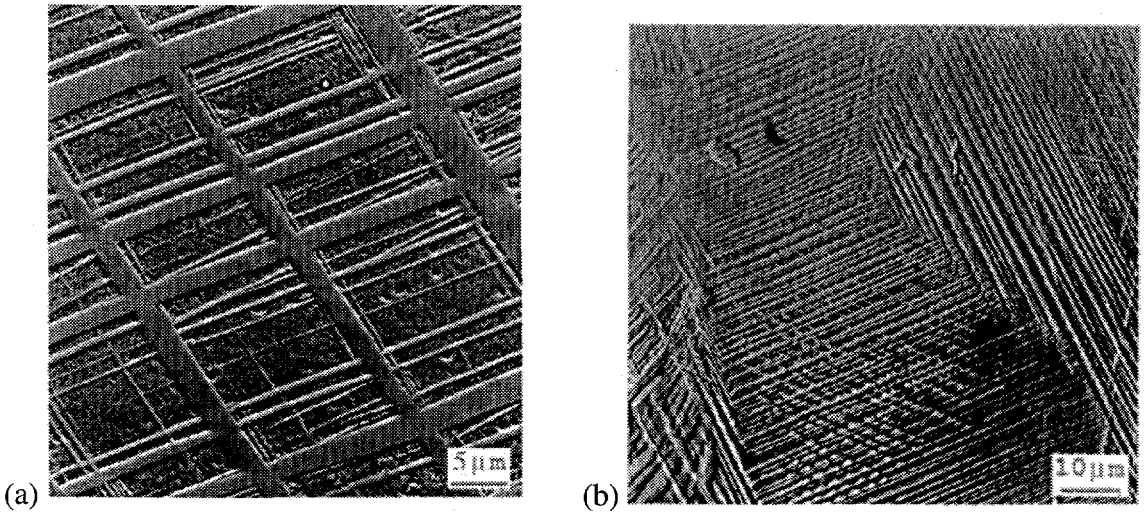
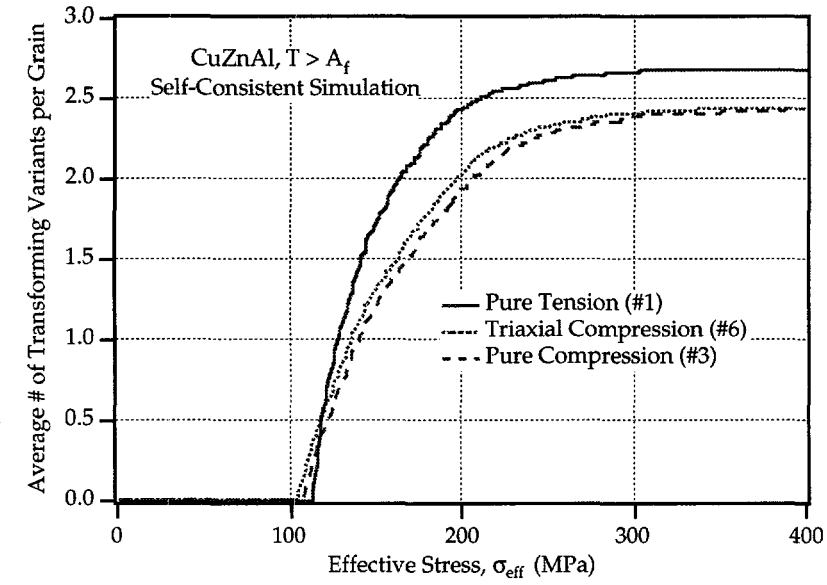
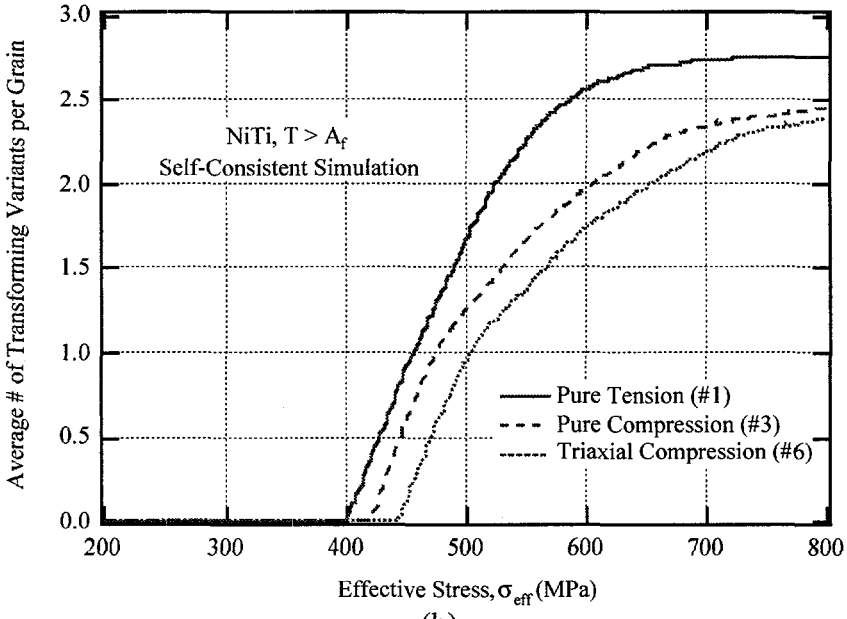


Figure 4. Scanning electron microscope image of martensite plates (a) in a single grain and (b) in several grains of CuZnAl [1].



(a)



(b)

Figure 5. Plot of the average number of active martensite variants versus the applied effective stress for a polycrystalline CuZnAl SMA [1], (b) results for NiTi .

CUTTING STATE IDENTIFICATION

B. S. Berger, I. Minis, M. Rokni, M. Papadopoulos, K. Deng, A. Chavalli

University of Maryland, Mechanical Engineering Department
College Park, MD 20742-3035 U.S.A.

Argonne National Laboratory
Argonne, Illinois 60439, U.S.A.

A Summary Reprinted from *Journal of Sound and Vibration* Vol. 200(1), B.S. Berger, et al., CUTTING STATE IDENTIFICATION, pp. 15-29, 1997, by permission of the publisher Academic Press Limited, London

ABSTRACT

Cutting states associated with the orthogonal cutting of stiff cylinders are identified through an analysis of the singular values of a Toeplitz matrix of third order cumulants of acceleration measurements. The ratio of the two pairs of largest singular values is shown to differentiate between light cutting, medium cutting, pre-chatter and chatter states. Sequences of cutting experiments were performed in which either depth of cut or turning frequency was varied. Two sequences of experiments with variable turning frequency and five with variable depth of cut, 42 cutting experiments in all, provided a database for the calculation of third order cumulants. Ratios of singular values of cumulant matrices find application in the analysis of control of orthogonal cutting

INTRODUCTION

The identification of cutting states, associated with the orthogonal cutting of stiff cylinders, is realized in the following through an analysis of the behavior of the singular values of a Toeplitz matrix of third order cumulants of acceleration measurements. A bispectral analysis of cutting tool acceleration measurements has shown, [3], that the cutting process is quadratically phase coupled. The determination of coefficients in an autoregressive approximation of the bispectrum, [20], involves the construction of an unsymmetric Toeplitz matrix, **R**, of third order cumulants. It is shown that the behavior of the dominant pairs of singular values of **R** provides a basis for the identification of cutting states. In particular, the ratio of the two pairs of largest singular values, the

R-ratio, is shown to differentiate between light cutting, medium cutting, pre-chatter and chatter states. Sequences of cutting experiments were performed in which either depth of cut or turning frequency was varied while all other cutting parameters were held constant. Two sequences of experiments with variable turning frequency and five with variable depth of cut, a total of forty-two cutting experiments, were studied. Results typical of the entire set are presented for a sequence of variable cutting depth and a sequence of variable turning frequency. The R-ratio evaluated at maxlag = 100, (4), is close to one for all cases of light cutting and two or greater for chatter. For intermediate states the ratio increases as the chatter state is approached.

EXPERIMENTAL APPARATUS

A schematic diagram of the experimental apparatus employed is shown in Figure 1 and consists of a Hardinge CNC lathe, a special force dynamometer (utilizing three Kistler 9068 force transducers) and its associated electronics, and a digital spectrum analyzer (Hewlett Packard 3566A) for data acquisition and real-time analysis.

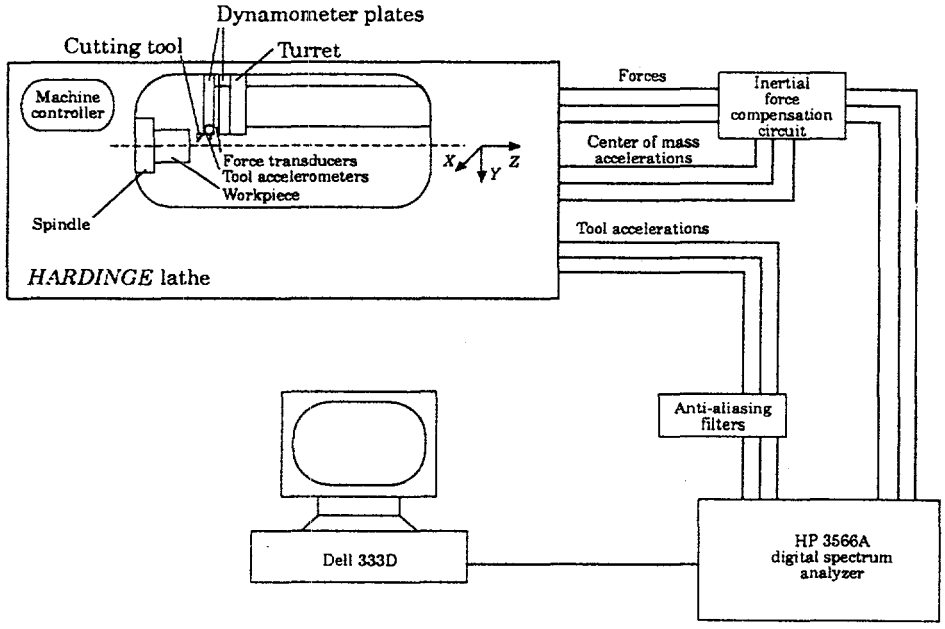


Figure 1. The experimental system.

THIRD ORDER RECURSION

Let $c_3(\tau_1, \tau_2) \equiv$ the third order cumulant of the real third order stationary random process X(k), k = 0, ±1, ±2. ... . If the mean of X(k) vanishes then $c_3(\tau_1, \tau_2) = m_3(\tau_1, \tau_2)$ where $m_3(\tau_1, \tau_2)$

$= E(X(k), X(k+\tau_1)\, X(k+\tau_2))$, E is the expected value, which may be estimated by

$$m_3(\tau_1,\tau_2) = (1/2n) \sum_{k=-n}^{+n} X(k)\; X(k+\tau_1)\; X(k+\tau_2) \tag{1}$$

where $n \to +\infty$. The bispectrum of X(k), $C_3(\omega_1, \omega_2)$ is defined by

$$C_3(\omega_1,\omega_2) = \sum_{\tau_1=-\infty}^{+\infty} \sum_{\tau_2=-\infty}^{+\infty} c_3(\tau_1,\tau_2)\; \exp\; [-j(\omega_1\tau_1+\omega_1\tau_2)] \tag{2}$$

$|C_3(\omega_1, \omega_2)| \equiv$ the bispectral index.

Consider an autoregressive, AR, estimation of the bispectrum, $C_3(\omega_1, \omega_2)$, (2) [16,17]. A p - th order AR process is described by

$$X(k) + \sum_{i=1}^{p} a(i)\; X(k-i) = W(k) \tag{3}$$

where it is assumed that W(k) is non-Gaussian, $E(W(k)) = 0$, $E(W^3(k)) = \beta$. Multiplying through (3), summing and noting (1) gives

$$c_3^x(-k,-l) + \sum_{i=1}^{p} a(i)\; c_3^x(i-k,\; i-l) = \beta\; \delta(k,l) \tag{4}$$

where k, l > 0. Letting k=l in the third order recursion equation, (4), with k = 0, ..., p yields p+1 equations for the p+1 unknowns a(i) and β; p+1 ≡ maxlag. In matrix notation

$$\boldsymbol{R}\; \boldsymbol{a} = \boldsymbol{b} \tag{5}$$

where

$$\boldsymbol{R} = \begin{vmatrix} g(o,o) & g(1,1) & \ldots g(p,p) \\ g(-1,-1) & g(o,o) & \ldots g(p-1,p-1) \\ \vdots & & \vdots \\ g(-p,-p) & g(-p+1,-p+1) & \ldots g(o,o) \end{vmatrix} \tag{6}$$

$g(i,j) \equiv c_3^x(i,j)$, $\mathbf{a} \equiv [1, a(1), ..., a(p)]^T$ and $\mathbf{b} \equiv [\beta, 0, ..., 0]^T$. **R** is in general a nonsymmetric Toeplitz matrix. A sufficient but not necessary condition for the representation in (5) to exist is the symmetry and positive definiteness of **R**. A discussion of this and related conditions is given in [17]. The bispectrum corresponding to (3) is given by, [4],

$$C_3^x(\omega_1,\omega_2) = \beta\ H(\omega_1)\ H(\omega_2)\ H^*(\omega_1+\omega_2) \tag{7}$$

where

$$H(\omega) = 1/(\ 1\ +\ \sum_{n=1}^{p}\ a(i)\ \exp\ (-j\ \omega\ n)) \tag{8}$$

and $H^*(\omega) \equiv$ complex conjugate of $H(\omega)$.

An estimate of the **R** matrix, (6), and bispectrum, (7), for a data set X(I), I=1,...,N may be formed [16,17], as follows:

1. Segment the data set into K records of M samples each. $X^i(k)$, k=1,2,...,M are data points associated with the i-th record.
2. Compute $c_3{}^x{}_{,i}$ (m,n) for the i-th record as

$$c_{3,i}^{x} = (1/M) \sum_{l=a}^{b}\ X^{(i)}(l)\ X^{(i)}(l+m)\ X^{(i)}(l+n) \tag{9}$$

where i = 1,2,...,K, a ≡ max (1,1-m,1-n) and b≡ min(M, M-m, M-n).

3. Average $c_3{}^x{}_{,i}$ (m,n) over all K records,

$$\hat{c}_3(m,n) = (1/K) \sum_{i=1}^{K}\ c_{3,i}^{x}(m,n) \tag{10}$$

to yield the estimate $\hat{c}_3(m,n)$ of $c_3(m,n)$. Form an estimated **R** matrix by replacing $c_3(m,n)$ by $\hat{c}_3(m,n)$ in (6). Estimated values of **a** follow from (5). These results implemented in [22] are subsequently applied to orthogonal cutting data.

SINGULAR VALUE DECOMPOSITION

If **A** is a real mxn matrix then there exist orthogonal matrices $\mathbf{U} \in R^{mxm}$ and $\mathbf{V} \in \mathbf{R}^{nxn}$ such that

$$\boldsymbol{U}^{T}\ \boldsymbol{A}\ \boldsymbol{V} = diag.(\sigma_1,\cdots,\sigma_q) \in R^{mxn} \tag{11}$$

where q = min(m,n), $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_q \geq 0$ are the singular values and R^{mxn} denotes a real mxn matrix. A criterion for selecting the autoregressive order, p, in (3) is given in [17,22]. p is chosen to equal the number of singular values of the R matrix which are above the noise floor. Note that if $\sigma_1 \geq \ldots > \sigma_r > \sigma_{r+1} = \ldots = \sigma_q = 0$ then rank (**A**) = r, [5,7].

Relationships between phase coupled trigonometric functions and the singular values of the corresponding **R** matrix were established through a study of three functions $f_i(t)$ where

$$f_1(t) = \cos(2\pi \cdot 100t+\phi_1) + \cos(2\pi \cdot 100t+\phi_2) + 0.2\cos(2\pi \cdot 200t+\phi_1+\phi_2) \quad \textbf{(12)}$$

$$f_2(t) = 0.9\cos(2\pi \cdot 90t+\phi_1) + 1.0\cos(2\pi \cdot 100t+\phi_2) + 0.2\cos(2\pi \cdot 190t+\phi_1+\phi_2) \quad \textbf{(13)}$$

$$f_3(t) = 1.0\cos(2\pi 90t+\phi_1) + 1.0\cos(2\pi 100t+\phi_2) + 1.0\cos(2\pi 190t+\phi_1+\phi_2) + 1.0\cos(2\pi 100t+\phi_2) + 1.0\cos(2\pi 110t+\phi_3) + 0.5\cos(2\pi 210t+\phi_2+\phi_3) \quad \textbf{(14)}$$

and ϕ_i are mutually independent and uniformly distributed over $[0, 2\pi]$. The $f_i(t)$ functions were sampled at 1024 Hz over an interval of 10 sec. **R** matrices were evaluated for each $f_i(t)$ by averaging over 10 one sec. intervals, (6), (9).

$f_1(t)$, (12) is an example of the self phase coupling of a 100 Hz frequency component. In the experimental data studied frequency components in the neighborhood of 100 and 200 Hz were always observed in the power spectra of cutting states close to chatter. A peak with frequency coordinates in the neighborhood of (100 Hz, 100 Hz), appeared in the bispectrum of cutting states in the neighborhood of chatter. The ratio of the mean of the largest pair of singular values to the mean of the second largest pair defines a non-dimensional ratio of invariants of **R**, the R-ratio. This ratio is shown as a function of maxlag for $f_1(t)$ in Figure 2(c). R-ratio $\simeq$ 2.0 for maxlag > 30.

$f_2(t)$, (13), involves the phase coupling of 90 and 100 Hz components. A bispectral peak at (100, 90) indicates phase coupling between the 90 and 100 Hz components. The mean of the first pair of singular values is nearly equal to the mean of the second pair of singular values for maxlag > 80. Note Figure 3(b) for maxlag > 90 for which 1.0 < R-ratio $\leq$ 1.2.

$f_3(t)$, (14), is the sum of a phase coupled component at 100 Hz and 110 Hz and a phase coupling of 90 and 100 Hz components. The bispectrum of $f_3(t)$ has peaks at (100, 110), (100, 90) and (110, 100), (90, 100) because of symmetry. In Figure 4(b) the R-ratio is plotted as a function of maxlag from which it is seen that R-ratio $\simeq$ 1.5 for maxlag > 80.

CUTTING STATE CHARACTERIZATION

Sequences of cutting experiments were performed in which either depth of cut or the turning frequency was varied with all other cutting parameters held constant. Singular values of **R**, (6), were computed for two sequences with variable turning frequency and five sequences with variable depth

of cut over a turning frequency range of 290-852 rpm. Each variable cutting depth sequence ended in chatter while each variable turning frequency sequence contained at least one chatter state. A total of 42 cutting experiments were performed. Typical sequences have been selected from the set.

For sequence 1, s-1, the turning frequency = 460 rpm, rake angle = 5°, surface speed = 90 m/min, feed rate = .007 in/rev, resampling rate = 1024 Hz, frequency cut-off = 1100 Hz and the depth of cut = 2.5, 2.6, 2.7 and 2.8 mm at which depth chatter was observed.

The R-ratio vs. maxlag is shown in Figure 5(b) for a depth of cut of 2.5 mm which corresponds to light cutting. The R-ratio, Figure 5(b), is close to 1.0. For $70 \leq$ maxlag ≤ 100, $1.12 \geq R \geq 1.08$. The behavior of the R-ratio as a function of maxlag has similarities with that of $f_2(t)$, (13), Figure 3(b). $f_2(t)$ contains two phase coupled trigonometric functions of 90 and 100 Hz which approximates phase coupling between the first natural frequency of the system at 98 Hz and a lower frequency component of the sideband structure.

Chatter was observed for a depth of cut of 2.8 mm. The R-ratio vs. maxlag is shown in Figure 6(b). One pair of singular values is dominant. For $20 \leq$ maxlag ≤ 100, $2.0 \leq$ R-ratio ≤ 2.4. The R-ratio as a function of maxlag is similar to that of $f_1(t)$, (12), Figures 2(b) and 6(b), which represents self phase coupling of the 100 Hz component. The presence of self-phase coupling in the time series is confirmed by peaks in the power spectrum at 100 and 200 Hz and a peak in the bicoherence index at (100, 100). These results are consistent across all sequences of experiments in which the depth of cut varies. For increasing depth of cut the R-ratio clearly differentiates between light cutting and chatter.

Two intermediate states with depths of cut of 2.6 and 2.7 mm complete the sequence s-1. The R-ratio vs. maxlag is shown in Figure 7(b), for the 2.6 mm case. For $50 \leq$ maxlag ≤ 110 the R-ratio ≈ 1.6, Figure 7(b). There is a similarity between Figure 4(b), R-ratio for $f_3(t)$ and Figure 7(b). The R-ratios are close to one another for $50 \leq$ maxlag ≤ 110.

CONCLUSIONS

The ratio of the mean of the two dominant pairs of singular values, R-ratio, evaluated for maxlag = 100, approximates one for light cutting, two or more for chatter and near chatter states and takes intermediate values for intermediate states, increasing from one to two as chatter is approached. This behavior was observed in an analysis of tool acceleration time series for five sequences of cutting experiments with increasing depth of cut and two sequences with variable turning frequency. For chatter and light and intermediate cutting the R-ratio is seen to be a constant or slowly changing function of maxlag for maxlag > 40 (4), Figures 5(b), 6(b), and 7(b).

ACKNOWLEDGEMENTS

The authors acknowledge the support of D.O.E. through DE-FG02-93ER14335 and the N.S.F. through GER-9354956. The encouragement of O. Manley, D.O.E. and P. Grootenhuis and D.J. Ewins, Imperial College of Science and Technology, London, is appreciated.

REFERENCES

3. B. S. Berger, I. Minis, K. Deng, Y. S. Chen, A. Chavali and M. Rokni 1996 *Journal of Sound and Vibration* 191(5), 986-992. Phase Coupling in orthogonal cutting.
4. D. R. Brillinger and M. Rosenblatt 1967a in Spectral Analysis of Time Series (B. Harris, editor). New York : John Wiley. Asymptotic theory of k-th order spectra. 1967b in Spectral Analysis of Time Series (B. Harris, editor). New York : John Wiley. Computation and interpretation of k-th order spectra.
5. G. H. Golub and C. F. Van Loan 1993 Matrix Computations. Baltimore : The Johns Hopkins University Press.
7. R. A. Horn and C. R. Johnson 1991 Topics in Matrix Analysis. Cambridge : Cambridge University Press.
16. C. L. Nikias and J. M. Mendel 1993 *IEEE Signal Processing Magazine* July, 10-37. Signal processing with higher-order spectra.
17. C. L. Nikias and A. P. Petropulu 1993 Higher-Order Spectra Analysis. Englewood Cliffs, New Jersey : Prentice Hall.
20. M. R. Raghuveer and C. L. Nikias 1985 *IEEE Transactions on Acoustics, Speech and Signal Processing* ASSP-33(4), 1213-1230. Bispectrum estimation : A parametric approach.
22. A. Swami, J. M. Mendel and C. L. Nikias 1993 Hi Spec Toolbox. Matick Massachusetts : Math Works, Inc.

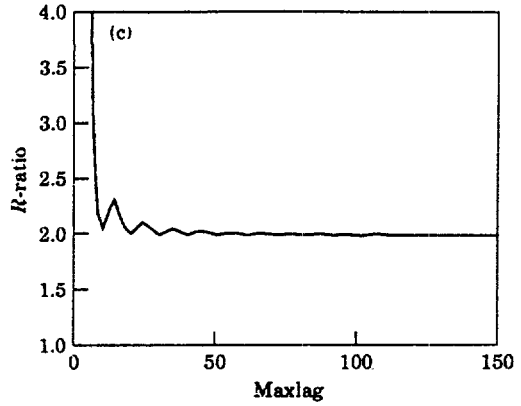


Figure 2. Test Function $f_1(t)$: (c) R-ratio vs. maxlag for $f_1(t)$.

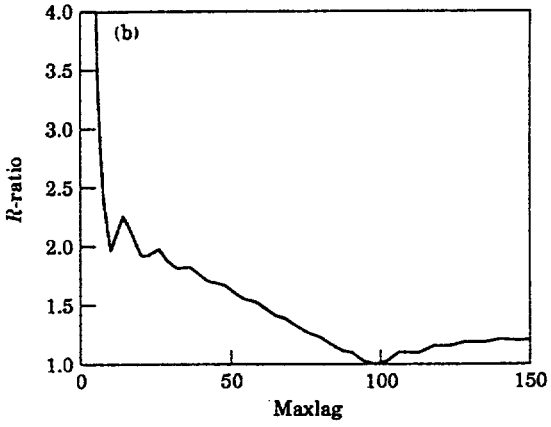


Figure 3. Test Function $f_2(t)$. (b) R-ratio vs. maxlag for $f_2(t)$.

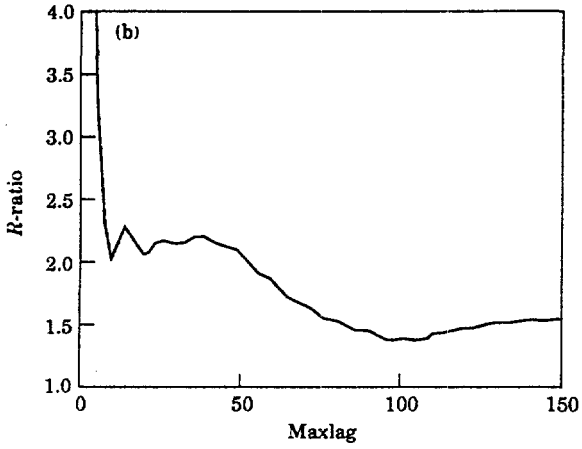


Figure 4. Test Function $f_3(t)$. (b) R-ratio vs. maxlag for $f_3(t)$.

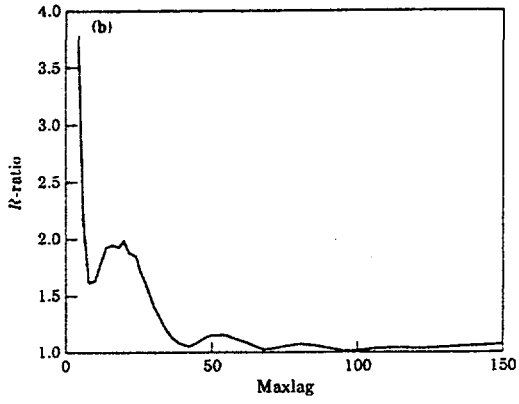


Figure 5. Data Set s-1. 2.5 mm. (b) R-ratio vs. maxlag.

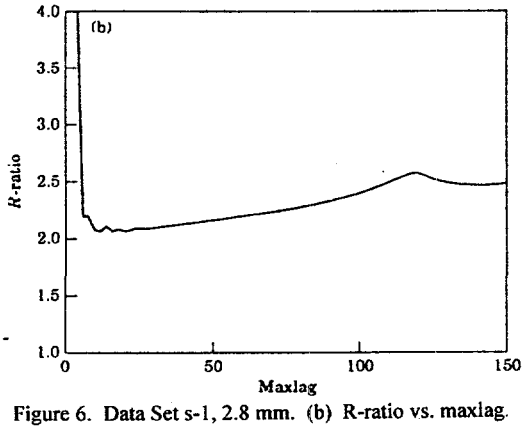


Figure 6. Data Set s-1, 2.8 mm. (b) R-ratio vs. maxlag.

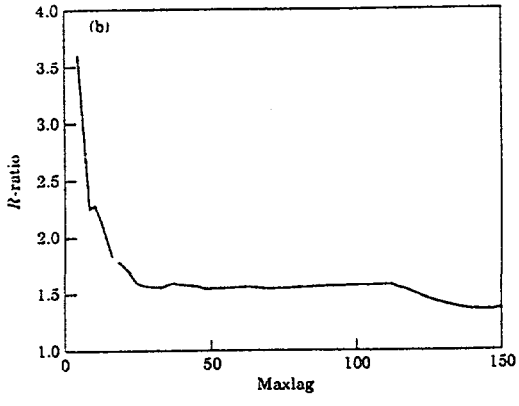


Figure 7. Data Set s-1, 2.6 mm. (b) R-ratio vs. maxlag.

MULTI-ROBOT MOTION CONTROL FOR COOPERATIVE OBSERVATION

Lynne E. Parker

Oak Ridge National Laboratory
Oak Ridge, TN 37831-6364

ABSTRACT

An important issue that arises in the automation of many security, surveillance, and reconnaissance tasks is that of monitoring (or observing) the movements of targets navigating in a bounded area of interest. A key research issue in these problems is that of sensor placement — determining where sensors should be located to maintain the targets in view. In complex applications involving limited-range sensors, the use of multiple sensors dynamically moving over time is required. In this paper, we investigate the use of a cooperative team of autonomous sensor-based robots for the observation of multiple moving targets. We focus primarily on developing the distributed control strategies that allow the robot team to attempt to minimize the total time in which targets escape observation by some robot team member in the area of interest. This paper first formalizes the problem and discusses related work. We then present a distributed approximate approach to solving this problem that combines low-level multi-robot control with higher-level reasoning control based on the ALLIANCE formalism. We analyze the effectiveness of our approach by comparing it to three other feasible algorithms for cooperative control, showing the superiority of our approach for a large class of problems.

INTRODUCTION

An important issue that arises in the automation of many security, surveillance, and reconnaissance tasks is that of monitoring (or observing) the movements of targets navigating in a bounded area of interest. A key research issue in these problems is that of sensor placement — determining where sensors should be located to maintain the targets in view. In the simplest version of this problem, the number of sensors and sensor placement can be fixed in advance to ensure adequate sensory coverage of the area of interest. However, in more complex applications, a number of factors may prevent fixed sensory placement in advance. For example, there may be little prior information on the location of the area to be monitored, the area may be sufficiently large that economics prohibit the placement of a large number of sensors, the available sensor range may be limited, or the area may not be physically accessible in advance of the mission. In the general case, the combined coverage capabilities of the available robot sensors will be insufficient to cover the entire terrain of interest. Thus, the above constraints force the use of multiple sensors dynamically moving over time.

In this paper, we investigate the use of a cooperative team of autonomous sensor-based robots for applications in this domain. We focus primarily on developing the distributed control strategies that allow the team to attempt to minimize the total time in which targets escape observation by some robot team member in the area of interest. Of course, many variations of this dynamic, distributed sensory coverage problem are possible. For example, the relative numbers and speeds of the robots and the targets to be tracked can vary, the availability of inter-robot communication can vary, the robots can differ in their sensing and movement capabilities, the terrain may be either

enclosed or have entrances that allow targets to enter and exit the area of interest, the terrain may be either indoor (and thus largely planar) or outdoor (and thus 3D), and so forth. Many other subproblems must also be addressed, including the physical tracking of targets (e.g. using vision, sonar, IR, or laser range), prediction of target movements, multi-sensor fusion, and so forth. Thus, while our ultimate goal is to develop distributed algorithms that address all of these problem variations, we first focus on the aspects of distributed control in homogeneous robot teams with equivalent sensing and movement capabilities working in an uncluttered, bounded area.

The following section defines the multitarget observation problem of interest in this paper, and is followed by a discussion of related work. We then describe our approach, discussing each of the subcomponents of the system. Next, we describe and analyze the results of our approach, compared to three other feasible algorithms for cooperative motion control. Finally, we offer concluding remarks.

PROBLEM DESCRIPTION

The problem of interest in this paper — the cooperative multi-robot observation of multiple moving targets (or *CMOMMT* for short) — is defined as follows. Given:

$\mathcal{S}$: a two-dimensional, bounded, enclosed spatial region, with entrances/exits
$\mathcal{R}$: a team of m robots with 360^0 field of view observation sensors that are noisy and of limited range
$\mathcal{O}(t)$: a set of n targets $o_j(t)$, such that $In(o_j(t), \mathcal{S})$ is true (where $In(o_j(t), \mathcal{S})$ means that target $o_j(t)$ is located within region $\mathcal{S}$ at time t)

Define an $m \times n$ matrix $A(t)$, where

$$a_{ij}(t) = \begin{cases} 1 & \text{if robot } r_i \text{ is monitoring target } o_j(t) \text{ in } \mathcal{S} \text{ at time } t \\ 0 & \text{otherwise} \end{cases}$$

We further define the *logical OR* operator over a vector H as:

$$\bigvee_{i=1}^{k} h_i = \begin{cases} 1 & \text{if there exists an } i \text{ such that } h_i = 1 \\ 0 & \text{otherwise} \end{cases}$$

We say that a robot is *monitoring* a target when the target is within that robot's observation sensory field of view. Then, the goal is to maximize:

$$\sum_{t=0}^{T} \sum_{j=1}^{n} \bigvee_{i=1}^{m} a_{ij}(t)$$

over time steps Δt under the assumptions listed below. In other words, the goal of the robots is to maximize the collective time during which targets in $\mathcal{S}$ are being monitored by at least one robot during the mission from $t = 0$ to $t = T$. Note that we do not assume that the membership of $\mathcal{O}(t)$ is known in advance.

In addressing this problem, we assume the following: Define *sensor_coverage*(r_i) as the area visible to robot r_i's observation sensors, for $r_i \in \mathcal{R}$. Then we assume that, in general,

$$\bigcup_{r_i \in \mathcal{R}} sensor\_coverage(r_i) \ll \mathcal{S}.$$

That is, the maximum area covered by the observation sensors of the robot team is much less than the total area to be monitored. This implies that fixed robot sensing locations or sensing paths will not be adequate in general, and that, instead, the robots must move dynamically as targets appear in order to maintain observational contact with them and to maximize the coverage of the area $\mathcal{S}$.

We further assume the following:

- The robots have a broadcast communication mechanism that allows them to send (receive) messages to (from) each other within the area $\mathcal{S}$.
- For all $r_i \in \mathcal{R}$ and for all $o_j(t) \in \mathcal{O}(t)$, $max\_v(r_i) > max\_v(o_j(t))$, where $max\_v(a)$ gives the maximum possible velocity of entity a, for $a \in \mathcal{R} \cup \mathcal{O}(t)$.
- Targets in $\mathcal{O}$ can enter and exit region $\mathcal{S}$ through distinct entrances/exits.
- The robot team members share a known global coordinate system.

To somewhat simplify the problem initially, we report here the results of the case of an omni-directional 2D sensory system (such as a ring of cameras or sonars), in which the robot sensory system is of limited range, but is available for the entire 360^o around the robot.

RELATED WORK

Research related to the multiple target observation problem can be found in a number of domains, including art gallery and related problems, multitarget tracking, and multi-robot surveillance tasks. While a complete review of these fields is not possible in a short paper, we will briefly outline the previous work that is most closely related to the topic of this paper.

The work most closely related to the *CMOMMT* problem falls into the category of the *art gallery* and related problems [1], which deal with issues related to polygon visibility. The basic art gallery problem is to determine the minimum number of guards required to ensure the visibility of an interior polygonal area. Variations on the problem include fixed point guards or mobile guards that can patrol a line segment within the polygon. Most research in this area typically utilizes centralized approaches to the placement of sensors, uses ideal sensors (noise-free and infinite range), and assumes the availability of sufficient numbers of sensors to cover the entire area of interest. Several authors have looked at the static placement of sensors for target tracking in known polygonal environments (e.g. [2]). These works differ from the *CMOMMT* problem, in that our robots must dynamically shift their positions over time to ensure that as many targets as possible remain under surveillance, and their sensors are noisy and of limited range.

Sugihara *et al.* [3] address the *searchlight scheduling problem*, which involves searching for a mobile "robber" (which we call *target*) in a simple polygon by a number of fixed searchlights, regardless of the movement of the target. They develop certain necessary and sufficient conditions for the existence of a search schedule in certain situations, under the assumption of a single target, no entrances/exits to the polygon, and fixed searcher positions

Suzuki and Yamashita [4] address the *polygon search* problem, which deals with searching for a mobile target in a simple polygon by a single mobile searcher. They examine two cases: one in which the searcher's visibility is restricted to k rays emanating from its position, and one in which the searcher can see in all directions simultaneously. Their work assumes no entrances/exits to the polygon and a single searcher.

LaValle *et al.* [5] introduces the visibility-based motion planning problem of locating an unpredictable target in a workspace with one or more robots, regardless of the movements of the target. They define a visibility region for each robot, with the goal of guaranteeing that the target will eventually lie in at least one visibility region. In LaValle *et al.* [6], they address the related question of maintaining the visibility of a moving target in a cluttered workspace by a single robot. They are also able to optimize the path along additional criteria, such as the total distance traveled. The problems they address in these papers are closely related to the problem of interest here. The primary difference is that their work does not deal with multiple robots maintaining visibility of multiple targets, nor a domain in which targets may enter and exit the area of interest.

Another large area of related research has addressed the problem of multitarget tracking (e.g. Bar-Shalom [7, 8], Blackman [9], Fox *et al.* [10]). This problem is concerned with computing the trajectories of multiple targets by associating observations of current target locations with previously detected target locations. In the general case, the sensory input can come from multiple sensory platforms. Our task in this paper differs from this work in that our goal is not to calculate the trajectories of the targets, but rather to find dynamic sensor placements that minimize the

collective time that any target is not being monitored (or observed) by at least one of the mobile sensors.

APPROACH

Overview

Since the *CMOMMT* problem can be shown to be NP-complete, and thus intractable for computing optimal solutions, we propose an approximate control mechanism that is shown to work well in practice. This approximate control mechanism is based upon our previous work, described in [11, 12], which defines a fully distributed, behavior-based software architecture called ALLIANCE that enables fault tolerant, adaptive multi-robot action selection. This architecture is a hybrid approach to robotic control that incorporates a distributed, real-time reasoning system utilizing behavioral motivations above a layer of low-level, behavior-based control mechanisms. This architecture for cooperative control utilizes no centralized control; instead, it enables each individual robot to select its current actions based upon its own capabilities, the capabilities of its teammates, a previous history of interaction with particular team members, the current state of the environment, and the robot's current sensory readings. ALLIANCE does not require any use of negotiation among robots, but rather relies upon broadcast messages from robots to announce their current activities. The ALLIANCE approach to communication and action selection results in multi-robot cooperation that gracefully degrades and/or adapts to real-world problems, such as robot failures, changes in the team mission, changes in the robot team, or failures or noise in the communication system. This approach has been successfully applied to a variety of cooperative robot problems, including mock hazardous waste cleanup, bounding overwatch, janitorial service, box pushing, and cooperative manipulation, implemented on both physical and simulated robot teams.

Our proposed approach to the *CMOMMT* problem is based upon the same philosophy of control that was utilized in ALLIANCE. In this approach, we enable each robot team member to make its own action selections, without the need for any centralized control or negotiation. The low-level, behavior based control of each robot calculates local force vectors that attract the robot to nearby targets and repel the robot from nearby teammates. Added above the low-level control is a higher-level reasoning system that generates weights to be applied to the force vectors. These weights are based upon previous experiences of the robot, and can be in the form of motivations of behavior or rule-based heuristics. The high-level reasoning system of an individual robot is thus able to influence the local, low-level control of that robot, with the aim of generating an improved collective behavior across robots when utilized by all robot team members.

Target and robot detection

Ideally, robot team members would be able to passively observe nearby robots and targets to ascertain their current positions and velocities. Research fields such as machine vision have dealt extensively with this topic, and have developed algorithms for this type of passive position calculation. However, since the physical tracking and 2D positioning of visual targets is not the focus of this research, we instead assume that robots use a global positioning system (such as GPS for outdoors, or the laser-based MTI indoor positioning system [13] that is in use at our CESAR laboratory) to determine their own position and the position of targets within their sensing range, and communicate this information to the robot team members within their communication range[1].

For each robot r_i, we define the *predictive tracking range* as the range in which targets localized by other robots $r_k \neq r_i$ can affect r_i's movements. Thus, a robot can know about two types of targets: those that are directly sensed or those that are "virtually" sensed through predictive tracking. When a robot receives a communicated message regarding the location and velocity of a sighted target that is within its predictive tracking range, it begins a predictive tracking of that target's location, assuming that the target will continue linearly from its current state. We

[1]This approach to communication places an upper limit on the total allowable number of robots and targets at about 400. Since the communication is $O(nm)$, we compute this upper limit by assuming a 1.6 Mbps Proxim radio ethernet system (such as the one in our laboratory) and assuming that messages of length 10 bytes per robot per target are transmitted every 2 seconds. With these numbers, we find that nm must be less than 4×10^4 bps to avoid saturation of the communication bandwidth.

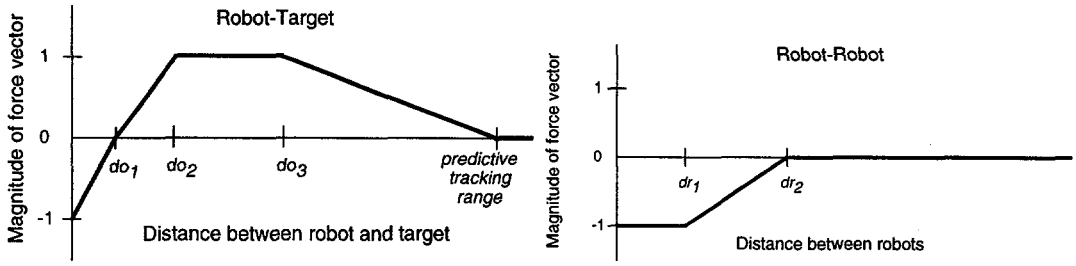


Figure 1: Functions defining the magnitude of the force vectors to nearby targets and robots.

assume that if the targets are dense enough that their position estimations do not supply enough information to disambiguate distinct targets, then existing tracking approaches (e.g. Bar-Shalom [8]) should be used to uniquely identify each target based upon likely trajectories.

Local force vector calculation

The local control of a robot team member is based upon a summation of force vectors which are attractive for nearby targets and repulsive for nearby robots. The first function in figure 1 defines the relative magnitude of the attractive forces of a target within the predictive tracking range of a given robot. Note that to minimize the likelihood of collisions, the robot is repelled from a target if it is too close to that target ($distance < do_1$). The range between do_2 and do_3 defines the preferred tracking range of a robot from an object. In practice, this range will be set according to the type of tracking sensor used and its range for optimal tracking. The attraction to the object falls off linearly as the distance to the object varies from do_2. The attraction goes to 0 beyond the predicted tracking range, indicating that this object is too far to have an effect on the robot's movements.

The second function of figure 1 defines the magnitude of the repulsive forces between robots. If the robots are too close together ($distance < dr_1$), they repel strongly. If the robots are far enough apart ($distance > dr_2$), they have no effect upon each other in terms of the force vector calculations. The magnitude scales linearly between these values.

One problem with using only force vectors, however, is that of local minima. As defined so far, the force vector computation is equivalent for all targets, and for all robots. Thus, we need to inject additional high-level reasoning control into the system to take into account more global information. This reasoning is modeled as predictive weights that are factored into the force vector calculation, and are described in the next subsection.

High-level reasoning control

To help resolve the problems of local minima, the higher-level reasoning control differentially weights the contributions of each target's force field on the total computed field. This higher-level knowledge can express any information or heuristics that are known to result in more effective global control when used by each robot team member locally. Our present approach expresses this high-level knowledge in the form of two types of probabilities: the probability that a given target actually exists, and the probability that no other robot is already monitoring a given target. Combining these two probabilities helps reduce the overlap of robot sensory areas toward the goal of minimizing the likelihood of a target escaping detection.

The probability that a target exists is modeled as a decay function based upon when the target was most recently seen, and by whom. In general, the probability decreases inversely with distance from the current robot. Beyond the predictive tracking range of the robot, the probability becomes zero.

The probability that no other robot is already monitoring a nearby target is based upon the target's position and the location of nearby robots. If the target is in range of another robot, then this probability is generally high. In the future, we plan to incorporate the ALLIANCE motivation of "impatience", if a nearby robot does not appear to be satisfactorily observing its local targets (perhaps due to faulty sensors). This impatience will effectively reduces the probability that the other robot is already monitoring nearby targets. In more complex versions of the *CMOMMT* problem, robots could also learn about the viewing capabilities of their teammates, and discount their teammates' observations if that teammate has been unreliable in the past.

The higher-level weight information is combined with the local force vectors to generate the commanded direction of robot movement. This direction of movement is given by:

$$\sum_{i=0}^{N}(FVO_i \times P(exists_i) \times P(NT_i)) + \sum_{j=0}^{M} FVR_j$$

where FVO_k is the force vector attributed to target o_k, $P(exists_k)$ is the probability that target o_k exists, $P(NT_k)$ is the probability that target o_k is not already being tracked, and FVR_l is the force vector attributed to robot r_l. This movement command is then sent to the robot actuators to cause the appropriate robot movements. We also incorporate a low-level obstacle avoidance behavior that overrides these movement commands if it would likely result in a collision.

EXPERIMENTAL RESULTS AND DISCUSSION

To evaluate the effectiveness of the algorithm we designed for the *CMOMMT* problem (which we will refer to as *A-CMOMMT*, we conducted experiments both in simulation and on a team of mobile robots. In the simulation studies, we compared four possible cooperative observation algorithms: (1) *A-CMOMMT* (high-level plus local control), (2) *Local control only*, (3) *Random/linear robot movement*, and (4) *Fixed robot positions*.

In all of these experiments, targets moved according to a "random/linear" movement, which causes the target to move in a straight-line until an obstacle is met, followed by random turns until the target is able to again move forward without collision. The *local control only* algorithm computed the motion of the robots by calculating the unweighted local force vectors between robots and targets. This approach was studied to determine the effectiveness of the high-level reasoning that is incorporated into the *A-CMOMMT* algorithm. The last two algorithms are control cases for the purposes of comparison: the *random/linear robot movement* approach caused robots to move according the the "random/linear" motion defined above, while the *fixed robot positions* algorithms distributed the robots uniformly over the area $\mathcal{S}$, where they maintained fixed positions. In both of these control approaches, robot movements were not dependent upon target locations or movements (other than obstacle avoidance).

We compared these 4 approaches by measuring the average value of the $A(t)$ matrix (see PROBLEM DESCRIPTION section) during the execution of the algorithm. Since the algorithm performance is expected to be a function f of the number of robots n, number of targets m, the range of a given robot's sensor r, and the relative size of the area $\mathcal{S}$, we collected data for a wide range of values of these variables. To simplify the analysis of our results, we defined the area $\mathcal{S}$ as the area within a circle of radius R, fixed the range of robot sensing at 2,600 units of distance, and included no obstacles within $\mathcal{S}$ (other than the robots and targets themselves, and the boundary of $\mathcal{S}$).

We collected data by varying n from 1 to 10, m from 1 to 20, and R from 1,000 to 50,000 units. For each instantiation of variables n, m, and R, we computed the average $A(t)$ value every $\Delta t = 2$ seconds of a run of length 2 minutes; we then repeated this process for 250 runs for each instantiation to derive an average $A(t)$ value for the given values of n, m, and R. In all runs of all 4 algorithms, the targets were placed randomly at the center of $\mathcal{S}$ within a circle of radius 1,000. In all runs of all algorithms (except for *fixed robot positions*), the robots were also placed randomly within the same area as the targets.

To analyze the results of these experiments, we speculated that the function $f(n, m, r, R)$ would be proportional to ratio of the total collective area that could be covered by the robot sensors (i.e. $n\pi r^2$) over the area that would be allotted to one target (call it a *target slot*), were $\mathcal{S}$ divided equally over all targets (i.e. $\frac{\pi R^2}{m}$), we have:

$$f(n, m, r, R) = \frac{n\pi r^2}{\frac{\pi R^2}{m}} = \frac{nmr^2}{R^2}$$

Thus, this function was used to compare the similarity of experiments that varied in their instantiations of n, m, and R.

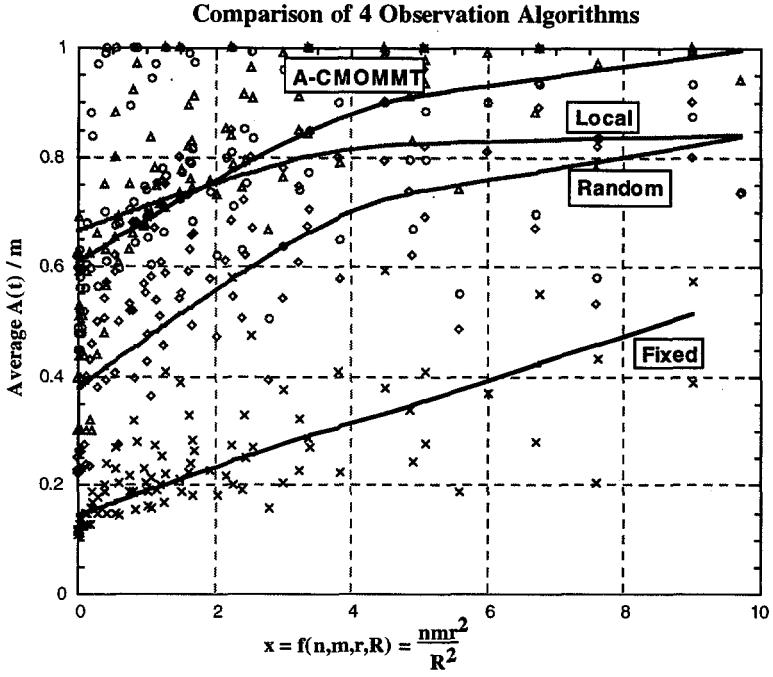


Figure 2: Comparison of 4 cooperative observation algorithms.

Since the optimum value of the average $A(t)$ for a given experiment depends upon the value of m (and, in fact, equals m), we normalized the experiments by plotting the average $A(t)/m$ which is the average percentage of targets that are within some robot's view at a given instant of time.

Figure 2 gives the results of our experiments, plotting the average $A(t)/m$ versus $f(n, m, r, R)$ for all of our experimental data. For each algorithm, we fit a curve to the data using the locally weighted Least Squared error method. Since there is considerable deviation in the data points for given values of $f(n, m, r, R)$, we computed the statistical significance of the results using the Student's t distribution, comparing the algorithms two at a time for all 6 possible pairings. In these computations, we used the null hypothesis: $H_0 : \mu_1 = \mu_2$, *and there is essentially no difference between the two algorithms.* Under hypothesis H_0:

$$T = \frac{\overline{X}_1 - \overline{X}_2}{\sigma\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \quad where \quad \sigma = \sqrt{\frac{n_1 S_1^2 + n_2 S_2^2}{n_1+n_2-2}}$$

Then, on the basis of a two-tailed test at a 0.01 level of significance, we would reject H_0 if T were outside the range $-t_{.995}$ to $t_{.995}$, which for $n_1 + n_2 - 2 = 250 + 250 - 2 = 498$ degrees of freedom, is the range -2.58 to 2.58. For the data given in figure 2, we found that we could reject H_0 at a 0.01 level of significance for all pairing of algorithms that show a visible difference in performance in this figure. Thus, we can conclude that the variation in performance of the algorithms illustrated by the fitted curves in figure 2 is significant.

We see from figure 2 that the *A-CMOMMT* and *local control only* algorithms perform better than the two naive control algorithms, which is expected since the naive algorithms use no information about target positions. Note that all approaches improve as the value of $f(n, m, r, R)$ increases, corresponding to a higher level of robot coverage available per target. The *random/linear robot movement* approach performed better than the *fixed robot positions*, most likely due to the proximity of the initial starting locations of the robots and objects in the *random/linear robot movement* approach. This seems to suggest that much benefit can be gained by learning areas of the environment $\mathcal{S}$ where targets are more likely to be found, and concentrate on locating robots in those areas.

Of more interest, we see that the *A-CMOMMT* approach is superior to the *local control only* approach for values of $f(n, m, r, R)$ greater than about 2; the *local control only* approach is slightly better for $f(n, m, r, R)$ less than 2. This means that when the fraction of robot coverage available per target is low (< 2), relative to the size of $\mathcal{S}$, then robots are better off *not* ignoring any targets, which is essentially what happens due to the high-level control of *A-CMOMMT*. Examples of experimental scenarios where the *local control only* approach is better than the *A-CMOMMT* approach are (n, m, R) = (2,1,5000-50000), (2,2,4000-50000), (3,1,5000-50000), (3,2,5000-50000), (3,3,8000-50000), and (3,4,8000-50000). However, for more complex cases, where the number of targets is much greater than the number of robots, and the environmental area is not "too large", we find that the higher-level reasoning provided by *A-CMOMMT* works better. Examples of scenarios where *A-CMOMMT* is better include (n, m, R) = (2,4,1000-5000), (2,6,1000-6000) (2,20,1000-10000),(3,3,1000-5000), (3,4,1000-6000), (3,6, 1000-7000), and (3,12,1000-11000). Note that *A-CMOMMT* approaches perfect performance as $f(n, m, r, R)$ reaches 10, whereas the results of the *random/linear robot movement* and *local control only* approaches begin to level off at around 85%. In continuing and future work, we are determining the impact of these results on multi-robot cooperative algorithm design.

We have also implemented the *A-CMOMMT* algorithm on a team of a team of four Nomadic Technologies robots to illustrate the feasibility of our approach for physical robot teams. We have demonstrated a very simple case of cooperative tracking using these robots. Refer to [14] for details.

CONCLUSIONS

Many real-world applications in security, surveillance, and reconnaissance tasks require multiple targets to be monitored using mobile sensors. We have presented an approximate, distributed approach based upon the philosophies of the ALLIANCE architecture and have illustrated its effectiveness in a wide range of cooperative observation scenarios. This approach is based upon a combination of high-level reasoning control and lower-level force vector control that is fully distributed across all robot team members and involves no centralized control. Empirical investigations of our cooperative control approach have shown it to be effective at achieving the goal of maximizing target observation for most experimental scenarios, as compared to three other feasible control algorithms.

ACKNOWLEDGEMENTS

This research is sponsored by the Office of Engineering Research, Basic Energy Sciences of the U.S. Department of Energy, under contract No. DE-AC05-96OR22464 with Lockheed Martin Energy Research Corporation.

REFERENCES

[1] J. O'Rourke. Art Gallery Theorems and Algorithms. Oxford University Press, 1987.

[2] Amy J. Briggs. Efficient Geometric Algorithms for Robot Sensing and Control. PhD thesis, Cornell University, 1995.

[3] Kzuo Sugihara, Ichiro Suzuki, and Masafumi Yamashita. The searchlight scheduling problem. SIAM Journal of Computing, 19(6):1024–1040, 1990.

[4] Ichiro Suzuki and Masafumi Yamashita. Searching for a mobile intruder in a polygonal region. SIAM Journal of Computing, 21(5):863–888, 1992.

[5] S. M. LaValle, D. Lin, L. J. Guibas, J-C. Latombe, and R. Motwani. Finding an unpredictable target in a workspace with obstacles. In submitted to 1997 International Conference on Robots and Automation, 1997.

[6] S. M. LaValle, H. H. Gonzalez-Banos, C. Becker, and J-C. Latombe. Motion strategies for maintaining visibility of a moving target. In submitted to 1997 International Conference on Robots and Automation, 1997.

[7] Yaakov Bar-Shalom. Tracking methods in a multitarget environment. IEEE Transactions on Automatic Control, AC-23(4):618–626, 1978.

[8] Yaakov Bar-Shalom. Multitarget Multisensor Tracking: Advanced Applications. Artech House, 1990.

[9] S. S. Blackman. Multitarget Tracking with Radar Applications. Artech House, 1986.

[10] G.C. Fox, R.D. Williams, and P.C. Messina. Parallel Computing Works. Morgan Kaufmann, 1994.

[11] Lynne E. Parker. ALLIANCE: An architecture for fault tolerant, cooperative control of heterogeneous mobile robots. In Proc. of the 1994 IEEE/RSJ/GI Int'l Conf. on Intelligent Robots and Systems (IROS '94), pages 776–783, Munich, Germany, Sept. 1994.

[12] Lynne E. Parker. Heterogeneous Multi-Robot Cooperation. PhD thesis, Massachusetts Institute of Technology, Artificial Intelligence Laboratory, Cambridge, MA, February 1994. MIT-AI-TR 1465 (1994).

[13] MTI Research Inc. Conac 3-D tracking system. Operating manual, Chelmsford, MA, 1995.

[14] Lynne E. Parker and Brad Emmons. Cooperative multi-robot observation of multiple moving targets. In Proceedings of the 1997 IEEE International Conference on Robotics and Automation, pages 2082–2089, Albuquerque, New Mexico, April 1997.

GLOBAL OPTIMIZATION FOR MULTISENSOR FUSION IN SEISMIC IMAGING

Jacob Barhen, Vladimir Protopopescu, and David Reister

Center for Engineering Systems Advanced Research
Oak Ridge National Laboratory
Oak Ridge, TN 37831-6355

ABSTRACT

The accurate imaging of subsurface structures requires the fusion of data collected from large arrays of seismic sensors. The fusion process is formulated as an optimization problem and yields an extremely complex "energy surface". Due to the very large number of local minima to be explored and escaped from, the seismic imaging problem has typically been tackled with stochastic optimization methods based on Monte Carlo techniques. Unfortunately, these algorithms are very cumbersome and computationally intensive. Here, we present TRUST - a novel deterministic algorithm for global optimization that we apply to seismic imaging. Our excellent results demonstrate that TRUST may provide the necessary breakthrough to address major scientific and technological challenges in fields as diverse as seismic modeling, process optimization, and protein engineering.

INTRODUCTION

In many geophysical tasks, seismic energy is detected by receivers which are regularly spaced along a grid that covers the explored domain. A source is positioned at some grid node to produce a shot. Time series data is collected from the detectors for each shot; then the source is moved to another grid node for the next shot. A major degradation of seismic signals usually arises from near-surface geologic irregularities [1, 2]. These include uneven soil densities, topography, and significant lateral variations in the velocity of seismic waves. The most important consequence of such irregularities is a *distorted image of the subsurface structure*, due to misalignment of signals caused by unpredictable delays in recorded travel times of seismic waves in a vertical neighborhood of every source and receiver. To improve the quality of the seismic analysis, timing adjustments (called "statics corrections") must be performed. One typically distinguishes between "*field statics*", which correspond to corrections that can be derived directly from topographic and well measurements, and "*residual statics*", which incorporate adjustments that must be inferred statistically from the seismic data. The common occurrence of severe residual statics (where the dominant period of the recorded data is significantly exceeded), and the significant noise contamination render the automatic identification of large static shifts extraordinarily difficult. Thus, *multisensor fusion* must be invoked [3]. This problem has generally been formulated in terms of global optimization

and, to date, Monte-Carlo techniques (e.g., simulated annealing, genetic algorithms) have provided the primary tools for seeking a potential solution.

The objective function associated with the task of fusing data from an array of seismic sensors depends on a very large number of parameters. Finding the extrema and, in particular, the absolute extremum of such a function turns out to be painstaking difficult. The primary difficulty stems from the fact that the global extremum, say minimum, of a real function is - despite its name - a local property. In other words, significant alteration of the location and magnitude of the global minimum can be carried out without affecting at all the locations and magnitudes of the other minima. Short of exhaustive search, it would then appear extraordinarily unlikely to design unfallible methods to locate the absolute minimum for an arbitrary function. In recent years there has been a remarkable surge of interest in global optimization [5 - 8]. Although significant progress has been achieved in breaking new theoretical ground [9 - 19], the need for efficient and reliable global optimization methods remains as urgent as ever. In particular, a major need exists for a breakthrough paradigm which would enable the accurate and efficient solution of *large-scale* problems. In response to that need, we have been focusing, at ORNL's Center for Engineering Systems Advanced Research (CESAR), on two innovative concepts, namely subenergy tunneling and non-Lipschitzian terminal repellers, to ensure escape from local minima in a fast, reliable, and computationally efficient manner. The generally applicable methodology is embodied in the TRUST algorithm [4], which is deterministic, scalable, and easy to implement. Benchmark results show that TRUST is considerably faster and more accurate than previously reported global optimization techniques. Hence, TRUST may provide the enabling means for addressing major scientific and technological challenges in fields as diverse as seismic modeling, process optimization, and protein engineering.

The classical theory of optimization started to develop almost concomitantly with classical mechanics by trying to find extremal values (minima or maxima) of certain functions that bear special physical meaning and practical significance. For instance, Newton studied projectile trajectories and obtained their maximum range by taking into account the friction with the atmosphere. He was also interested in minimizing resistance by modifying the shape of an object propelled through water. The Bernoulli brothers, who were active in Switzerland between 1670 and 1720, discovered that the shortest time of descent between two points under gravity is achieved not on the straight line joining the two points, but on a convex curve, called brachistocrone. Another famous optimization problem is to find the greatest area enclosed between a straight line and an arbitrary curve of fixed length joining two points on the line. By Virgil's account (Aeneid, Book I, line 367), Queen Dido solved the problem by determining the shape of the curve and the position of the points, thereby founding Carthage.

The completion of the main body of classical physics around the turn of the century came with the realization that many natural processes take place according to extremal principles such as: (i) the principle of stationary (minimum) action in mechanics and electrodynamics; (ii) the principle of minimal potential energy in stable mechanical equilibrium states; (iii) the principle of maximal entropy in isolated thermodynamic systems at equilibrium; and (iv) the principle of motion along geodesics (Fermat's principle in geometrical optics and Einstein's principle in relativity theory). Thenceforth extremal principles and, more generally, optimization problems have been perceived as a systematic and elegant framework for addressing and solving more complex problems with applications to economy, investment policies, and social or political negotiations. In these domains, optimization is, in turn, used to determine "the best" model for a complex situation , to make "the best" choice within a given model, and to solve the associated, purely technical, sub-problems that

occur in the mathematical analysis and implementation of the model. In this context, optimality is, almost always, to be obtained under certain constraints and/or at the expense of a certain price.

The generic *global optimization* problem can be stated as follows. The overall performance of a system is described by a multivariate function, called the objective function. Optimality of the system is reached when the objective function attains its global extremum, which can be a maximum or a minimum, depending on the problem under consideration. From an algorithmic perspective, however, there is essentially no difference between the two.

THE TRUST ALGORITHM

We now define the global optimization problem considered in more rigorous terms. Let $f(\mathbf{x}) : \mathcal{D} \rightarrow \mathcal{R}$ be a function with a finite number of discontinuities, and $\mathbf{x}$ be an n-dimensional state vector. At any discontinuity point, $\mathbf{x}^{\delta}$, the function $f(\cdot)$ is required to satisfy the inequality $\lim_{\mathbf{x}\rightarrow\mathbf{x}^{\delta}} \inf f(\mathbf{x}) \geq f(\mathbf{x}^{\delta})$ (lower semicontinuity condition). Hereafter, $f(\mathbf{x})$ will be referred to as the objective function, and the set $\mathcal{D}$ as the set of feasible solutions (or the solution space). The goal is to find location of the global minimum, i.e. the value $\mathbf{x}^{gm}$ of the state variables which minimizes $f(\mathbf{x})$,

$$f(\mathbf{x}^{gm}) = \min\{f(\mathbf{x}) \mid \mathbf{x} \in \mathcal{D}\} \ . \tag{1}$$

Without loss of generality, we shall take $\mathcal{D}$ as the hyperparallelepiped

$$\mathcal{D} = \{x_i \mid \beta_i^- \leq x_i \leq \beta_i^+ \ ; \ \ i = 1, 2, \ldots, n\} \ . \tag{2}$$

where β_i^- and β_i^+ denote, respectively, the lower and upper bound of the i-th state variable.

We define the *subenergy tunneling* transformation of the function $f(\mathbf{x})$ by the following nonlinear monotonic mapping:

$$E_{sub}(\mathbf{x}, \mathbf{x}^*) = \log(1/[1 + \exp(-\hat{f}(\mathbf{x}) - a)]) \ . \tag{3}$$

In Eq. (3), $\hat{f}(\mathbf{x}) = f(\mathbf{x}) - f(\mathbf{x}^*)$, a is a constant that affects the asymptotic behavior, but not the monotonicity, of the transformation, and $\mathbf{x}^*$ is a fixed value of $\mathbf{x}$, whose selection will be discussed in the sequel. Whenever f is differentiable, the derivative of $E_{sub}(\mathbf{x}, \mathbf{x}^*)$ with respect to $\mathbf{x}$ is given by

$$\partial E_{sub}(\mathbf{x}, \mathbf{x}^*)/\partial \mathbf{x} = (\partial f(\mathbf{x})/\partial \mathbf{x})(1/[1 + \exp(\hat{f}(\mathbf{x}) + a)]) \ , \tag{4}$$

which yields

$$\partial E_{sub}(\mathbf{x}, \mathbf{x}^*)/\partial \mathbf{x} = 0 \quad \Leftrightarrow \quad \partial f(\mathbf{x})/\partial \mathbf{x} = 0 \ . \tag{5}$$

It is clear that $E_{sub}(\mathbf{x}, \mathbf{x}^*)$ has the same discontinuity and critical points as $f(\mathbf{x})$, and the same relative ordering of the local and global minima. In other words, $E_{sub}(\mathbf{x}, \mathbf{x}^*)$ is a transformation of $f(\mathbf{x})$ which preserves all properties relevant for optimization. In addition, this transformation

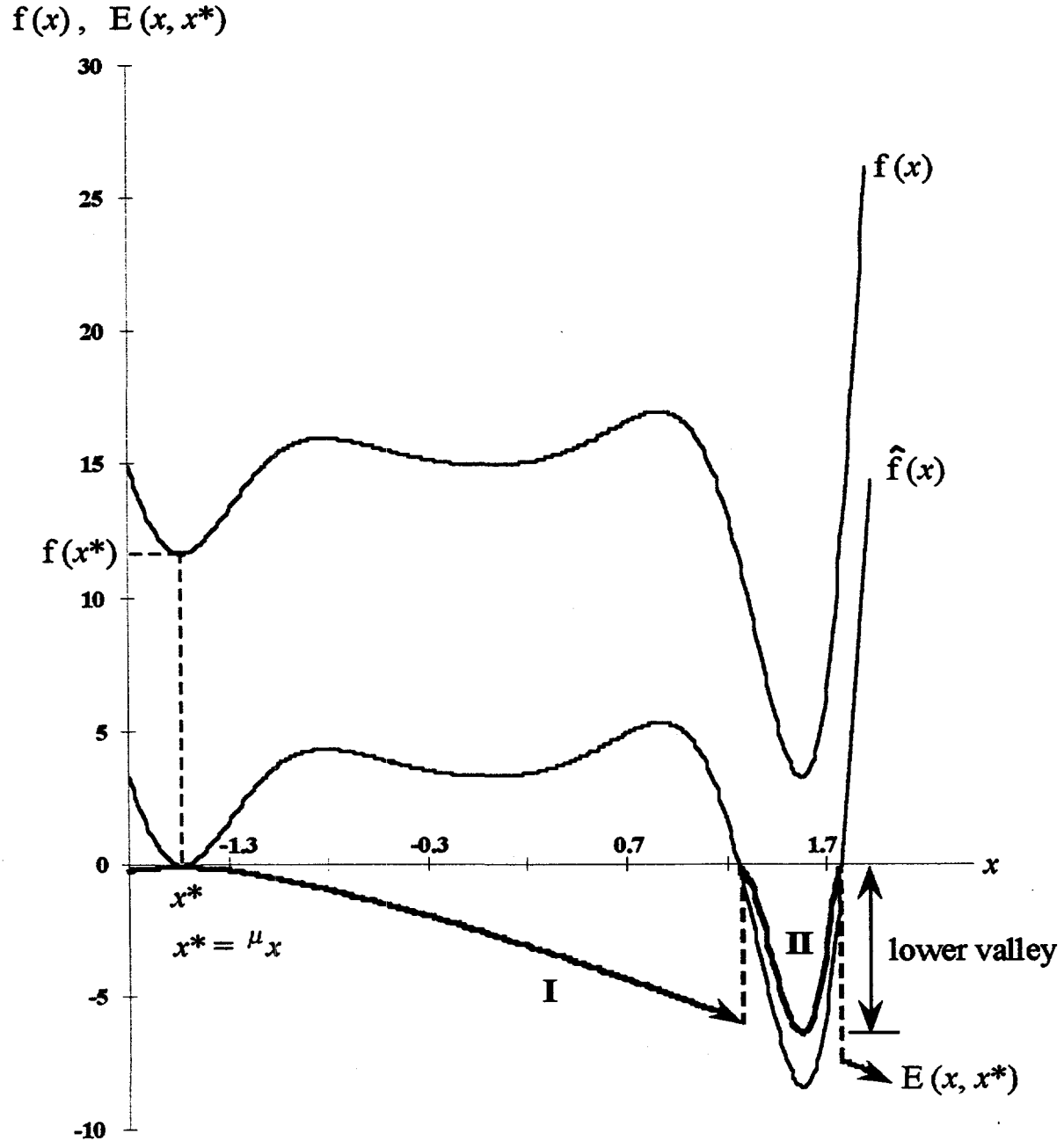


Figure1 . Operation of TRUST, illustrated on the function $f(x) = 4x^2 e^{2\alpha(x-1)} \sin[\frac{\pi}{8}(4x^2+3)]$, with $\alpha \simeq 1.22$.

is designed to ensure that: (i) $E_{sub}(\mathbf{x}, \mathbf{x}^*)$ quickly approaches zero for large positive $\hat{f}(\mathbf{x})$; and (ii) $E_{sub}(\mathbf{x}, \mathbf{x}^*)$ rapidly tends toward $\hat{f}(\mathbf{x})$, whenever $\hat{f}(\mathbf{x}) \ll 0$.

An equilibrium point $\mathbf{x}_{eq}$ of the dynamical system $\dot{\mathbf{x}} = \mathbf{g}(\mathbf{x})$ is termed an attractor (repeller) if no (at least one) eigenvalue of the $n \times n$ matrix $\mathcal{M}$, $\mathcal{M} = \partial \mathbf{g}(\mathbf{x}_{eq})/\partial \mathbf{x}$ has a positive real part. Typically, a certain amount of regularity (Lipschitz condition) is required to guarantee the existence of a unique solution for each initial condition $\mathbf{x}(0)$ and, in those cases, the system's relaxation time to an attractor, or escape time from a repeller, is theoretically infinite. If the regularity condition at equilibrium points is violated, singular solutions are induced, such that each solution approaches a *terminal attractor* or escapes from a *terminal repeller* in finite time. The above concepts are at the foundation of our Terminal Repeller Unconstrained Subenegy Tunneling (TRUST) global optimization algorithm.

Let $f(\mathbf{x})$ be a function one wishes to globally minimize over $\mathcal{D}$. We define the TRUST *virtual objective function*

$$\begin{aligned} E(\mathbf{x}, \mathbf{x}^*) &= \log(1/[1 + \exp(-\hat{f}(\mathbf{x}) - a)]) - (3/4)\rho(\mathbf{x} - \mathbf{x}^*)^{4/3}\theta(\hat{f}(\mathbf{x})) \\ &= E_{sub}(\mathbf{x}, \mathbf{x}^*) + E_{rep}(\mathbf{x}, \mathbf{x}^*). \end{aligned} \tag{6}$$

In the above expression $\theta(\cdot)$ denotes the Heaviside function, that is equal to one for positive values of the argument and zero otherwise. The first term on the right-hand side of Eq. (6) corresponds to the subenergy tunneling function; the second term is referred to as the repeller energy term. The parameter $\rho > 0$ quantifies the strength of the repeller. Application of gradient descent to $E(\mathbf{x}, \mathbf{x}^*)$ results in the dynamical system$(i = 1, ..n)$

$$\dot{x}_i = -(\partial f(\mathbf{x})/\partial x_i)(1/[1 + \exp(\hat{f}(\mathbf{x}) + a)]) + \rho(x_i - x_i^*)^{1/3}\theta(\hat{f}(\mathbf{x})) \ . \tag{7}$$

Figure 1 illustrates the main characteristics of TRUST for a one-dimensional problem objective function $E(x, x^*)$. A schematical representation of a sufficiently smooth $f(x)$ is shown, which has three local minima, one of which is the global minimum. We assume that the solution flows in the positive direction (i.e., away from the left boundary), and that the local minimum at $x = {}^{\mu}x$ is encountered by a local minimization method, gradient descent for example. The task under consideration is to escape this local minimum, in order to reach the valley of another minimum with a lower value. We set $x^* = {}^{\mu}x$; then the objective function in Eq. (6) performs the following transformation (see Figure 1):

- the offset function $\hat{f}(x) = f(x) - f(x^*)$ creates the curve parallel to $f(x)$, such that the local minimum at $x = x^*$ intersects with the x-axis tangentially;
- the term $E_{sub}(x, x^*)$ forms the portion of the thick line denoted by **II** (i.e., the lower valley) as a result of the properties of the subenergy transformation;
- the repeller energy term $E_{rep}(x, x^*)$ essentially constitutes the portion of the thick line denoted by **I**;
- finally, as the complete thick line (i.e., **I** and **II**) shows, the virtual objective function $E(x, x^*)$, which is a superposition of these two terms, creates a discontinuous but well-defined function with a *global maximum* located at the previously specified local minimum ${}^{\mu}x$.

To summarize, as seen in Figure 1, $E(x, x^*)$ of Eq. (7) transforms the current local minimum of $f(x)$ into a global maximum, but preserves all lower local minima. Thus, when gradient descent is applied to the function $E(x, x^*)$, the new dynamics, initialized at a small perturbation from the local minimum of $f(x)$ (i.e., at $x = x^* + d$, with $x^* = {}^{\mu}x$), will escape this critical point (which is also the global maximum of $E(x, x^*)$) to a lower valley of $f(x)$ with a lower local minimum. It is important to note that the discontinuity of $E(x, x^*)$ does not affect this desired operation, since the flow of the gradient descent dynamics follows the gradient of $E(x, x^*)$, which is well-defined at every point in the region. It is clear that if gradient descent were to be applied to the objective function $f(x)$ under the same conditions, escaping the local minimum at $x = {}^{\mu}x$ would not be accomplished.

Hence, application of gradient descent to the function $E(x, x^*)$ as defined in Eq. (6), as opposed to the original function $f(x)$, results in a system that has a *global descent property*, i.e., the new system escapes the encountered local minimum to another one with a lower functional value. This is the main idea behind constructing the TRUST virtual objective function of Eq. (6). Additional details and formal derivations can be found in [4, 15, 18].

BENCHMARKS AND COMPARISONS TO OTHER METHODS

This section presents results of benchmarks carried out to assess the TRUST algorithm using several standard test functions taken from the literature. A description of each test function is given in Table 1. In Tables 2-3, the performance of TRUST is compared to the best competing global optimization methods, where the term "best" indicates the best widely reported reproducible results the authors could find for the particular test function. The criterion for comparison is the number of function evaluations.

One of the primary limitations of conventional global optimization algorithms is their lack of stopping criteria. This limitation is circumvented in benchmark problems, where the value and coordinates of the global minima are known in advance. The achievement of a desired accurracy (e.g., $\epsilon = 10^{-6}$) is then considered as a suitable termination condition [6]. For consistent comparisons, this condition has also been used in TRUST, rather than its general stopping criterion described earlier. For each function, corners of the domain were taken as initial conditions; each reported result then represents the average number of evaluations required for convergence to the global minimum of the particular function. The TRUST calculations were performed using the value $a = 2$, for which the subenergy tunneling transformation achieves its most desirable asymptotic behavior [15]. The dynamical equations were integrated using an adapitive scheme, that, within the basin of attraction of a local minimum, considers the local minimum as a terminal attractor. Typical base values for the key parameters Δ_t and ρ were 0.05 and 10., respectively.

In Table 2, the benchmark labels, i.e. BR (Branin), CA (Camelback), GP (Goldstein-Price), RA (Rastrigin), SH (Shubert) and H3 (Hartman), refer to the test functions specified in Table 1. The following abbreviations are also used: SDE is the stochastic method of Aluffi-Pentini [9]; EA is the annealing evolution algorithms of Yong, Lishan, and Evans [17] and Schneider [19]; MLSL is the multiple level single linkage method of Kan and Timmer [10]; IA is the interval arithmetic technique of Ratschek and Rokne [19]; TUN is the tunneling method of Levy and Montalvo [11]; and TS refers to the Taboo Search scheme of Cvijovic and Klinowski [16]. The results demonstrate that TRUST is substantially faster than these state-of-the-art methods.

Table 1. Standard Test Functions used for global optimization benchmarks.

Name	Definition	Domain
Branin	$f(\mathbf{x}) = [x_2 - (5.1/4\pi^2)x_1^2 + (5/\pi)x_1 - 6]^2 + 10(1 - 1/8\pi)\cos x_1 + 10$	$-5. \leq x_1 \leq +10.$ $0. \leq x_2 \leq +15.$
Camelback	$f(\mathbf{x}) = [4 - 2.1x_1^2 + (x_1^4/3)]\,x_1^2 + x_1x_2 + (-4 + 4x_2^2)x_2^2$	$-3. \leq x_1 \leq +3.$ $-2. \leq x_2 \leq +2.$
Goldstein-Price	$f(\mathbf{x}) = [1 + (x_1 + x_2 + 1)^2 (19 - 14x_1 + 3x_1^2 - 14x_2 + 6x_1x_2 + 3x_2^2)]$ $\times[30 + (2x_1 - 3x_2)^2 (18 - 32x_1 + 12x_1^2 + 48x_2 - 36x_1x_2 + 27x_2^2)]$	$-2. \leq x_i \leq +2.$
Rastrigin	$f(\mathbf{x}) = x_1^2 + x_2^2 - \cos(18x_1) - \cos(18x_2)$	$-1. \leq x_i \leq +1.$
Shubert	$f(\mathbf{x}) = \left\{\sum_{i=1}^{5} i\cos[(i+1)x_1 + i]\right\}\left\{\sum_{i=1}^{5} i\cos[(i+1)x_2 + i]\right\}$	$-10. \leq x_i \leq +10.$
Hartman*	$f(\mathbf{x}) = \sum_{i-1}^{i=4} c_i \exp[-\sum_{j=1}^{j=N} a_{ij}(x_j - p_{ij})^2]$	$0. \leq x_i \leq 1.$
Styblinski and Tang	$f(\mathbf{x}) = \frac{1}{2}\sum_{i=1}^{i=2}(x_i^4 - 16x_i^2 + 5x_i) + \sum_{i=3}^{i=5}(x_i - 1)^2$	$-4.6 \leq x_i \leq +4.6$

(*)The values of the parameters are given in ([6], p. 185).

Table 2. Number of function evaluations required by different methods to reach a global minimum of Standard Test Functions.

Method	BR	CA	GP	RA	SH	H3
SDE	2700	10822	5439	–	241215	3416
EA	430	–	460	5917	–	–
MLSL	206	–	148	–	–	197
IA	1354	326	–	–	7424	–
TUN	–	1469	–	–	12160	–
TS	492	–	486	540	727	508
TRUST	55	31	103	59	72	58

Table 3. Number of function evaluations and precision for Styblinski and Tang function. Global minimum FSA and SAS results taken from Ref. [13].

Method	FSA	SAS	TRUST	Exact
Cost	100,000	3,710	89	n/a
x_1	-2.702844	-2.903506	-2.90353	-2.903534
x_2	-3.148829	-2.903527	- 2.90353	-2.903534
x_3	1.099552	1.000241	1.00004	1.
x_4	1.355916	0.999855	0.99997	1.
x_5	1.485936	1.000194	0.99997	1.

In Table 3, FSA is the fast simulated annealing algorithm of Szu [12], and SAS denotes the stochastic approximation paradigm of Styblinski and Tang [13]. As can be observed, TRUST is not only much faster, but produces very consistent and accurate results. Therefore, it seemed the ideal candidate for the solution of the notoriously difficult problem of multisensor fusion for seismic imaging, formulated as residual statics optimiation.

RESIDUAL STATICS CORRECTIONS FOR SEISMIC DATA

Statics optimization is typically done in a surface consistent manner to seismic traces corrected for *normal moveout* [3]; consequently, the correction time shifts depend only on the shot and receiver positions, and not on the ray path from shot to receiver. Shot corrections $\mathbf{S}$ correspond to wave propagation times from the shot locations to a reference plane, while the receiver corrections $\mathbf{R}$ are propagation times from the reference plane to receiver locations. From an operational perspective, data D_{ft} are provided by trace ($t = 1, \ldots N_t$), and sorted to midpoint offset coordinates (common midpoint *stacking*). For each trace, the data consist of the complex Fourier components ($f = 1, \ldots N_f$) of the collected time series. Each trace t corresponds to seismic energy travel from a source s_t to a receiver r_t via a midpoint k_t. Assuming the availability of N_k common midpoints, we seek statics corrections $\mathbf{S}$ and $\mathbf{R}$ that maximize the total power E in the stacked data:

$$E = \sum_k \sum_f |\sum_t \exp[2\pi i f (S_{s_t} + R_{r_t})] D_{ft} \delta_{kk_t}|^2 \ . \tag{8}$$

The above expression highlights the multimodal nature of E which, even for relatively low dimensional $\mathbf{S}$ and $\mathbf{R}$, exhibits a very large number of local minima. This is illustrated in Figure 2.

To assess the performance of TRUST, we considered a problem involving 77 shots and 77 receivers. A dataset consisting of 1462 synthetic seismic traces folded over 133 common midpoint gathers was obtained from CogniSeis Corporation (J. DuBose). It uses 49 Fourier components for data representation. Even though this set is somewhat smaller than typical collections obtained during seismic surveys by the oil industry, it is representative of the extreme complexity underlying residual statics problems. To derive a quantitative estimate of TRUST's impact, let E_k denote the total contribution to the stack power arising from midpoint k, and let B_k refer to the upper bound of E_k in terms of $\mathbf{S}$ and $\mathbf{R}$. Using a polar coordinates representation for the trace data D_{ft}, i.e., writing $D_{ft} = \alpha_{ft} \exp(iw_{ft})$, we can prove that

$$B_k = \sum_f (\sum_t \alpha_{ft} \delta_{kk_t})^2 \ . \tag{9}$$

The TRUST results, illustrated in Figure 3, show the dramatic improvement in the coherence factor of each common gather. This factor is defined as the ratio $\kappa_k = E_k / B_k$, and characterizes the overall quality of the seismic image.

CONCLUSIONS

TRUST is a novel methodology for unconstrained global function optimization, that combines the concepts of subenergy tunneling and non-Lipschitzian "terminal repellers." The evolution of a deterministic nonlinear dynamical system incorporating these concepts provides the computational

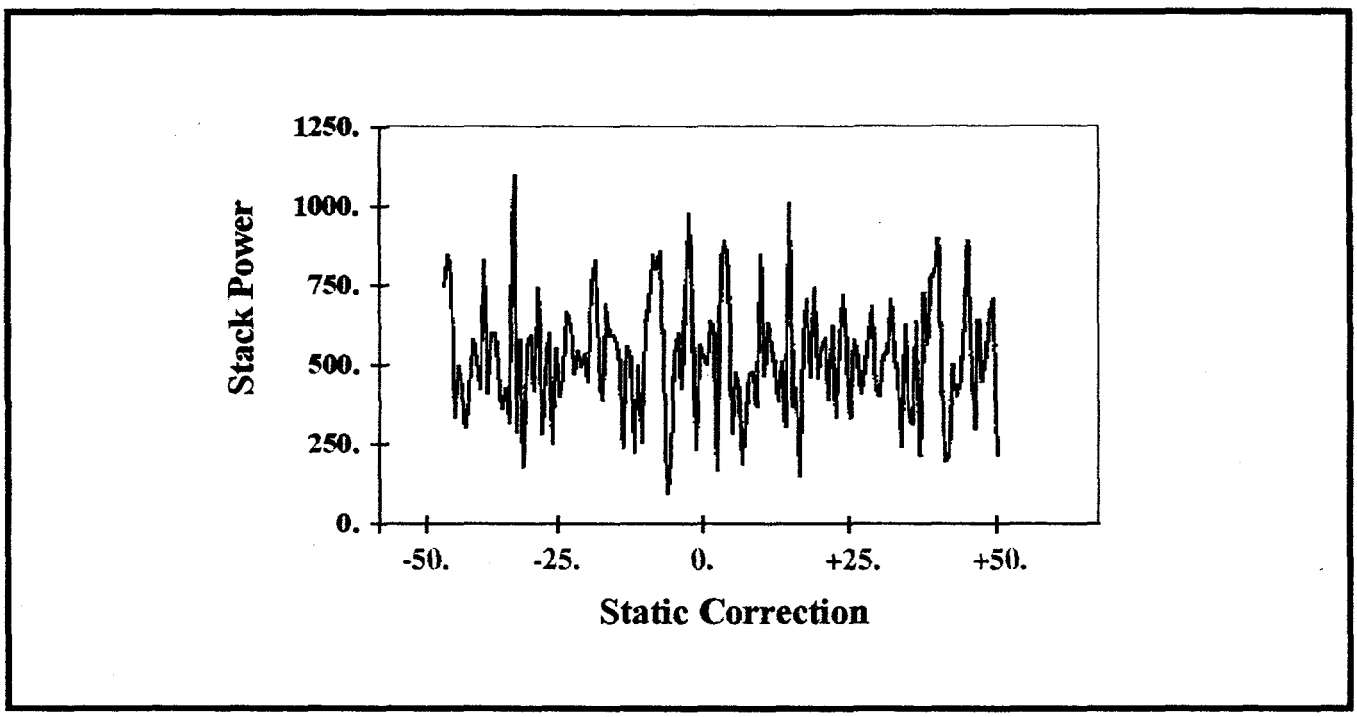


Figure 2. One-dimensional slice through a 154-dimensional objective function associated with a residual statics problem.

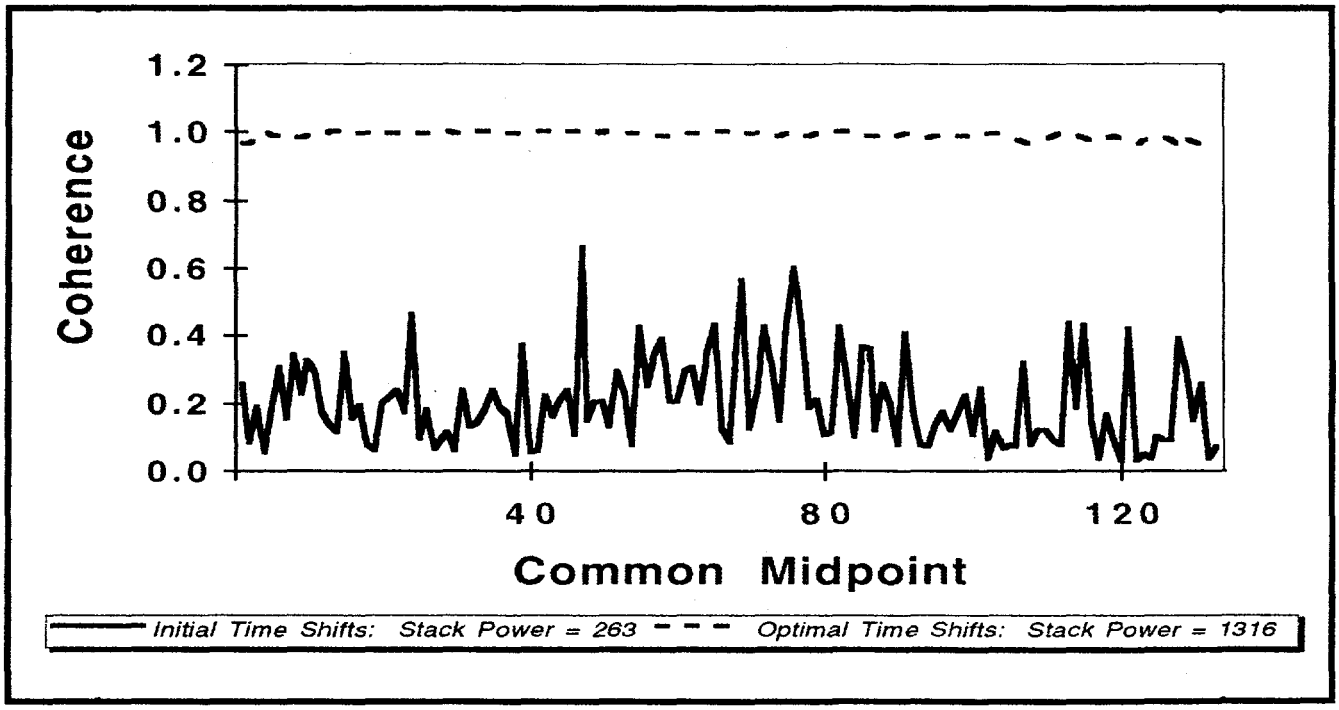


Figure 3. The coherence factors, i.e., the dimensionless ratios E_k/B_k, are plotted for each common gather using the initial and the optimal time shifts ("residual statics"). Ideally, at the global optimum, these ratios should be equal to one.

mechanism for reaching the global minima. The benchmark results demonstrate that TRUST is substantially faster, as measured by the number of function evaluations, than other global optimization techniques for which reproducible results have been published in the open literature. The application of TRUST to the problem of multisensor fusion for accurate seismic imaging (residual statics corrections) proves that the method is not a mere academic exercise for toy problems, but has the robustness and consistency required by large-scale, real-life applications.

Acknowledgments

The research described in this paper was performed at the Center for Engineering Systems Advanced Research, Oak Ridge National Laboratory. Funding for the effort was provided by the Engineering Research Program of the Office of Basic Energy Sciences, and by the Advanced Computing Technology Initiative Program of the U.S. Department of Energy, under contract DE-AC05-96OR22464 with Lockheed Martin Energy Research, Corp.

References

[1] Rothman, D., *Geophysics*, **50**(12), 2784-2796 (1985).

[2] DuBose, J., *Geophysics*, **58**(3), 399-407 (1993).

[3] Yilmaz, O., *Seismic Data Processing*, Society of Exploration Geophysicists, 1988.

[4] Barhen, J., V. Protopopescu, and D. Reister, *Science*, **276**, 1094-1097 (1997).

[5] Ratschek, H. and J. Rokne,*New Computer Methods for Global Optimization*, Ellis Horwood, 1988.

[6] Törn, A. and A. Zilinskas, *Global Optimization*, Springer-Verlag, 1989.

[7] Horst, R., P. M. Pardalos, and N. V. Thoai, Kluwer Academin Publishers, 1996.

[8] Floudas, C. A. and P. M. Pardalos, *State of the Art in Global Optimization: Computational Methods and Applications*, Kluwer Academic Publishers, 1996.

[9] Aluffi-Pentini, F., V. Parisi, and F. Zirilli, *Journal of Optimization Theory and Applications*, **47**, 1–15 (1985).

[10] Kan, A. H. and G. T. Timmer, in *Numerical Optimization*, eds. P. T. Boggs *et al*, pp. 245–262 SIAM, 1985.

[11] Levy, A. V. and A. Montalvo, *SIAM Journal on Scientific and Statistical Computing*, **6**, 15–29 (1985).

[12] Szu, H. and R. Hartley, *Physics Letters*, **A 122**, 157–162 (1987).

[13] Styblinski, M. A. and T. S. Tang, *Neural Networks*, **3**, 467–483 (1990).

[14] Ammar, H. and Y. Cherruault, *Math. Comp. and Modeling*, **18**, 17–21 (1993).

[15] Cetin, B., J. Barhen, and J. Burdick, *J. Opt. Theory and Appl.*, **77**, 97–126 (1993).

[16] Cvijovic, D. and J. Klinowski, *Science*, **267**, 664-666 (1995).

[17] Yong, L., K. Lishan, and D. J. Evans, *Parallel Computing*, **21**, 389-400 (1995).

[18] Barhen, J. and V. Protopopescu, in *State of the Art in Global Optimization*, C.A. Floudas and P.M. Pardalos eds., pp. 163-180, Kluwer Academic Press, 1996.

[19] Schneider, J., and P. Schuchhardt, *Biol. Cybern*, **74**, 203-207 (1996).

ALGORITHMS FOR FUSION OF MULTIPLE SENSORS HAVING UNKNOWN ERROR DISTRIBUTIONS

Nageswara S. V. Rao

Oak Ridge National Laboratory
Oak Ridge, TN 37831-6364

ABSTRACT

The sensor S_i, $i = 1, 2 \ldots, N$, of a multiple sensor system outputs $Y^{(i)} \in \Re$ in response to input $X \in \Re$ according to an unknown probability distribution $P_{Y^{(i)}|X}$. For a fusion rule $f : \Re^N \mapsto \Re$ the expected square error is given by $I(f) = E[(X - f(Y))^2]$. When only a training sample is available, f^* that minimizes $I(.)$ over a family of functions $\mathcal{F}$ cannot be computed since the underlying distributions are unknown. We consider methods to compute an estimator $\hat{f}$ such that $I(\hat{f}) - I(f^*) < \epsilon$ with probability $1 - \delta$, for any $\epsilon > 0$ and $0 < \delta < 1$. We present a general method based on the scale-sensitive dimension of $\mathcal{F}$. We then review two computational methods based on the Nadaraya-Watson estimator, and the finite dimensional vector spaces.

INTRODUCTION

In a number of engineering applications, there has been an increased need for solving difficult sensor fusion problems. The fuser is very critical in these problems since an inappropriate fuser can render the system worse than the worst individual sensor. Additionally, the fuser must be efficiently computable in order to be of practical use. Early sensor fusion methods require either independence of sensor errors or closed-form analytical expressions for error densities. Under the first condition a general majority rule suffices, while under the second condition the Bayesian methods can be used to design the fuser. Furthermore, there have been only a limited number of studies on the computational aspects of sensor fusion problems. In practical applications, however, independence can seldom be assured and, in fact, may not be satisfied. The fusion rules are typically obtained from a specific function class which can be chosen to make the estimation problem simple, while the underlying distributions cannot be so chosen since they depend on the sensors. As a result, the problem of obtaining the probability densities required by the Bayesian methods can be more difficult than the fusion problem itself (in an information theoretic sense). When sensors are available for operation, one can collect "empirical data" by sensing objects with known parameters. Such data can then be exploited to solve the fusion rule estimation problems

under very general conditions as shown in [1]. In this paper, we generalize the results of [1] in terms of function classes by using the scale-sensitive dimension [2], and also by removing the requirement of sensor error densities.

Consider a system of N sensors such that corresponding to input $X \in \Re$, the sensor S_i, $i = 1, 2, \ldots, N$, outputs $Y^{(i)} \in \Re$ according to an *unknown* distribution $P_{Y^{(i)}|X}$. An independently and identically distributed (iid) training n-sample $(X_1, Y_1), (X_2, Y_2), \ldots, (X_n, Y_n)$ is given where $Y_i = (Y_i^{(1)}, Y_i^{(2)}, \ldots, Y_i^{(N)})$ and $Y_i^{(j)}$ is the output of S_j in response to input X_i. We consider the *expected square error*

$$I(f) = \int [X - f(Y))]^2 dP_{Y,X}, \tag{1.1}$$

where $Y = (Y^{(1)}, Y^{(2)}, \ldots, Y^{(N)})$, to be minimized over a family of fusion rules $\mathcal{F}$, based on the given n-sample. In general, the *expected best* fusion rule f^* that minimizes $I(.)$ over $\mathcal{F}$ cannot be computed since the underlying distributions are unknown. We consider conditions under which, based on a sufficiently large sample, an estimator $\hat{f}$ can be computed such that

$$P[I(\hat{f}) - I(f^*) > \epsilon] < \delta, \tag{1.2}$$

where $\epsilon > 0$ and $0 < \delta < 1$. Informally, Eq. (1.2) states that the "error" of $\hat{f}$ is within ϵ of the optimal error (of f^*) with arbitrary high probability $1 - \delta$, *irrespective* of the underlying sensor distributions.

The sensor fusion problem (1.1) is solved under the criteria (1.2) in [1] using empirical estimation methods of Vapnik [3] when $\mathcal{F}$ has a finite capacity. The computational problems associated with this approach are intractable even for simple function classes. Under additional conditions, the stochastic approximation algorithms are shown to solve this problem [4], but, these conditions are hard to verify in practical cases. Sample size estimates to ensure the criterion (1.2) using feedforward sigmoidal neural networks are derived in [5] based on three different properties. A polynomial-time solution is obtained in [6] using the classical Nadaraya-Watson estimator when: (a) the densities corresponding to $P_{Y^{(i)}|X}$ exist and are smooth, and (b) the function class itself is smooth. When the function class $\mathcal{F}$ forms a finite dimensional vector space, finite sample results as well as polynomial-time computation are obtained in [7]. A review of the last two methods will be presented in this paper.

PRELIMINARIES

Let S be a set equipped with a pseudometric ν. The *covering number* $N(\epsilon, \nu, S)$ is defined as the smallest number of closed balls of radius ϵ, and centers in S, whose union covers S. Let d_∞^m be a specific pseudometric defined on $[0,1]^m$ such that for $a, b \in [0,1]^m$, we have $d_\infty^m(a,b) = \max_{1 \le i \le m} |a_i - b_i|$ where $a = (a_1, a_2, \ldots, a_m)$ and $b = (b_1, b_2, \ldots, b_m)$.

Let $\mathcal{F}$ be a class of $[0,1]$-valued functions on some domain set D and let ρ be a positive real number. We say that $\mathcal{F}$ *P_ρ-shatters* a set $A \subseteq D$ if there exists a function $s : A \mapsto [0,1]$ such that for every $E \subseteq A$ there exists some $f_E \in \mathcal{F}$ satisfying: for every $x \in A - E$, $f_E(x) \le s(x) - \rho$, and for every $x \in E$, $f_E(x) \ge s(x) + \rho$. Let the P_ρ-dimension of $\mathcal{F}$, denoted by P_ρ-dim $(\mathcal{F})$, be the maximal cardinality of a set $A \subseteq D$ that is P_ρ-shattered by $\mathcal{F}$.

Let Q denote the unit cube $[0,1]^N$ and $\mathcal{C}(Q)$ denote the set of all continuous functions defined on Q. The modulus of smoothness of $f \in \mathcal{C}(Q)$ is defined as

$$\omega_\infty(f;r) = \sup_{\|y-z\|_\infty < r,\ y,z \in Q} |f(y) - f(z)|$$

where $\| y - z \|_\infty = \max_{i=1}^{M} |y_i - z_i|$. For $m = 0, 1, \ldots$, let Q_m denote a family of diadic cubes (Haar system) such that $Q = \bigcup_{J \in Q_m} J$, $J \cap J' = \emptyset$ for $J \neq J'$, and the N-dimensional volume of J, denoted by $|J|$, is 2^{-Nm}. Let $1_J(y)$ denote the indicator function of $J \in Q_m$: $1_J(y) = 1$ if $y \in J$, and $1_J(y) = 0$ otherwise. For given m, we define the map P_m on $\mathcal{C}(Q)$ as follows: for $f \in \mathcal{C}(Q)$, we have $P_m(f) = P_m f$ defined by $P_m f(y) = \frac{1}{|J|} \int_J f(z)dz$ for $y \in J$ and $J \in Q_m$ [8]. Note that $P_m f : Q \mapsto [0,1]$ is a discontinuous (in general) function which takes constant values on each $J \in Q_m$. The *Haar kernel* is given by

$$P_m(y,z) = \frac{1}{|J|} \sum_{J \in Q_m} 1_J(y) 1_J(z)$$

for $y, z \in Q$.

GENERAL SOLUTIONS FOR FUSER DESIGN

In this section, we consider conditions for solving the general sensor fusion problem in (1.1) under the criterion (1.2). Consider the *empirical error* of $f \in \mathcal{F}$ given by

$$I_{emp}(f) = \frac{1}{l} \sum_{i=1}^{n} [X_i - f(Y_i)]^2 \tag{3.1}$$

based on the sample $(X_1, Y_1), (X_2, Y_2), \ldots, (X_n, Y_n)$. To approximate $f^* \in \mathcal{F}$ that minimizes the expected error in (1.1), we minimize instead the empirical error in (3.1) to obtain the *empirical best* fusion rule $\hat{f}$. The following theorem presents an estimate of the sample size to ensure the condition (1.2) when $\mathcal{F}$ has finite scale-sensitive dimension [2] and $X \in [0,1]$.

Theorem 1 *Let f^* and $\hat{f}$ denotes the expected best and empirical best fusion rules chosen from a function class $\mathcal{F}$ with range $[0,1]$. Given an iid sample of size*

$$\frac{5040}{\epsilon^2} \max \left\{ d \ln^2 \frac{50d}{\epsilon}, \ln \frac{48}{\delta} \right\}$$

where $d = P_{\epsilon/4}$-dim $(\mathcal{F})$, we have $P[I(\hat{f}) - I(f^*) > \epsilon] < \delta$.

Proof: From Vapnik [9] we have $P\left\{ I(\hat{f}) - I(f^*) > \epsilon \right\} < P\left\{ \sup_{g \in \mathcal{G}} |I_{emp}(f) - I(f)| > \epsilon/2 \right\}$, where $\mathcal{G} = \{(x - f(y))^2 : f \in \mathcal{F}\}$.

Let $\mathcal{G}_{2n} = \{g(X_1, Y_1), g(X_2, Y_2), \ldots, g(X_{2n}, Y_{2n}) | g \in \mathcal{G}\} \subseteq [0,1]^{2n}$, based on an iid sample of size $2n$. From Lemma 3.3 and 3.4 of [2], we have.

$$\begin{aligned} P\left\{\sup_{g\in\mathcal{G}} |I_{emp}(f) - I(f)| > \epsilon/2\right\} &\leq 24nE_{X^{2n}}[N(\epsilon/12, d_\infty^{2n}, \mathcal{G}_{2n})]e^{-\epsilon^2 n/144} \\ &\leq 48n\left(\frac{4608n}{\epsilon^2}\right)^{d\log_2(96en/(d\epsilon))} e^{-\epsilon^2 n/144} \end{aligned}$$

where $d = P_{\epsilon/4}\text{-dim}\ (\mathcal{F})$. By equating the right hand side to δ, we obtain our sample size estimate. The derivation closely follows that of Theorem 1 of [2]. □

The result of Theorem 1 is more general than that in [1] which is based on the capacity of $\mathcal{F}$ [9] in that finiteness of capacity implies that of scale-sensitive dimension but not vice versa. This theorem can be generalized in a straight forward manner to handle the cases: (a) $Y^{(i)}$ is a multi-dimensional vector from $\Re^d$, and/or (b) $X \in [0, \tau]$, $\tau > 0$. The cost function can also be generalized to Lipschitz cost functions with an appropriate change in the sample size (see [10]).

The sample bound of Theorem 1 is based on uniform convergence of empirical means to their expectations for function classes, which are available from the empirical process theory [11, 12] and its applications to machine learning [3, 13]. Results of this kind are available based on a number of characterizations of $\mathcal{F}$ such as pseudo-dimension [11], fat VC-dimension [14], scale-sensitive dimension [2], graph dimension [15], and Euclidean parameters [12], which can be used to obtain sample size estimates along the lines of Theorem 1. Finiteness of these parameters is only sufficient for the "learnability" of bounded functions, while that of the scale sensitive dimension is both necessary and sufficient [2]. Moreover, the latter is only such deterministic quantity known to us, while other similar quantities are based on expected capacity or entropy [3].

A solution based on Theorem 1 simply requires that $\hat{f}$ minimize the empirical error, and does not specify methods to *compute* it. The problem of computing $\hat{f}$ in this general framework is intractable; for example in the special case that $\mathcal{F}$ is set of feedforward neural networks with threshold hidden units, this problem is NP-complete even for simple architectures. In the next sections, we consider more restrictive cases where $\mathcal{F}$ is chosen to be a special class to make the computational problems easier.

FUSERS BASED ON NADARAYA-WATSON ESTIMATOR

We now present a polynomial-time (in sample size n) computable estimator which guarantees the criterion (1.2) under additional smoothness conditions. Given an n-sample, the Nadaraya-Watson estimator based on Haar kernels is defined by

$$\hat{f}_{m,n}(y) = \frac{\sum_{j=1}^{n} X_j P_m(y, Y_j)}{\sum_{j=1}^{n} P_m(y, Y_j)} = \frac{\sum_{Y_j \in J} X_j}{\sum_{Y_j \in J} 1_J(Y_j)}$$

for $y \in J$ [16] (see also Engel [17]). The second expression indicates that $\hat{f}_{m,n}(y)$ is the mean of the function values corresponding to Y_j's in J that contains y. This property is the key to efficient computation of the estimate [18].

Theorem 2 [6] *Consider a family of functions $\mathcal{F} \subseteq \mathcal{C}(Q)$ with range $[0,1]$ such that $\omega_\infty(f;r) \leq kr$ for some $0 < k < \infty$. We assume that: (i) there exists a family of densities $\mathcal{P} \subseteq \mathcal{C}(Q)$; (ii) for each $p \in \mathcal{P}$, $\omega_\infty(p;r) \leq kr$; and (iii) there exists $\mu > 0$ such that for each $p \in \mathcal{P}$, $p(y) > \mu$ for all $y \in [0,1]^N$. Suppose that the sample size, n, is larger than*

$$\frac{2^{2m+4}}{\epsilon_1^2}\left[\left(\frac{k2^m}{\epsilon_1}\left[\left(\frac{k2^m}{\epsilon_1}-1\right)^{N-1}+1\right]+m\right)\ln\left(2^{m+1}k/\epsilon_1\right)+\ln\left(\frac{2^{2m+6}}{(\delta-\lambda)\epsilon_1^4}\right)\right]$$

where $\epsilon_1 = \epsilon(\mu-\epsilon)/4$, $0 < \beta < \frac{N}{2(N+1)}$, $m = \lceil \frac{\log n\beta}{N} \rceil$ and $\lambda = b\left(\frac{2}{\epsilon}\right)^{1/N+1-1/2\beta} + b\left(\frac{2}{\epsilon_1}\right)^{1/N+1-1/2\beta}$. Then for any $f \in \mathcal{F}$, we have $P\left[I(\hat{f}_{m,n}) - I(f^*)| > \epsilon\right] < \delta$.

The value of $\hat{f}_{m,n}(y)$ can be computed in $O((\log n)^N)$ time after a preprocessing step in $O(n(\log n)^{N-1})$ time (see [18]). The smoothness conditions required in Theorem 2 are not very easy to verify in practice. However, this estimators is found to perform well in a number of applications including those that do not have smoothness properties.

VECTOR SPACE METHODS

We now consider the case when $\mathcal{F}$ forms a finite dimensional vector space. The advantages of vector space methods over the existing methods are three-fold: (a) the sample size estimate is a simple function of the dimensionality of $\mathcal{F}$, (b) the estimate can be easily computed by well-known least square methods in polynomial time, and (c) no smoothness conditions are required on the functions or distributions.

Theorem 3 [7] *Let f^* and $\hat{f}$ denote the expected best and empirical best fusion functions chosen from a vector space $\mathcal{F}$ of dimension d_V and range $[0,1]$. Given an iid sample of size*

$$\frac{512}{\epsilon^2}\left[d_V \ln\left(\frac{64e}{\epsilon}+\ln\frac{64e}{\epsilon}\right)+\ln(8/\delta)\right],$$

we have $P[I(\hat{f}) - I(f^*) > \epsilon] < \delta$.

Let $\{f_1, f_2, \ldots, f_{d_V}\}$ be a basis of $\mathcal{F}$ such that $f \in \mathcal{F}$ can be written as $f(y) = \sum_{i=1}^{d_V} a_i f_i(y)$ for $a_i \in \Re$. Then consider $\hat{f} = \sum_{i=1}^{d_V} \hat{a}_i f_i(y)$ such that $\hat{a} = (\hat{a}_1, \hat{a}_2, \ldots, \hat{a}_{d_V})$ minimizes the cost expressed as (with abuse of notation)

$$I_{emp}(a) = \frac{1}{n}\sum_{k=1}^{n}\left(X_k - \sum_{i=1}^{d_V} a_i f_i(Y_k)\right)^2,$$

where $a = (a_1, a_2, \ldots, a_{d_V})$. Now $I_{emp}(a)$ can be written in the quadratic form $a^T C a + a^T D$, where $C = [c_{ij}]$ is a positive definite symmetric matrix, and D is a vector. This problem can be solved in polynomial-time using quadratic programming methods [19].

The potential functions of Aizerman *et al.* [20], where $f_i(y)$ is of the form $exp((y-\alpha)^2/\beta)$ for suitably chosen constants α and β, constitute an example of the vector space methods. Note that the above sample size is valid only for the method that minimizes $I_{emp}(.)$ and is not valid for the original incremental algorithm of the potential functions.

The two-layer sigmoidal networks of Kurkova [21], where the unknown weights are only in the output layer, constitute another example for the vector space methods. The specific form of these networks enables us to express each network in the form $\sum_{k=1}^{d_V} a_i \eta_i(y)$ where $\eta_i(.)$'s are universal.

APPLICATION

We consider the problem of recognizing a door (an opening) wide enough for a mobile robot to move through. The mobile robot (TRC Labmate) is equipped with an array of four ultrasonic and four infrared Boolean sensors on each of four sides as shown in Figure 1. The sensors are periodically polled while the robot is in motion. This example deals with only the problem of detecting a wide enough door when the sensor array of any side is facing it. The ultrasonic sensors return a measurement corresponding to distance to an object within a certain cone as illustrated in Figure 1. The infrared sensors return Boolean value based on the light reflected by an object in the line-of-sight of the sensor; white smooth objects are detected due to high reflectivity, while objects with black or rough surface are generally not detected. Both ultrasonic and infrared sensors are unreliable. It is very difficult to derive accurate probabilistic models for these sensors since it requires a detailed knowledge of the physics and engineering of the device as well as a priori statistical information. Thus a Bayesian solution to this problem is very hard to implement. We employ the Nadaraya-Watson estimator to derive a non-linear relationship between the width of the door and the sensor readings. Here the training sample is generated by actually recording the measurements while the sensor system is facing the door. Positive examples are generated if the door is wide enough for the robot, and the sensory system is facing the door. Negative examples are generated when the door is not wide enough or the sensory system is not correctly facing a door (wide enough or not). The robot is manually located in various positions to generate the data. Consider the sensor array of a particular side of the mobile robot. Here Y_1, Y_2, Y_3, Y_4 correspond to the normalized distance measurements from the four ultrasonic sensors, and Y_5, Y_6, Y_7, Y_8 correspond to the Boolean measurements of the infrared sensors. X is 1 if the sensor system is correctly facing a wide enough door, and is 0 otherwise. The training data included 6 positive examples and 12 negative examples. The test data included 3 positive examples and 7 negative examples. The Nadaraya-Watson estimator predicted the correct output in all examples of test data.

CONCLUSIONS

We presented recent results on a general sensor fusion problem, where the underlying sensor error distributions are not known, but a sample is available. We presented a general method for obtaining a fusion rule based on scale-sensitive dimension of the function class.

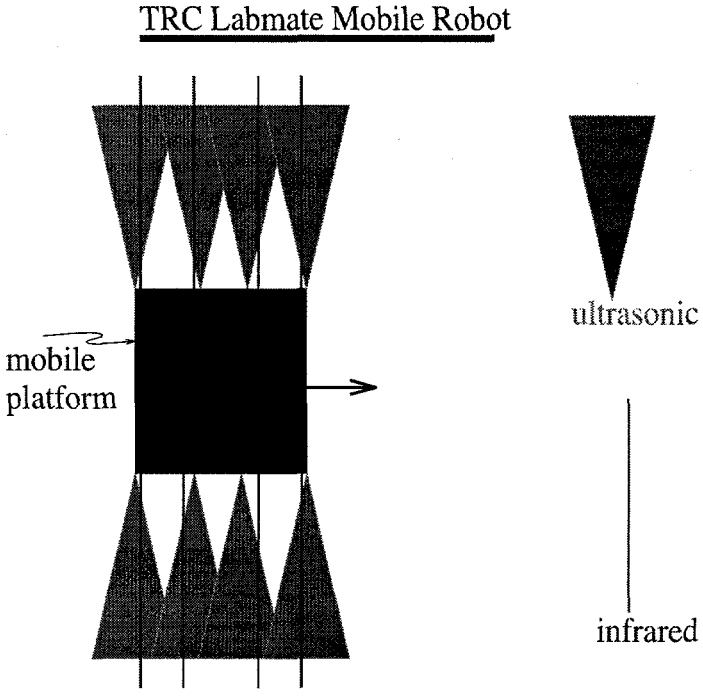


Figure 1: Schematic of sensory system (only the side sensor arrays are shown for simplicity).

Two computationally viable methods are reviewed based on the Nadaraya-Watson estimator, and the finite dimensional vector spaces.

Several computational issues of the fusion rule estimation are open problems. It would be interesting to obtain necessary and sufficient conditions under which polynomial-time algorithms can be used to solve the fusion rule estimation problem under the criterion (1.2). Also, conditions under which the composite system is "significantly" better than best sensor would be extremely useful. Finally, lower bound estimates for various sample sizes will be very important in judging the optimality of sample size estimates.

ACKNOWLEDGEMENTS

This research is sponsored by the Engineering Research Program of the Office of Basic Energy Sciences, of the U.S. Department of Energy, under Contract No. DE-AC05-96OR22464 with Lockheed Martin Energy Research Corp., and by the Office of Naval Research under order N00014-96-F-0415.

REFERENCES

[1] N. S. V. Rao. Fusion methods for multiple sensor systems with unknown error densities. *Journal of Franklin Institute*, 331B(5):509–530, 1994.

[2] N. Alon, S. Ben-David, N. Cesa-Bianchi, and D. Hausler. Scale-sensitive dimensions, uniform convergence, and learnability. In *Proc. of 1993 IEEE Symp. on Foundations of Computer Science*, 1993.

[3] V. N. Vapnik. *The Nature of Statistical Learning Theory.* Springer-Verlag, New York, 1995.

[4] N. S. V. Rao. Fusion rule estimation in multiple sensor systems using training. In H. Bunke, T. Kanade, and H. Noltemeier, editors, *Modelling and Planning for Sensor Based Intelligent Robot Systems*, pages 179–190. World Scientific Pub., 1995.

[5] N. S. V. Rao. Fusion methods in multiple sensor systems using feedforward neural networks. *Intelligent Automation and Soft Computing*, 1996. submitted.

[6] N. S. V. Rao. Nadaraya-Watson estimator for sensor fusion. *Optical Engineering*, 36(3):642–647, 1997.

[7] N. S. V. Rao. Fusion rule estimation using vector space methods. In *Proceedings of SPIE Conference on Sensor Fusion: Architecture and Applications*. 1997.

[8] Z. Ciesielski. Haar system and nonparametric density estimation in several variables. *Probability and Mathematical Statistics*, 9:1–11, 1988.

[9] V. Vapnik. *Estimation of Dependences Based on Empirical Data*. Springer-Verlag, New York, 1982.

[10] N. S. V. Rao and V. Protopopescu. Function estimation by feedforward sigmoidal networks with bounded weights. 1997. manuscript, submitted for publication.

[11] D. Pollard. *Empirical Processes: Theory and Applications*. Institute of Mathematical Statistics, Haywood, California, 1990.

[12] M. Talagrand. Sharper bounds for Gaussian and empirical processes. *Annals of Probability*, 22(1):28–76, 1994.

[13] D. Haussler. Decision theoretic generalizations of the PAC model for neural net and other learning applications. *Information and Computation*, 100:78–150, 1992.

[14] M. J. Kearns and R. E. Schapire. Efficient distribution-free learning of probabilistic concepts. *Journal Computer and Systems Sciences*, 48(3):464–, 1994.

[15] R. Dudley. Universal Donsker classes and metric entropy. *Annals of Probability*, 15:1306–1326, 1987.

[16] B. L. S. Prakasa Rao. *Nonparametric Functional Estimation*. Academic Press, New York, 1983.

[17] J. Engel. A simple wavelet approach to nonparametric regression from recursive partitioning schemes. *Journal of Multivariate Analysis*, 49:242–254, 1994.

[18] N. S. V. Rao and V. Protopopescu. On PAC learning of functions with smoothness properties using feedforward sigmoidal networks. *Proceedings of the IEEE*, 84(10):1562–1569, 1996.

[19] S. A. Vavasis. *Nonlinear Optimization*. Oxford University Press, New York, 1991.

[20] M. A. Aizerman, E. M. Braverman, and L. I. Rozonoer. *Extrapolative problems in automatic control and method of potential functions*, volume 87 of *American Mathematical Society Translations*, pages 281–303. 1970.

[21] V. Kurkova. Kolmogorov's theorem and multilayer neural networks. *Neural Networks*, 5:501–506, 1992.

AN ALGORITHM FOR NOISY IMAGE SEGMENTATION

(Extended Abstract)

Ying Xu, Victor Olman, and Edward C. Uberbacher
Computer Science and Mathematics Division
Oak Ridge National Laboratory
Oak Ridge, TN 37831-6364

Abstract

This paper presents a segmentation algorithm for gray-level images and addresses issues related to its performance on noisy images. It formulates an image segmentation problem as a partition of an image into (arbitrarily-shaped) connected regions to minimize the sum of gray-level variations over all partitioned regions, under the constraints that (1) each partitioned region has at least a specified number of pixels, and (2) two adjacent regions have significantly different "average" gray-levels. To overcome the computational difficulty of directly solving this problem, a minimum spanning tree representation of a gray-level image has been developed. With this tree representation, an image segmentation problem is effectively reduced to a tree partitioning problem, which can be solved efficiently. To evaluate the algorithm, we have studied how noise affects the performance of the algorithm. Two types of noise, transmission noise and Gaussian additive noise, are considered, and their effects on both phases of the algorithm, construction of a tree representation and partition of a tree, are studied. Evaluation results have shown that the algorithm is stable and robust in the presence of these types of noise.

1 Introduction

Image segmentation is one of the most fundamental problems in low-level image processing. The problem is to partition (segment) an image into connected regions of similar textures or similar colors/gray-levels, with adjacent regions having significant dissimilarity. Many algorithms have been proposed to solve this problem (see surveys [1, 2]). Most of these algorithms fit into two categories: (1) boundary detection-based approaches, which partition an image by discovering closed boundary contours, and (2) region clustering-based approaches, which group "similar" neighboring pixels into clusters. Rigorous mathematical solutions to the image segmentation problems are generally difficult to achieve due to their (intrinsic) computational complexity. Hence many researchers have exploited either probabilistic/stochastic methods, which guarantee only asymptotic results, or heuristic methods while sacrificing the mathematical rigor.

In this paper, we present an efficient region-based segmentation algorithm. We formulate an image segmentation problem as a partition of an image into a number (not predetermined) of arbitrarily-shaped connected regions to minimize the sum of gray-level variations over all partitioned regions under the constraints that (1) each partitioned region has at least a specified

number of pixels, and (2) two adjacent regions have significantly different "average" gray-levels. To overcome the computational difficulty of directly solving this problem, we have developed a minimum spanning tree representation of an image. The minimum spanning tree representation, though simple, captures the essential information of an image for the purpose of segmentation, and it facilitates a fast segmentation algorithm. The technical contribution of our approach includes (1) a new spanning tree representation of an image that captures all the key information for the purpose of segmentation, and (2) a fast and mathematically rigorous tree partitioning algorithm.

To evaluate the algorithm, we have studied how two types of noise, transmission noise and Gaussian additive noise, affect the performance of the algorithm. We have shown, both analytically and experimentally, that (1) both types of noise have very little effect on the minimum spanning tree construction algorithm, i.e., the property that an originally homogeneous region corresponds to one subtree of the spanning tree will generally not be affected by noise; (2) transmission noise, in general, has less effect on the performance of our tree-partitioning algorithm than Gaussian additive noise does.

2 Image Segmentation: the problem formulation

Consider a gray-level image I. Each pixel x of I has a gray level $\mathcal{G}(x) \in [0, K]$. An image *segmentation problem* can be naturally formulated as follows: find a partition $\{I_1, ..., I_k\}$ of I with each I_i being a connected region of I, such that

$$\text{minimize} \qquad \sum_{i=1}^{k} \sum_{x_i^j \in I_i} (average(I_i) - \mathcal{G}(x_i^j))^2$$

$$\text{subject to:} \quad (1) \quad \|I_i\| \geq L, \text{ for each } I_i,$$
$$(2) \quad |average(I_i) - average(I_{i'})| \geq D, \text{ for all adjacent } I_i \text{ and } I_{i'}.$$

where $\|\cdot\|$ denotes the cardinality of a set, $average(I_i)$ denotes the average gray-level of region I_i, and L and D are two (application-dependent) parameters.

Though this formulation captures the intuition of segmenting an image it is computationally difficult to solve due to two reasons: (1) segmenting a 2-D object to optimize some non-trivial function is always difficult, and (2) explicit calculation of averages implicitly requires to consider all the possible partitions. Two strategies have been developed to overcome these difficulties: a tree representation of an image, and an approximation scheme to avoid explicit calculation of averages.

2.1 Spanning tree representation of an image

For a given image I, we define a weighted (undirected) planar graph $G(I) = (V, E)$ as follows: The vertex set $V = \{$ all pixels of $I \}$ and the edge set $E = \{(u, v) | u, v \in V \text{ and } distance(u, v) \leq \sqrt{2} \}$, with $distance(u, v)$ representing the Euclidean distance in terms of the coordinates of the image array; Each edge $(u, v) \in E$ has a weight $w(u, v) = |\mathcal{G}(u) - \mathcal{G}(v)|$.

A *spanning tree* T of a connected graph $G(I)$ (note that $G(I)$ is connected) is a connected subgraph of $G(I)$ such that (1) T contains every vertex of $G(I)$, and (2) T does not contain cycles. A *minimum* spanning tree is a spanning tree with a minimum total weight.

A minimum spanning tree of a weighted graph can be found using greedy methods, like in the classical Kruskal's algorithm [3]: the initial solution is a singleton set containing an edge with

the smallest weight, and then the current partial solution is repeatedly expanded by adding the next smallest weighted edge (from the unconsidered edges) under the constraint that no cycles are formed until no more edges can be added. For the above defined planar graph $G(I)$, a minimum spanning tree can be constructed in $O(\|V\| \log(\|V\|))$ time and in $O(\|V\|)$ space.

A key property of a minimum spanning tree representation obtained by Kruskal's algorithm is that *pixels of a homogeneous region are connected in the tree structure only through pixels of this region*, i.e., pixels of a homogeneous region form a (connected) subtree of the minimum spanning tree. The following theorem formalizes this statement.

Consider an object A in a given image I. Let $G(I)$ be the planar graph representation of I and T be its minimum spanning tree obtained by Kruskal's algorithm. A is called *T-connected* if every pair of pixels of A are connected in T only through pixels of A. We use $G(A)$ to denote the subgraph of $G(I)$ induced by the pixels of A. A set of edges C of $G(A)$ is called a *cutset* if deleting C divides $G(A)$ into two unconnected parts.

Theorem 1 *A is not T-connected if and only if there exists a cutset C of $G(A)$ and a path P in $G(I)$ that has its two end vertices on two sides of the cut of $G(A)$ and has its remaining vertices outside of $G(A)$ such that every edge of P has smaller[1] weight than every edge of C.* □

2.2 An approximation scheme

To formulate the image segmentation problem in a natural and intuitive way, we have explicitly used the average gray-levels of a region in the problem formulation, which makes the computation difficult. This subsection presents an approximation scheme to avoid the explicit calculation of averages.

Consider the following formulation of an image segmentation problem. Given an image I and two parameters D and L, find a partition $\{I_1, ..., I_k\}$ of I with each I_i being a connected region of I, and a $g_i \in \mathcal{R}$ (real value) for each I_i, such that

$$\text{minimize} \qquad \sum_{i=1}^{k} \sum_{x_i^j \in I_i} (g_i - \mathcal{G}(x_i^j))^2$$

$$\text{subject to:} \quad (1) \quad \|I_i\| \geq L, \text{ for each } I_i,$$
$$(2) \quad |g_i - g_{i'}| \geq D, \text{ for all adjacent } I_i \text{ and } I_{i'}.$$

The relationship between this formulation, which does not involve explicit calculation of averages, and the original one can be intuitively described as follows: if a solution to this formulation is stable around the given parameter D, then the two formulations are equivalent. This can be stated more rigorously as in the following theorem. Let

$$F(k, I, g) = \sum_{i=1}^{k} \sum_{x_i^j \in I_i} (g_i - \mathcal{G}(x_i^j))^2,$$

and

$$R(D, L) = \{(k, I, g) | \text{ which satisfies constraints (1) and (2)}\},$$

[1] We ignore the case of equality for the simplicity of discussion.

where $I = \bigcup_{i=1}^{k} I_i$ and $g = (g_1, ..., g_k)$. Hence the above formulation can be rewritten as

$$\min_{k,I,g}\{F(k,I,g)|(k,I,g) \in R(D,L)\}.$$

Theorem 2 *For the given parameters D and L, if there is an $\epsilon > 0$ such that*

$$\min_{k,I,g}\{F(k,I,g)|(k,I,g) \in R(d,L)\} = F_0 \tag{1}$$

for some constant F_0, for all $d \in [D, D+\epsilon]$, then any minimum solution $I^* = \{I_1^*, ..., I_k^*\}$ and $g^* = \{g_1^*, ..., g_k^*\}$ to $\min_{k,I,g}\{F(k,I,g)|(k,I,g) \in R(D+\epsilon,L)\}$ has $g_i^* = average(I_i^*)$, for all $i \in [1,k]$. □

Note that g_i's, as defined above, are real values $\in [0, \mathcal{K}]$. To facilitate a fast algorithm, we restrict g_i's to integer values $\in [0, \mathcal{K}]$. Now we can give the tree-based image segmentation problem as follows. Given a minimum spanning tree representation T of an image and two parameters D and L, find a partition $\{T_1, ..., T_k\}$ of T with each T_i being a connected subtree of T, and an integer $g_i \in [0, \mathcal{K}]$ for each T_i, such that

$$\begin{array}{lll} \text{minimize} & \sum_{i=1}^{k} \sum_{x_i^j \in T_i} (g_i - \mathcal{G}(x_i^j))^2 & \\ & & (P) \\ \text{subject to:} & (1) \quad \|T_i\| \geq L, \text{ for each } T_i, & \\ & (2) \quad |g_i - g_{i'}| \geq D, \text{ for all adjacent } T_i \text{ and } T_{i'}. & \end{array}$$

To estimate how close the approximation problem is to the original problem, we have the following result:

$$\frac{E(\sum_{i=1}^{k} \|T_i\|(average(T_i) - g_i)^2)}{E(\sum_{i=1}^{k} \sum_{x_i^j \in T_i} (average(x_i^j) - \mathcal{G}(x_i^j))^2)} \leq k/N, \tag{2}$$

which indicates the minimum value of the approximation problem is fairly close to the minimum value of the original optimization problem, where $E()$ represents the mathematical expectation.

3 A Tree-based Image Segmentation Algorithm

A dynamic programming algorithm is developed to solve the optimization problem (P). The algorithm first converts the given tree into a *rooted* tree by selecting an arbitrary vertex as the root. Hence the *parent-children* relation is defined. We assume that the vertices of T are labeled consecutively from 1 to $\|T\|$ with the tree root labeled as 1. We use T^i to denote the subtree rooted at vertex i. For each tree vertex i, the dynamic programming algorithm solves a number of constraint version of the problem (P) on T^i by combining solutions to the "corresponding" problems on T^j's, with j's being i's children. It does this repeatedly in such a bottom-up fashion and stops when it reaches the tree root.

Let $score(i,k,g)$ denote the minimum value of (P) on T^i, under the additional constraint that the partitioned subtree of T^i containing i has at least k vertices and is mapped to a fixed value g, for $k \in [0, L]$ and $g \in [0, \mathcal{K}]$. These quantities can be efficiently calculated using the following lemma and can be used to construct an optimum partition of T.

Lemma 1 *(a) If $i_1, i_2, ..., i_n$ are the children of vertex i, $n \leq 8$ and $1 \leq k \leq L$, we have*

$$score(i,k,g) = \min \sum_{j=1}^{n} score(i_j, k_j, g) + (g - \mathcal{G}(i))^2,$$
$$for\ k = \sum_{j=1}^{n} k_j, k_j \geq 0, \quad when\ \|T^i\| \geq L$$
$$scores(i,k,g) = \begin{cases} \sum_{p \in \mathcal{D}(i)} (g - \mathcal{G}(p))^2, & \|T^i\| = k \\ +\infty, & \|T^i\| \neq k \end{cases}$$
$$when\ \|T^i\| < L$$

where $\mathcal{D}(i)$ is the set of all i's descendants, i is defined to be $\in \mathcal{D}(i)$ and $score(i_j, 0, g)$ is defined to be

$$\min_{|g'-g| \geq D} score(i_j, L, g').$$

(b) $\min_g score(1, L, g)$ is a minimum solution of (P), where 1 represents the tree root. □

Based on Lemma 1, we can solve the optimization problem (P) by calculating $score()$ for each tree vertex in a bottom-up fashion using the recurrence from Lemma 1(a), and stopping at the tree root.

Theorem 3 *$\min_g score(1, L, g)$ can be correctly calculated by the above algorithm in $O(\max\{(\|T\| - L), 1\}\mathcal{K}(\log(\mathcal{K}) + L^2))$ time and in $O(\|T\|L\mathcal{K})$ space.* □

4 Algorithm Evaluation on Noisy Images

Potentially noise affects the algorithm's performance in both stages of the algorithm: spanning tree construction and tree partitioning. We will show that noise has greater effects on the performance in the tree partitioning stage than in the spanning tree construction stage. In this study, we consider two types of noise: transmission noise and Gaussian additive noise.

The model for generating *transmission noise* is defined as follows: each pixel of the image has a probability $\mathcal{P}$ to keep its original gray level during transmission and the probability $1 - \mathcal{P}$ to randomly change to arbitrary gray level $\in [0, \mathcal{K}]$. *Gaussian additive noise* adds to each pixel independently a random normal value (using the floor function for real-to-integer conversion) according to a normal distribution $N(0, \sigma^2)$ censored to $[-\mathcal{K}/2, \mathcal{K}/2]$.

4.1 Effect of noise on tree representation

One basic premise for our image segmentation algorithm to be effective is that each object, given as a homogeneous region in an image, is represented as one subtree of the spanning tree representation. In the following, we show how noise affects this property. Theorem 1 provides the basic framework for such a study.

To estimate how probable the if-and-only-if condition in Theorem 1 is we have conducted the following computer simulation. The experiment is done on a 256-gray-level image I having one object A in the center of the image. I is a 256×256 image and A is a 30×30 square. The background has a uniform gray level 100 and A has a uniform gray level 150. We add transmission noise and Gaussian additive noise, respectively, to I as follows. When adding transmission noise, each pixel of I has a probability 0.3 to keep its original gray level and the probability 0.7 to

randomly and uniformly change to arbitrary gray level $\in [0, 255]$. When adding Gaussian additive noise, each pixel of I is added by a value $\lfloor \delta + 1/2 \rfloor$ (modulo 256), where δ is random number generated according to the normal distribution $N(0, \sigma^2)$ censored to [-128, 128] with $\sigma = 50$.

For each type of noise, we estimated the probability that there exist a path P connecting two pixels a and b, and a cutset C of A separating a and b such that every edge of P has smaller weight than every edge of C, where a and b are two randomly chosen pixels both of which are 5-pixels from the left boundary of A and are at least 5 pixels from the lower and upper boundaries of A, and P has at least 20 edges.

We have observed, for this particular experiment, that the probability that there exist such a P and a cutset C is very small ($< 10^{-3}$), for both types of noise. This experiment suggested that both types of noise have very little effect on the property that a homogeneous region corresponds to one subtree of the minimum spanning tree constructed by Kruskal's algorithm.

4.2 Effect of noise on tree partitioning

Though both types of noise have little effect on the property that a homogeneous region corresponds to one subtree of the spanning tree representation they could affect the tree partitioning result in a form we call *corrosions*. Consider an object A in a given image and its representing subtree T_A. With noise, T_A may contain a subtree (or subtrees) that has a (significantly) different average gray level than the rest of T_A, and contains more than enough vertices ($\geq L$) to be partitioned into a separate region. This subsection presents a comparative study on how the two types of noise affect the formation of corrosions.

Let $g(A)$ be the (uniform) gray level of A before noise is added. We compare the probabilities, P_1 and P_2, that a connected region A' of A will have its gray level changed to the same value $g(A)+k$, for any $k \neq 0$, when transmission noise and Gaussian additive noise are added, respectively. Let p_k denote the probability that one pixel of A' changes its gray level from $g(A)$ to $g(A) + k$ when Gaussian additive noise is added. For the simplicity of discussion, we assume that $g(A) = 0$, hence $k \in [1, \mathcal{K}]$. Recall $\mathcal{P}$ denotes the probability that a pixel keeps its original gray level in the presence of transmission noise. It can be shown by a simple calculation that

$$P_1 = \left(\frac{1-\mathcal{P}}{\mathcal{K}-1}\right)^n (\mathcal{K}-1), \qquad \text{and} \qquad P_2 = \sum_{k=1}^{\mathcal{K}} p_k^n,$$

where $n = \|A'\|$ (note that $P_2 = \sum_{k=1}^{\mathcal{K}} p_k^n$ is true for any type of independent noise). Theorem 4 shows the relationship between P_1 and P_2, which can be proved using Jensen's inequality [5] (page 433).

Theorem 4 *For any $\mathcal{N} \in [2, \mathcal{K}]$ and $n > 0$, when $\sum_{k=0}^{\mathcal{N}} p_k = 1$ and $p_0 = \mathcal{P}$,*

$$\sum_{k=1}^{\mathcal{N}} p_k^n \geq \left(\frac{1-\mathcal{P}}{\mathcal{N}-1}\right)^n (\mathcal{N}-1).$$

□

Theorem 4 implies that transmission noise is the least possible to form corrosions among all possible forms of noises (including Gaussian additive noise) when $\mathcal{P}$ or p_0 is fixed.

4.3 Tests on noisy images

This subsection presents a case-study on an aerial image of 202x503 pixels and with 256 gray levels, and on how noise of different types affects the performance of the segmentation algorithm. Throughout this study, the same set of parameters D and L are used. Segmentation on each image takes less than 1 CPU minute on a SPARC-20 workstation. Figure 1 gives the test examples on the image with noise added. For each figure, the image on the left is the original image with added noise and the one on the right represents the segmentation results.

Table 1 summarizes the performance of algorithm and the effect of the averaging operation on the two types of noise. Each entry of the first row represents the correlational coefficient between the original image and the image with noise, and each entry of the second row represents the correlational coefficient between the segmentation result of the original image and the segmentation of the noisy image.

Table 1: Performance summary of segmentations

	Transmission noise				**Gaussian additive noise**			
	$\mathcal{P} = 0.1$	$\mathcal{P} = 0.3$	$\mathcal{P} = 0.5$	$\mathcal{P} = 0.7$	$\sigma = 40$	$\sigma = 60$	$\sigma = 80$	$\sigma = 100$
noisy image	0.86	0.62	0.41	0.24	0.82	0.69	0.57	0.47
segmentation	0.95	0.89	0.80	0.70	0.87	0.84	0.81	0.76

Acknowledgements

This research was supported by the United States Department of Energy, under contract DE-AC05-84OR21400 with Lockheed Martin Energy Systems, Inc. The authors would like to thank Dr. Reinhold C. Mann for many helpful discussions related to the work presented in this paper.

References

[1] N. Pal and S. Pal, "A review on image segmentation techniques", *Pattern Recognition*, Vol. 26, No. 9, pp. 1277 - 1294, 1993.

[2] R. M. Haralick and L. G. Shapiro, "Image segmentation techniques", *Computer Vision, Graphics, and Image Processing*, Vol. 29, pp. 100 - 132, 1982.

[3] J. B. Jr. Kruskal, "On the shortest spanning subtree of a graph and the traveling salesman problem", *Proc. Amer. Math Soc*, Vol. 7, No. 1, pp. 48 - 50, 1956.

[4] Y. Xu and E. C. Uberbacher, "2-D Image Segmentation Using Minimum Spanning Trees", *Image and Vision Computing*, Vol. 15 pp. 47 - 57, 1997.

[5] P. J. Bickel and K. A. Doksum, *Mathematical Statistics: Basic Ideas and Selected Topics*, Holden-Day Inc., 1977.

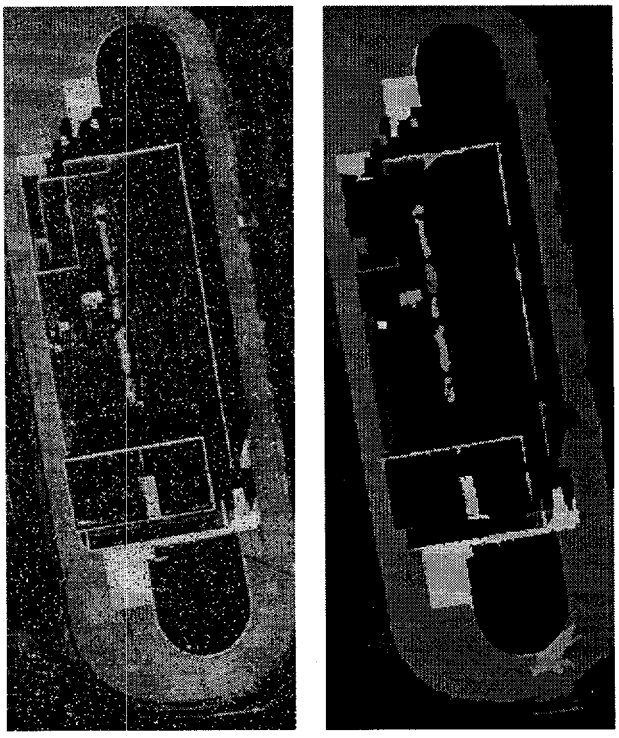


Figure 1: (a) Aerial image with added transmission noise and $\mathcal{P} = 0.1$.

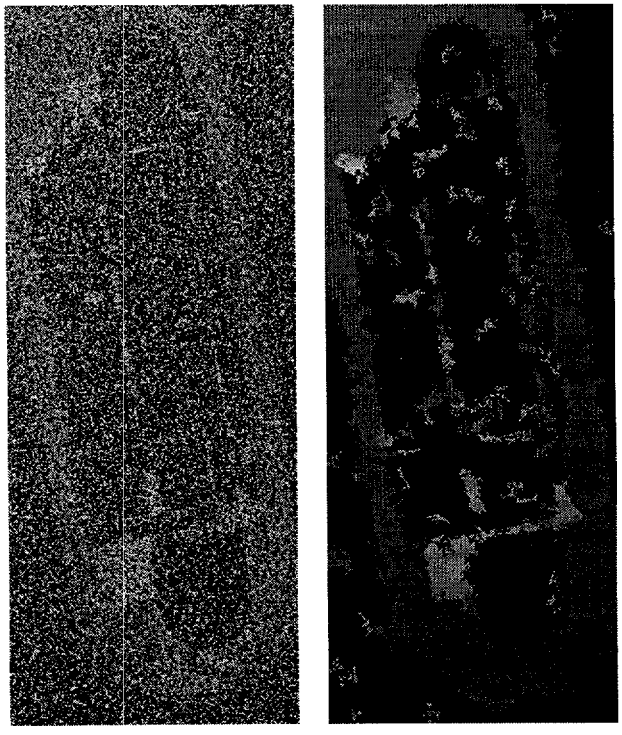


Figure 1: (b) Aerial image with added transmission noise and $\mathcal{P} = 0.7$.

ADAPTATION WITH DISTURBANCE ATTENUATION IN NONLINEAR CONTROL SYSTEMS

Tamer Başar

Coordinated Science Laboratory and the
Department of Electrical and Computer Engineering
University of Illinois
Urbana, IL 61801-2307

ABSTRACT

We present an optimization-based adaptive controller design for nonlinear systems exhibiting parametric as well as functional uncertainty. The approach involves the formulation of an appropriate cost functional that places positive weight on deviations from the achievement of desired objectives (such as tracking of a reference trajectory while the system exhibits good transient performance) and negative weight on the energy of the uncertainty. This cost functional also translates into a disturbance attenuation inequality which quantifies the effect of the presence of uncertainty on the desired objective, which in turn yields an interpretation for the optimizing control as one that optimally attenuates the disturbance, viewed as the collection of unknown parameters and unknown signals entering the system dynamics. In addition to this *disturbance attenuation* property, the controllers obtained also feature *adaptation* in the sense that they help with identification of the unknown parameters, even though this has not been set as the primary goal of the design. In spite of this adaptation/identification role, the controllers obtained are not of certainty-equivalent type, which means that the identification and the control phases of the design are not decoupled.

INTRODUCTION AND PROBLEM DESCRIPTION

We consider in this paper the problem of control of partially unknown, uncertain nonlinear systems so that the system output tracks (at least asymptotically) a given reference trajectory while all internal states remain bounded and the system exhibits acceptable transient performance. The uncertainty is due to the presence of unknown deterministic signals entering the system dynamics, and unknown (and unmeasurable) noise in the measurements. To capture the presence of all these factors that impact the overall performance of the system, and to quantify various tradeoffs that exist, we base our control design on the minimization of a carefully selected cost functional, which leads to a systematic construction of *robust adaptive controllers* that attenuate the disturbances optimally. These adaptive controllers have distinguishable identifier and control dynamics, which however are not decoupled, and hence the controllers are not certainty equivalent — in contrast to many existing designs in the literature. This "noncertainty equivalence" structure, which comes about naturally as a result of the optimization procedure, brings with it many appealing features such as robustness to unmodeled dynamics, attenuation of disturbances, and excellent transient performance.

To introduce the approach adopted in this paper in general terms, consider the n-dimensional dynamic system described by

$$\dot{x} \;=\; f(x,\theta) + G(x,\theta)u + \sigma(x)w, \quad x(0) = x_0 \tag{1}$$

where θ is a p-dimensional unknown parameter vector, x is the n-dimensional state, u is an r-dimensional control, w is a q-dimensional unknown disturbance, f, G and σ are appropriate dimensional vectors and matrices, continuous in x, and with f and G linear in θ. Let us assume for the moment that the control uses state feedback with memory, that is for some measurable function μ,

$$u(t) \;=\; \mu(t, x_{[0,t]})\,, \tag{2}$$

and that the objective is for an m-dimensional ($m \leq n$) output of the system,

$$z \;=\; h(x), \tag{3}$$

to track a given m-dimensional reference trajectory, z_r, in spite of the presence of the disturbance w, and regardless of what the true value of θ is. Hence, what is being sought is a controller that achieves the desired objective (of tracking) while attenuating the disturbances at the output of the error systems, which is the tracking error, and at the same time keeping all internal states of the system bounded. A criterion that captures this objective is now given in the following.

Let us first introduce the notation

$$\|y\|_t^2 \;:=\; \int_0^t |y(\tau)|^2 d\tau, \qquad |y(\tau)|_Q^2 \;:=\; y'(\tau)Qy(\tau)\,,$$

for any vector-valued $\mathcal{L}_2$ function y, where $'$ stands for transpose of a vector (or a matrix), the latter is the square of a weighted Euclidean norm of $y(\tau)$, where Q is a positive definite weighting matrix, and the former is the square of the $\mathcal{L}_2$ norm of the function $y(\tau)$ restricted to interval $[0,t]$. Then, consider for each $t > 0$,

$$\mathcal{I}_t(\mu) \;=\; \sup_{w,\theta,x_0} \frac{\|z - z_r\|_t^2 + \tilde{\ell}_t(x_{[0,t]}; u_{[0,t]})}{\|w\|_t^2 + |\theta - \bar{\theta}|_{Q_0}^2 + \ell_0(x_0, \theta - \bar{\theta})} \tag{4}$$

as the performance index to be minimized by the controller μ for each $t > 0$. Here

$$\tilde{\ell}_t \;:=\; \int_0^t \ell(x_{[0,\tau]}; u(\tau), \tau) d\tau$$

is a nonnegative integral cost on the state and the control, $z - z_r$ is the tracking error, $\bar{\theta}$ is an initial estimate for θ, Q_0 is a positive definite matrix, and ℓ_0 is a nonnegative cost on x_0 and $\theta - \bar{\theta}$, vanishing at $x_0 = 0$ and $\theta = \bar{\theta}$. Note that $\mathcal{I}_t$ involves a maximization operation with respect to the unknowns, w, θ and x_0, and hence characterizes a worst case scenario. By minimizing this index with respect to μ we would be minimizing the worst-case effect of w, θ and ℓ_0 on the tracking error $z - z_r$, the state x and the control u. Now let

$$\inf_{\mu} \mathcal{I}_t(\mu) \;=:\; \gamma_t^{*^2}, \quad t > 0\,,$$

and pick $\gamma > 0$ such that $\gamma > \gamma_t^*$ for all $t > 0$. Let μ_γ be a controller that achieves a better (lower) level of disturbance attenuation than γ for all $t > 0$, that is

$$\mathcal{I}_t(\mu_\gamma) \;\leq\; \gamma^2. \tag{5}$$

Then, (4) implies that for all $w \in \mathcal{L}_2[0,\infty)$, $\theta \in \mathbf{R}^p$, $x_0 \in \mathbf{R}^n$, the following dissipation inequality holds, for all $t > 0$, with $u = \mu(\cdot)$:

$$J_\gamma^t(\mu;\omega) := \|z - z_r\|_t^2 + \tilde{\ell}_t(x_{[0,t]}; u_{[0,t]}) - \gamma^2\|w\|_t^2 - \gamma^2\,|\theta - \bar{\theta}|_{Q_0}^2 - \gamma^2\,\ell_0(x_0, \theta - \bar{\theta}) \;\leq\; 0. \tag{6}$$

Denote the left hand side of this inequality for an arbitrary μ and $\omega := (w, \theta, x_0)$ by $\mathcal{I}_\gamma^t(\mu;\omega)$. Then, clearly, a μ_γ satisfying (5) can be obtained by solving the *minmax* problem:

$$\inf_{\mu} \sup_{\omega} J_\gamma^t(\mu;\omega), \qquad t > 0. \tag{7}$$

This can be viewed as a zero-sum differential game between two players [2], with the minimizer choosing μ and the maximizer ω, and the quantity in (7) is the upper value of such a game. Note that whenever this value is bounded, it has to be zero, since by picking $w = 0$, $\theta = \bar{\theta}$ and $x_0 = 0$ the maximizer can force it to be nonnegative, and on the other hand we know from inequality (5) that it cannot be positive.

Our approach to this problem is based on the recognition that the supremization part of (7) can be broken into two sequential supremizations,

$$\sup_{\omega} J_\gamma^t(\mu;\omega) \quad = \quad \sup_{\theta, x_{[0,t]}} \quad \sup_{(w_{[0,\infty)} | x_{[0,t]}, \theta)} J_\gamma^t(\mu;\omega) \tag{8}$$

where the inner supremization is over all disturbance trajectories consistent with the observed state trajectory $x_{[0,t]}$ and for a fixed value of the parameter vector θ, and the outer supremization is over all possible values of $\theta \in \mathbf{R}^p$ and all continuous state trajectories $x_{[0,t]}$. If the controller did not have access to full state measurements, but only to partial measurements, possibly corrupted with (unknown) noise, such as

$$y(t) \quad = \quad h(x,\theta) + n(x)w$$

where h and n are continuous functions of their arguments, then (8) would be replaced by the more general relationship

$$\sup_{\omega} J_\gamma^t(\mu;\omega) \quad = \quad \sup_{\theta, x_t, y_{[0,t]}} \quad \sup_{(x_0, w_{[0,\infty)} | y_{[0,t]}, \theta, x(t) = x_t)} J_\gamma^t(\mu;\omega) \tag{9}$$

with the inner and outer supremizations interpreted in a similar way (as in (8)). Now, this splitting of J_γ^t into two parts leads to a sequential design procedure that generates worst-case identifiers and robust adaptive controllers. The inner supremization (maximization) is the *worst-case identification step* which can be solved using the recently developed tool of *cost-to-come function* [4,5] which leads for some important classes of problems (as to be discussed shortly) to closed-form expressions for an identifier for the unknown parameters and an estimator for the unmeasured states. During this identification step the control $u_{[0,t]}$, generated by μ, can be regarded as an open-loop time function since it is merely a causal function of the given output waveform $y_{[0,t]}$ (or of the state $x_{[0,t]}$, if state measurements are available). After thus completing the inner supremization, we then proceed with the outer supremization of J_γ^t over all measurement waveforms $y_{[0,t]}$, terminal states x_t and parameter values θ, while structuring the control in such a way that J_γ^t remains nonpositive. This is the *control design step* which leads to a robust disturbance attenuating controller.

The problem just formulated above can also be viewed as a nonlinear H^∞ control problem with partial state information [1], by adjoining to the system dynamics (1) the natural parameter dynamics

$$\dot{\theta} = 0, \qquad \theta(0) = \theta_0 \tag{10}$$

where now θ_0 is the unknown parameter vector. This H^∞ control problem is one with partial state information even if full state measurements are available, because θ is now considered a part of the extended system dynamics, which is not measured directly. Nonlinear H^∞ control problems with partial state information are known to be inherently difficult to solve, and generally they do not admit finite-dimensional solutions [1,3]. It turns out, however, that (as shown in our recent research [9,10]) for some special subclasses of the robust adaptive control problems formulated above, finite-dimensional closed-form solutions do exist; this will be discussed also in the following sections.

Our approach to robust adaptive control as delineated above is inherently different from other existing approaches which are either Lyapunov-based or estimation-based. The former places restrictions on the selection of parameter update laws, whereas the latter (which generally makes use of the "certainty equivalent" principle) uses a wide variety of estimation/identification tools, among which are the standard gradient and least-squares algorithms. Any stabilizing controller can in fact be combined with any such identifier,

as long as the identifier guarantees certain boundedness properties independently of the controller module. This modularity feature has made estimation-based schemes more popular in linear adaptive control than their Lyapunov-based counterparts, but efforts to extend this to nonlinear systems have failed to a large extent. The source of this failure is mainly the fact that nonlinear systems exhibit different instability characteristics (than linear systems) such as finite escape. Various measures have been taken to overcome this difficulty [6,7,8], but the designs have involved certainty-equivalent controllers, which are known to have weaknesses in the framework of nonlinear systems, in particular as regards robustness against model uncertainty and external disturbance inputs.

The approach presented above, and discussed in some detail (as length restrictions permit) in the following sections, is a direct optimization-based approach that brings in robustness as an essential component of the design procedure. To carry out the details of the two-step procedure outlined above, and to obtain explicit expressions for the optimally disturbance attenuating controllers, we focus on a special, but important, class of systems where the dynamics (1) are in triangular form, the control is of dimension *one*, and the system output is the first component of the state. Such systems are called "systems in parametric strict feedback form," and one of their appealing features is that a recursive technique called *backstepping* can be used to construct an optimizing controller. Formulation of this specific problem is provided in the next section, followed by presentation of some explicit results.

SYSTEMS IN PARAMETRIC-STRICT-FEEDBACK FORM

In view of the discussion above, consider now the class of single input-single output (SISO) nonlinear systems described by (as a special case of (1), and by a possible abuse of notation):

$$\begin{aligned}
\dot{x}_1 &= x_2 + f_1(x_1) + \phi_1'(x_1)\theta_1 + \sigma_1'(x_1)w_1 \\
&\vdots \quad \vdots \\
\dot{x}_{n-1} &= x_n + f_{n-1}(x_1,\ldots,x_{n-1}) + \phi_{n-1}'(x_1,\ldots,x_{n-1})\theta_{n-1} \\
&\qquad +\sigma_{n-1}'(x_1,\ldots,x_{n-1})w_{n-1} \\
\dot{x}_n &= f_n(x_1,\ldots,x_n) + \phi_n'(x_1,\ldots,x_n)\phi_n + b(x_1,\ldots,x_n)u + \sigma_n'(x_1,\ldots,x_n)w_n \\
z &= x_1.
\end{aligned} \tag{11}$$

where $w := (w_1', \ldots, w_n')'$ is the q-dimensional disturbance input, where w_i is of dimension q_i, $i = 1,\ldots,n$; $\theta = (\theta_1', \ldots, \theta_n')$ is the p-dimensional vector of unknown parameters, where θ_i is of dimension p_i, $i = 1,\ldots,n$; z is the scalar output; and the nonlinear functions f_i, ϕ_i, σ_i, $i = 1,\ldots,n$, are known and satisfy the triangular structure depicted above. We assume that

A1 f_i, ϕ_i, $\sigma_i \in \mathcal{C}^{n-i+1}$, $i = 1,\ldots,n$; $b \in \mathcal{C}^1$, $(1/b) \in \mathcal{C}^1$.

A2 $\sigma_i'(x)\sigma_i(x) > c$, $\forall x \in \mathbf{R}^n$, $i = 1,\ldots,n$, for some $c > 0$.

A3 The reference trajectory $z_r \in \mathcal{C}^n$, and both z_r and its first n derivatives are uniformly bounded on $[0,\infty)$.

Let us first endow the controller (2) with also the derivative of the state, $\dot{x}$, under which the inner maximization of (8) can be performed to yield [5] the identifier dynamics (for θ):

$$\dot{\hat{\theta}}_i = \Sigma_i\phi_i(\sigma_i'\sigma_i)^{-1}(\dot{x}_i - \mathcal{X}_i - \phi_i'\hat{\theta}_i); \qquad \hat{\theta}_i(0) = \bar{\theta}_i \tag{12}$$

$$\dot{\Sigma}_i = -\Sigma_i(\phi_i(\sigma_i'\sigma_i)^{-1}\phi_i' - Q_i)\Sigma_i; \qquad \Sigma_i(0) = Q_{0i}^{-1} \tag{13}$$

$$i = 1,\ldots,n$$

where

$$\mathcal{X}_i := f_i + x_{i+1}, \qquad i = 1,\ldots,n-1; \qquad \mathcal{X}_n := f_n + bu \tag{14}$$

and Q_i, Q_{0i} are positive-definite matrices constituting the i-th diagonal blocks of Q and Q_{oi}, respectively. The identifier $\hat{\theta}_i$, $i = 1,\ldots,n$, above is asymptotically convergent to the true value of the parameter vector θ

provided that $\Sigma_i(t) > 0\ \forall t \in [0,\infty)$, $i = 1,\ldots,n$, which is a *persistency of excitation* condition, that can be guaranteed by restricting the disturbance inputs to a particular set [5]. This set can be made unconstrained, and equal to $\mathcal{L}_\infty$, by choosing the design matrix Q_i as

$$Q_i = c_i\phi_i(\sigma_i'\sigma_i)^{-1}\phi_i', \qquad c_i \in [0,1), \quad i = 1,\ldots,n. \tag{15}$$

The identification error $\tilde{\theta} := \theta - \hat{\theta}$ satisfies:

$$\dot{\tilde{\theta}}_i = -\Sigma_i\phi_i(\sigma_i'\sigma_i)^{-1}\sigma_i'v_i, \qquad i = 1,\ldots,n \tag{16}$$

where

$$v_i := w_i + \sigma_i(\sigma_i'\sigma_i)^{-1}\phi_i'\tilde{\theta}_i, \quad i = 1,\ldots,n \tag{17}$$

is a transformed disturbance input. This converts the original attenuation problem with respect to w to a new (but equivalent) attenuation problem with respect to v, with dynamics described by (in place of (11)):

$$\dot{x}_i = \mathcal{X}_i + \phi_i'\hat{\theta}_i + \sigma_i'v_i, \quad i = 1,\ldots,n \tag{18}$$

$$\dot{\hat{\theta}}_i = \Sigma_i\phi_i(\sigma_i'\sigma_i)^{-1}\sigma_i'v_i, \quad i = 1,\ldots,n \tag{19}$$

along with (13). Note that there is no parametric uncertainty here, and hence the problem has been converted to one with perfect state measurements, where x_i, $\hat{\theta}_i$ and Σ_i, $i = 1,\ldots,n$, constitute the new states. This new nonlinear H^∞ control problem (with perfect state measurements) corresponds to the outer maximization problem in (8), where the cost to be maximized is

$$\int_0^t \left((z - z_r)^2 + \ell(\tau, x_{[0,\tau]}) - \gamma^2 \sum_{i=1}^{n} \mid \tilde{\theta}_i \mid^2_{Q_i} - \gamma^2|v|^2 \right) d\tau - \gamma^2 \sum_{i=1}^{n} \mid \tilde{\theta}_i(t) \mid^2_{\Sigma_i^{-1}(t)} - \ell_0(x(0), \theta - \bar{\theta}) \tag{20}$$

where we have dropped the control dependence in ℓ, and have absorbed γ^2 in ℓ_0.

This is now the control design step, which we carry out under assumptions **A1-A3**. The design procedure here is *backstepping*, which proceeds as follows: We first consider the first subsystem ($i = 1$), and treat x_2 as an input to this system. Introducing the transformed variable $y_1 := x_1 - z_r$, one can show that this decoupled scalar system with x_2 as an input can be made to achieve arbitrarily small levels of disturbance attenuation by picking x_2 appropriately. A corresponding value function for (20) for only this subsystem is $\bar{V}_1(y_1) = \frac{1}{2}y_1^2$. However, since x_2 is not a control input this is not exact, and hence we proceed to the next subsystem ($i = 2$) and choose x_3 as the new control input, where the dynamics for x_2 are now replaced by the dynamics of y_2, which stands for the difference between x_2 and its ideal value, had it been the control variable at step 1. At this step, again there exist choices for x_3 that make the attenuation level arbitrarily close to zero, with a corresponding value function being $\bar{V}_2 = \frac{1}{2}(y_1^2 + y_2^2)$. Again x_3 is not the true control and hence this result is not exact ..... Proceeding in this manner, we arrive at step n at the last subsystem where the real control appears, a proper choice for which makes the overall attenuation level again arbitrarily small. Because of space limitations, expressions for the construction of this controller (which are quite lengthy) are not given here; they can be found in an internal report available from the author. These steps now lead us to the following theorem.

Theorem 1. *Consider the nonlinear system described by (11) and with the performance index (4) where $\tilde{\ell}_t$ does not depend on u. Let assumptions* **A1-A3** *hold, and disturbances belong to a set (say $\mathcal{W}$) that makes $\Sigma_i(t)$ positive definite for all $t \geq 0$, $i = 1,\ldots,n$. Finally, let the derivative of x also be available for feedback. Then:*

(i) The control law generated by the backstepping procedure outlined above achieves asymptotic tracking with an arbitrarily small level of disturbance attenuation, γ, for all $w \in \mathcal{W}$.

(ii) For any $w_{[0,\infty)} \in \mathcal{L}_\infty$, $x(0)$, θ and $t \geq 0$, if the covariance matrices Σ_i's are uniformly upper bounded on $[0,t]$, then the expanded state vector ζ (consisting of x, $\hat{\theta}$ and Σ) is uniformly bounded on $[0,t]$, and Σ_i's are further uniformly bounded from below by some positive-definite matrices.

(iii) For any uncertainty triple in the set $\mathcal{W}$ such that $w_{[0,\infty)} \equiv 0$, if the covariance matrices Σ_i, $i = 1, \ldots, n$, become uniformly upper bounded on $[0, \infty)$, then the parameter estimates are uniformly bounded and the transformed state variable $y := (y_1, \ldots, y_n)'$ converges to zero as $t \to \infty$; if, in addition, the reference signal z_r is persistently exciting, i.e., $\lim_{t\to\infty} \lambda_{\max}\Sigma_i = 0$, $i = 1, \ldots, n$, then ζ converges to zero as $t \to \infty$.

Remark 1. The controller above is not a certainty equivalent controller, that is it does not correspond to the controller obtained by assuming full knowledge of parameter values and then replacing the true values of the parameters by their estimates. It is, however, asymptotically certain equivalent, as $\Sigma_i \to 0$, $i = 1, \ldots, n$, under a specified choice of the design parameters. ◇

A disadvantage of the controller presented above is that it depends (through the identifier dynamics) on the derivative of the state, which may not be available. To remove this dependence, so as to obtain a controller under the original measurement scheme (2), we first consider a noise-perturbed measurement:

$$u(t) = \mu(t, y(t)), \qquad y(t) = x(t) + \epsilon v(t),$$

where ϵ is a small positive parameter and v is an unknown disturbance. The identifier dynamics corresponding to this measurement (as the counterpart of (12)-(13)), and after v is set equal to zero are [5]:

$$\dot{\hat{\theta}}_i = \Sigma_i \phi_i (\sigma_i'\sigma_i)^{-1/2} \frac{1}{\epsilon}(x_i - \hat{x}_i); \qquad \hat{\theta}_i(0) = \bar{\theta}_i \quad . \tag{21}$$

$$\dot{\Sigma}_i = -\Sigma_i(\phi_i(\sigma_i'\sigma_i)^{-1}\phi_i' - Q_i)\Sigma_i; \qquad \Sigma_i(0) = Q_{0i}^{-1} \tag{22}$$

$$\dot{\hat{x}} = \chi_i + \phi_i'\hat{\theta}_i + \frac{1}{\epsilon}(\sigma_i'\sigma_i)^{1/2}(x_i - \hat{x}_i) : \qquad \hat{x}_i(0) = x_i(0) \tag{23}$$

$$i = 1, \ldots, n$$

where we now have additional dynamics representing the estimate for x. An appropriate choice for the design matrix Q_i in this case turns out to be

$$Q_i = \Sigma_i^{-1}\Delta_i\Sigma_i^{-1} + \check{Q}_i; \quad \Delta_i > \kappa_Q I_{p_i} \tag{24}$$

for some symmetric matrices Δ_i and $\check{Q}_i$. Now, introducing

$$e := (x - \hat{x})/\epsilon$$

whose i-th component is $e_i = (x_i - \hat{x}_i)/\epsilon$, we can equivalently write (21)-(22) as (using also the specific choice made for Q_i):

$$\dot{\hat{\theta}}_i = \Sigma_i \phi_i (\sigma_i'\sigma_i)^{-1/2} e_i \tag{25}$$

$$\dot{\Sigma}_i = -\Sigma_i(\phi_i(\sigma_i'\sigma_i)^{-1}\phi_i' - \check{Q}_i)\Sigma_i + \Delta_i \tag{26}$$

$$\epsilon\dot{e}_i = -(\sigma_i'\sigma_i)^{1/2} e_i + \phi_i'\tilde{\theta}_i + \sigma_i'\omega_i, \tag{27}$$

which involves singularly perturbed dynamics. It should be noted that formally setting $\epsilon = 0$ in (27) and substituting the resulting expression for e_i into (25) yields precisely the identifier dynamics (19). Using this limiting relationship (which can be made precise using singular perturbations analysis), and the same backstepping design tool as in the earlier case, we obtain a robust disturbance attenuating controller in exactly the same form as in Theorem 1 but with the identifier now generated by (25)-(27). For a precise statement of this result, which would be the counterpart of Theorem 1 here, we first introduce a class of admissible uncertainties, $\mathcal{W}_C$, as the counterpart of $\mathcal{W}$ introduced in Theorem 1. For some arbitrary positive constant C, let

$$\mathcal{W}_C := \{(x(0), \theta, w_{[0,\infty)}) : \lambda_{\max}\Sigma_i(t) \le C, |x(0)| \le C, |\theta| \le C, |w(t)| \le C, \forall t \in [0, \infty), \forall i = 1, \ldots, n\} \tag{28}$$

Theorem 2. *Consider the nonlinear system described by (11) with perfect state (but not derivative) information, and with performance index (4) where $\tilde{\ell}_t$ does not depend on u and is positive definite. Let assumptions* **A1-A3** *hold, Q_i be given by (24), and $\mathcal{W}_C$ be as defined by (28). Then:*

(i) There exists a positive scalar $\epsilon_0 > 0$ such that for all $\epsilon \in (0, \epsilon_0]$, the control law of Theorem 1, with identifier (25)-(27), achieves asymptotic tracking with disturbance attenuation level γ for any uncertainty triple in the set $\mathcal{W}_C$. Furthermore, the closed-loop signals generated by the overall system are uniformly bounded on $[0, \infty)$.

(ii) For any uncertainty triple in the set $\mathcal{W}_C$ such that $w_{[0,\infty)} \equiv 0$, the expanded state vector, including the system state, and both slow and fast parameter errors, converges to 0 as $t \to \infty$ for any $\epsilon \in (0, \epsilon_0]$.

Remark 2. The passage from Theorem 1 to Theorem 2 has involved (in order to avoid the use of derivative information, and singularity in the optimization problem under pure state measurements) the introduction of small noise in the measurement equation, obtaining a controller along with a worst-case identifier under this noise-perturbed measurement, and then setting the disturbance (noise) entering the measurement equation to *zero*. The resulting identifier dynamics still depend on the small parameter ϵ multiplying the measurement disturbance even after the disturbance has been eliminated. This way, any performance achieved under derivative information can be achieved by using only state information. We should also note that in this case, due to the requirement that the error covariance matrices be bounded away from zero, the robust adaptive controller will not be certainty equivalent, even asymptotically. ◇

THE CASE OF OUTPUT MEASUREMENTS

Let us now turn to the case of output measurements, that is the case when not all state variables but only a subset of them is available for control purposes. In particular, let us consider in the context of the parametric strict feedback form (11) only the output, z, to be available. As in the previous section, let us first assume that the derivative of z is also measurable and is available for control purposes. Then, the first subsystem of (11) serves as the measurement equation:

$$\dot{z} - f_1(z) \;=\; x_2 + \phi_1'(z)\theta_1 + \sigma_1'(z)w_1, \tag{29}$$

through which noisy information is available on x_2 and θ_1 — with the noise being due to the presence of the disturbance w_1. Denote the remaining components of x by x_F, and let ξ denote the extended state $(\theta', x_F')'$, which satisfies an equation of the form

$$\dot{\xi} \;=\; A\xi + f + Hw, \quad \xi(0) = (\theta', x_F(0)')' \tag{30}$$

with obvious definitions for A, f and H. It should be noted that A depends on u (linearly) , and the dependence of f on x_F is in a lower triangular form. In terms of this notation, (29) can be rewritten as

$$\dot{z} - f_1(z) \;=\; C'(z)\xi + \sigma_i'(z)w \tag{31}$$

where again the definition of C should be obvious.

Now, with (30) serving as a state equation and (31) as the measurement equation, the inner supremization of (9) becomes an H^∞ filtering problem which can be solved using the theory of [1, chapter 7], leading to the following optimal (worst-case) observer and error covariance equations:

$$\dot{\hat{\xi}} \;=\; A\xi + f + (\gamma^2\Sigma C + L)N(\dot{z} - f_1 - C'\hat{\xi}), \quad \hat{\xi}(0) = \begin{pmatrix} \hat{\theta}_0 \\ \hat{x}_{F_0} \end{pmatrix} \tag{32}$$

$$\begin{aligned} \dot{\bar{\Sigma}} \;=\; & (A - LNC')\bar{\Sigma} + \bar{\Sigma}(A - LNC')' - \bar{\Sigma}(\gamma^2 CNC' - Q)\bar{\Sigma} + \gamma^{-2}(HH' - LNL'), \\ & \bar{\Sigma}(0) = \gamma^{-2}\text{blockdiag}(\Sigma_0 \;\; \pi_0) \end{aligned} \tag{33}$$

where $\hat{\theta}_0$ and $\hat{x}_{F_0}$ denote the initial (*a priori*) estimates for θ and $x_F(0)$, respectively, $L := H\sigma_1$, $N := (\sigma_1'\sigma_1)^{-1}$, and Q is a nonnegative-definite matrix, serving as a Euclidean weighting on the estimation error

$\xi - \hat{\xi}$, which is a part of $\tilde{\ell}_t$ in (4) (or equivalently (6)). To ensure boundedness of parameter error, it is generally useful to add to the right hand side of (32) a smooth function that forces the parameter estimate (the first p components of $\hat{\xi}$) to stay within an a priori known set θ_0 (where all the parameters lie); see [10] for details. This then completes the design of the identifier/estimator, and brings us to the control design stage (i.e., the outer maximization in (9)). The combined state, estimator and error covariance dynamics are again in strict feedback form, which makes it possible to apply the backstepping tool of the previous sections; the details are lengthy and have not been included here due to page limitations.

The procedure outlined above leads to a controller that depends not only on z but also on $\dot{z}$. To remove the dependence on $\dot{z}$ we again follow a procedure similar to that carried out in the previous section, to go from derivative measurements to the state measurement case. We introduce a new measurement, y, which is a noise-perturbed version of z: $y = z + \epsilon v$, where v is a scalar unknown disturbance, and ϵ is a small positive parameter. The inner maximization problem of (9) can be solved as in the derivative measurement case, to which we subsequently apply singular perturbations analysis to obtain estimators that are well-defined when $v \equiv 0$ and ϵ is small. Then, the solution of the outer maximization problem again involves backstepping, leading to a robust adaptive controller which uses only the given scalar output measurement. Under some technical conditions, one can then prove a result similar to Theorem 2, assuring asymptotic tracking property of the derived controller for sufficiently small $\epsilon > 0$, with an arbitrary level of disturbance attenuation, and for all uncertainty belonging to a given set. Furthermore, all closed-loop signals remain bounded for all time.

ACKNOWLEDGEMENT

This work was performed under the auspices of the U.S. Department of Energy. Contributions of Z. Pan and I.E. Tezcan to this effort are also acknowledged.

REFERENCES

1. T. BAŞAR and P. BERNHARD, *H^∞-Optimal Control and Related Minimax Design Problems: A Dynamic Game Approach*, Birkhäuser, Boston, MA, 2nd edition (1995).

2. T. BAŞAR and G. J. OLSDER, *Dynamic Noncooperative Game Theory*, Academic Press, London, 2nd edition (1995).

3. J. A. BALL, J. W. HELTON, and M. WALKER, "H^∞ control for nonlinear systems with output feedback," *IEEE Trans. Automatic Control*, 38(4):546-559 (1993).

4. G. DIDINSKY and T. BAŞAR, "Minimax adaptive control of uncertain plants," *Proceedings of the 33rd IEEE Conference on Decision and Control*, 2839-2844, Orlando, FL (1994).

5. G. DIDINSKY, Z. PAN, and T. BAŞAR, "Parameter identification for uncertain plants using H^∞ methods," *Automatica*, 31(9):1227-1250 (1995).

6. I. KANELLAKOPOULOS, P. V. KOKOTOVIĆ, and A. S. MORSE, "Systematic design of adaptive controllers for feedback linearizable systems," *IEEE Trans. Automatic Control*, 36:1241-1253 (1991).

7. M. KRSTIĆ, I. KANELLAKOPOULOS, and P. V. KOKOTOVIĆ, *Nonlinear and Adaptive Controllers Design*, Wiley, NY (1995).

8. M. KRSTIĆ and P. V. KOKOTOVIĆ, "Adaptive nonlinear design with controller-identifier separation and swapping, *IEEE Trans. Automatic Control*, 40(3):426-440 (1995).

9. Z. PAN and T. BAŞAR, "Adaptive controller design for tracking and disturbance attenuation in parametric-strict-feedback nonlinear systems," *Proc. 13th IFAC World Congress*, San Francisco, CA, F:323-328 (1996).

10. I. E. TEZCAN and T. BAŞAR, "Disturbance attenuating adaptive controllers for parametric strict feedback nonlinear systems with output measurements," *Proc. 1997 American Control Conference*, Albuquerque, NM (1997).

USE OF LASER DIODES IN CAVITY RING-DOWN SPECTROSCOPY

R. N. Zare, B. A. Paldus, Y. Ma, and J. Xie

Department of Chemistry, Stanford University
Stanford, CA 94305-5080, USA

ABSTRACT

We have demonstrated that cavity ring-down spectroscopy (CRDS), a highly sensitive absorption technique, is versatile enough to serve as a complete diagnostic for materials process control. In particular, we have used CRDS in the ultraviolet to determine the concentration profile of methyl radicals in a hot-filament diamond reactor; we have applied CRDS in the mid-infrared to detect 50 ppb of methane in a N_2 environment; and, we have extended CRDS so that we can use continuous-wave diode laser sources. Using a laser diode at 810 nm, we were able to achieve a sensitivity of 2×10^{-8} cm^{-1}. Thus, CRDS can be used not only as an *in situ* diagnostic for investigating the chemistry of diamond film deposition, but it can also be used as a gas purity diagnostic for any chemical vapor deposition system.

INTRODUCTION

Present-day technology is dominated by the synthesis of materials, ranging from biocompatible plastics, to metal-semiconductor heterostructures for lasers used in telecommunications, to silicon oxides and nitrides that provide the backbone of the electronics industry. Materials process control is rapidly becoming more important in industry, and is triggering fundamental research of materials and their chemistries.

Diamond films, because of their mechanical hardness, high thermal conductivity, and excellent optical properties are commercially important in a wide set of applications, ranging from the more traditional tool coating to integrated circuit fabrication to even modern sound system manufacturing. Diamond deposition by plasma, oxy-acetylene flame, and hot-filament chemical vapor deposition is a rapidly growing technology. Intense interest exists in the study of the basic reaction mechanisms in both the gas-phase and surface chemistries, because presently diamond synthesis remains more an art than an empirical process. New laser diagnostics developed during this study are being directly applied to various diamond deposition environments, such as inductively coupled plasma torch and hot-filament chemical vapor deposition, under the continuing

collaboration with the nonequilibrium plasma chemistry program of Prof. Charles H. Kruger at the High Temperature Gas Dynamics Laboratory, Stanford University.

Our principal diagnostic tool is based on cavity ring-down spectroscopy (CRDS). CRDS is a high-sensitivity absorption technique with potential for absolute concentration measurements of trace gases and impurities[1]. CRDS is usually practiced by coupling a pulsed laser source into a high-finesse optical resonator (Fabry-Perot cavity) that encloses the sample of interest, and detecting the decay of light in the resonator. Under many conditions, the decay is exponential, and a plot of the ring-down lifetime versus frequency gives the absorption spectrum[2]. The ring-down lifetime is controlled by the resonator finesse, and changes wherever the sample absorbs the wavelength of the incident radiation.

Most diagnostics used in research, however, tend to rely on expensive equipment that is difficult to maintain. To increase the utility of our diagnostics, we have begun to investigate practical schemes for CRDS. In particular, laser diodes, owing to their small size, low cost and relative ease of use, have begun to play a more dominant role in our research, and will open the possibility of portable diagnostics.

A MODEL SYSTEM: DIAMOND FILM GROWTH

A particularly suitable system for study of energy-related phenomena is the diamond film reactor, where the growth mechanism directly involves plasma chemistry. Two of the commonly used diamond film deposition methods are a CVD reactor using hot-filament chemical vapor deposition (HFCVD) or an inductively coupled atmospheric plasma torch. Both techniques are already under investigation at the Stanford High Temperature Gas Dynamics Laboratory. In order to understand the elementary growth mechanisms involved in diamond deposition, data bases of information about the numerous radicals present (e.g., hydrocarbon radicals as CH_3, CH_2, CH, C_2H, C_2, etc., or atomic hydrogen) are being compiled and will be used in future computer modeling, and subsequent numerical simulation of the complex plasma chemistry (e.g., gas-phase reactions of atomic hydrogen with hydrocarbon radicals or diamond interface reactions of atomic hydrogen selectively with graphite).

A CRDS setup has been designed to measure trace radical species generated in a hot-filament reactor for diamond deposition[3]. The methyl (CH_3) radical is an important free radical present during the initial stages of hydrocarbon combustion: it is believed to be a precursor for diamond growth by CVD. *In situ* measurements of methyl radical concentrations (cf. schematic diagram of reactor in figure 1a) have been carried out under various conditions[4,5]. Typically, a mixture of H_2 of CH_4 is flowed through the previously evacuated reactor. A tungsten filament is positioned vertically inside the reactor chamber and is resistively heated to a specified brightness temperature. Methyl radical absorption is observed near 216 nm, where feature lines are a few nm wide (cf. figure 1b)[3]. It is also important that the ground-state population of the absorber molecule is not significantly depleted by excitation during the time the laser pulse is circulating inside the optical cavity. In our experiment, for 216 nm light pulse of energy about 0.2 mJ and TEM_{00} mode radius $w = 250$ mm, for mirrors reflectivity $R = 0.991$, and for CH_3 absorption cross-section $s < 10^{-17}$ cm^2 / molecule, the fraction of molecules excited by the laser pulse inside the cavity is less than $3x10^{-3}$, which is sufficient for accurate CRDS measurements.

A profile of CH_3 absolute concentration near the hot filament has been determined by CRDS using a topological method - Abel inversion of the spatial profile of CH_3 absorbance (cf. figure 1c)[3,4,5].

This approach allowed us to estimate the uncertainty in the inverted profile. The error bars represent one standard deviation. The shaded part of the figure indicates radial distances from the filament where the gas temperature is between 1250 K and 2000 K. Based on a hydrogen diffusion model, methyl concentration should peak at the filament. It was unexpectedly observed, however, to peak about 5 mm from the filament. This behavior can possibly be explained by the Soret effect or dissociation of methyl near the filament (cf. figure 1d)[4,5].

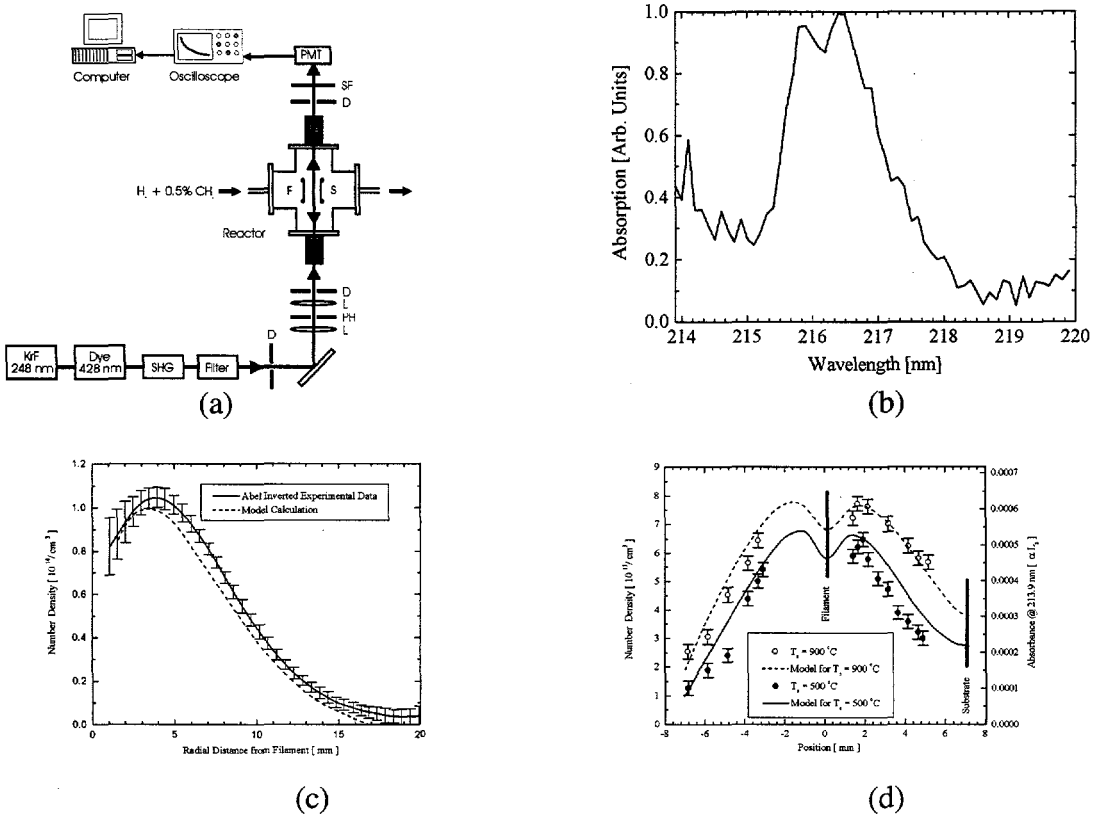


Figure 1: (a) CRDS setup for radical concentration measurements, (b) spectrum of methyl absorption at 216 nm, (c) radial distribution of CH_3, and (d) spatial profiles of the measured number density within the hot-filament reactor at two different substrate temperatures.

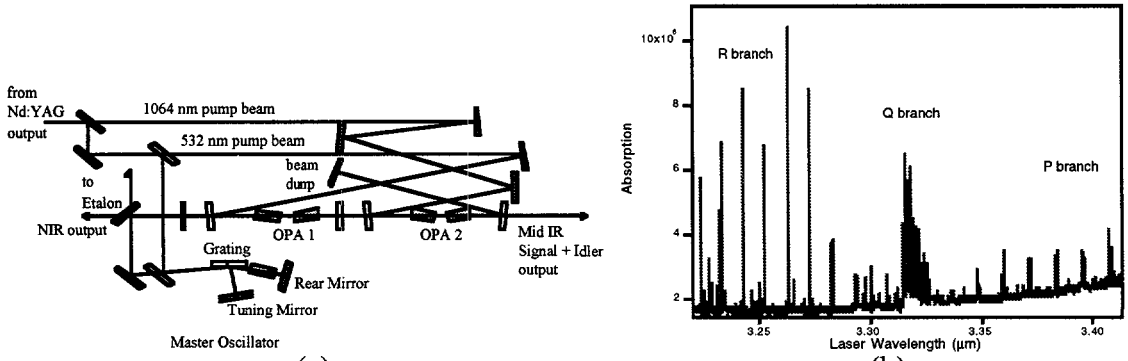
EXTENSIONS TO THE MID-INFRARED

The 1.5 to 10 μm region of the electromagnetic spectrum is rich in rovibrational transitions forming molecular "fingerprints" that are well known to be a means for identifying and characterizing specific species. This region is therefore rather ideal for mapping species concentration or temperature gradients in hot-filament reactors and arc jets. We have begun to exploit the high sensitivity, linearity, and simplicity in quantifying number densities provided by CRDS in the mid-infrared.

The application of CRDS to a problem presupposes the existence and availability of suitable light sources and cavity mirrors. With the advent of nonlinear optical devices, it has recently become possible to obtain tunable coherent light sources in the mid-infrared based on optical parametric oscillators (OPO)[2]. Simultaneously, highly reflecting mirrors with only minute scattering and

absorption losses have become available for wavelengths in the visible and the near infrared regions.

Our light source is a Nd:YAG laser-pumped OPO system (Continuum Mirage 3000) that can generate nearly Fourier transform-limited nanosecond Gaussian pulses with a manufacturer-specified bandwidth of 500 MHz (0.017 cm^{-1}) at a repetition rate of 10 Hz[2]. The wavelength can be tuned continuously from 1.5 to 4.0 μm, with the pulse energy decreasing from 8 mJ to 1 mJ, respectively. The OPO system architecture is shown in figure 2a[2].



(a) (b)

Figure 2: (a) Continuum Mirage 3000 OPO system diagram, and (b) absorption spectrum of a 100 ppm CH_4 in N_2 mixture at 50 Torr pressure.

We are currently pursuing CRDS studies of the well-known methane fundamental C-H stretching mode (n3), that occurs around 3.17 mm, and should serve as a good reference for future calibrating purposes. A typical absorption spectrum is given in figure 2b. All recorded spectra showed a very strong absorption, allowing us to record methane lines below 10^{-8} Torr partial pressure in N_2.

We have also applied our OPO system to the measurement of water vapor in various types of flames, to demonstrate the effectiveness of CRDS as a diagnostic tool for hostile environments such as flames, discharges, flashes, or plasmas[6]. A strong need exists for spectroscopic methods that can serve as remote diagnostics in these environments because they remain difficult to characterize, owing to their wide range of extreme physical conditions: high temperatures and consequently strong luminous background, sharp gradients in both temperature and density, and a reactive medium with ions, electrons and a variety of free radicals or intermediate states. CRDS, a laser-based spectroscopy, which is noninvasive, species specific, and spatially resolved, is ideally suited for probing environments like these.

We have measured the spectrum of water vapor in air from 810 to 820 nm, from atmospheric pressure to 20 mTorr, with a resolution of 0.001 nm (0.015 cm^{-1}). This demonstrates a nominal measurement sensitivity (with R=99.99% mirrors) to absorption coefficients as low as 1.7 x 10^{-7} cm^{-8}. We have also been able to extract accurate species partial pressure measurements of water vapor in a regulated cell (figure 3)[6]. We have subsequently measured a similar spectra of water vapor generated at the tip of a propane torch flame (T = 2000 K), and at various heights above a controlled plane methane-air burner[6]. By using the HITEMP database, we can extract rotational temperatures of water vapor at different heights above the plane burner surface. Figure 4 compares spectra of water vapor at room temperature to those in the propane flame, while figure 5 illustrates changes in the water spectrum, caused by the decreasing temperature gradient in a controlled flame.

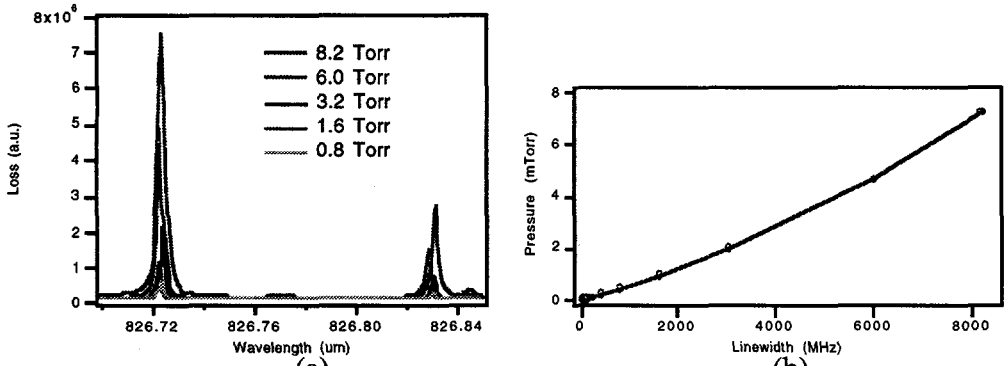


Figure 3: (a) Absorption spectra of water vapor at various cell pressures, and (b) variation of linewidth with pressure.

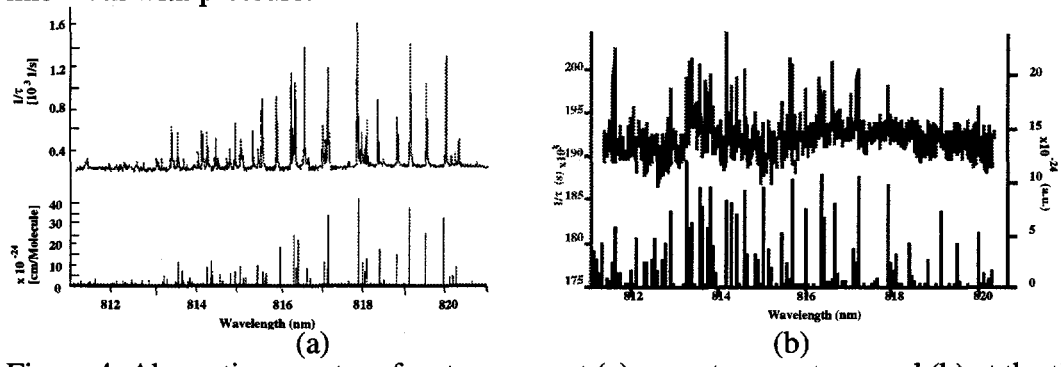


Figure 4: Absorption spectra of water vapor at (a) room temperature, and (b) at the tip of a propane torch. Measured spectra are at the top, while spectra from HITRAN96 are on the bottom.

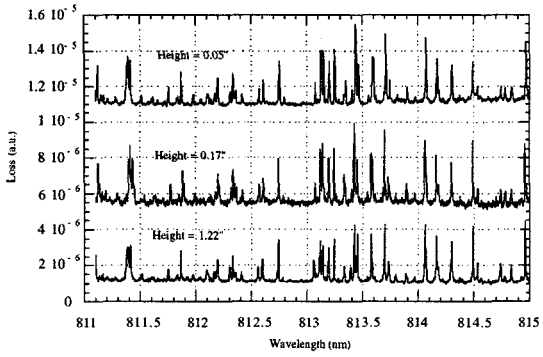


Figure 5: Absorption spectra of water vapor at various heights in a controlled methane burner.

MINIATURIZATION WITH LASER DIODE SOURCES

Much of current ring-down spectroscopy still relies on fairly costly laser sources. As solid state lasers (e.g., Ti:Sapphire lasers, Nd:Yag-pumped OPOs, and ECDLs) have gained in reliability, tuning range, and output power, they have started to replace the more traditional tunable dye lasers, although they are no less expensive. Simultaneously, semiconductor laser diodes (LDs) have also been improving in power, wavelength coverage, and reliability. The rapid growth of the communications industry in recent years has resulted in the availability of tunable UV-, near- and mid-infrared LDs at a rapidly diminishing cost (< $2000). In fact, owing to their compactness,

low cost, durability, high wallplug efficiency, and compatibility with both fiber and silicon technologies, infrared laser diodes seem to be an ideal light source for realizing practical CRDS systems.

Early attempts demonstrated difficulties in applying LD sources to CRDS: whenever a LD beam is reflected directly back into the laser, as is inevitable in a linear cavity configuration, even under optical isolation, the optical feedback results in phase fluctuations and mode hopping of the LD. In fact, at higher feedback levels, a wide variety of effects ranging from linewidth broadening to complete `coherence collapse' (linewidth > 10 GHz) is often observed and is illustrated in figure 6a[7]. The inherent problem is the formation of `external cavities' by reflective optics with the back facet of the LD that affect both the gain and phase relations of the LD. Thus, whenever back reflection is allowed, the lasing characteristics of the become highly dependent on uncontrollable experimental parameters, most notably the external cavity length.

Several solutions exist to this coupling problem. A LD with a high quality (but expensive) AR coated output facet can function as a gain medium in an external cavity; the feedback from a linear cavity configuration can be completely eliminated by using a ring resonator structure, as will be investigated in the future; or, the external cavity effect can be controlled by placing an acousto-optic modulator (AOM) inside the external cavity, thereby stabilizing the time-averaged behavior of the LD. The last approach, first demonstrated by Martin et al.[7] as a useful scheme for stabilizing LDs in the presence of direct back reflections, was the point of departure for our LD research.

By placing an AOM between the laser diode and the input mirror of the ring-down cavity, the AOM can be used not only to switch the CW beam into and out off the first order diffraction, but simultaneously control LD linewidth. The AOM driving power determines the diffraction efficiency and hence the amount of feedback to the LD. The external cavity length fixes the maximum achievable linewidth for each feedback level (cf. figure 6b)[8]. The first order diffraction feedback drives the LD phase and stabilizes linewidth. Finally, the linewidth can be further enhanced by introducing nonfrequency-shifted that cyclically chirps the LD output through multiple external cavity modes, at twice the AOM driver frequency (cf. figure 6c)[8]. The flexibility in achievable LD linewidth in turn enables many different CRDS applications.

Using the AOM stabilization scheme for a laser diode source, shown in figure 7, we were able to perform CRDS on water vapor present in ambient air or in an evacuated optical cavity[8]. LD linewidth control was performed with feedback from both first and zeroth orders. Spectra of water vapor in room air and at 5 Torr are given in figure 8[8]. Spectra were obtained in one continuous scan. Spectra at ambient pressure used maximum zeroth order feedback (47.6 dB) to achieve the largest possible linewidth (240 - 500 MHz) and cavity coupling. Spectra at low pressures (< 100 Torr) used less zeroth order coupling (58 dB) to achieve a narrower laser linewidth (180 - 240 MHz) and to avoid convolution of the laser line with the absorption line. Scan step size in both cases remained limited to 0.001 nm resolution by the current step resolution (0.1 mA) of the LD driver. No baseline adjustments have been made, and the overall baseline noise results from the excitation of multiple transverse modes in the cavity, which were used to improve light throughput. Nonetheless, our detection limit of 2×10^{-8} cm^{-1} remains quite respectable for an inexpensive LD source, especially when compared to pulsed CRDS.

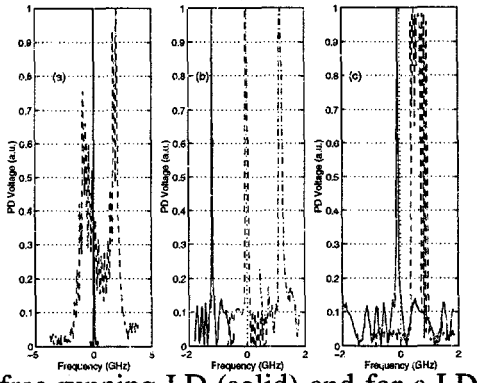


Figure 6: (a) Linewidth for a free-running LD (solid) and for a LD under feedback (dashed). (b) LD linewidth as a function of external cavity length for only first order feedback: L_{ext}=215 cm (solid), L_{ext}=100 cm (dashed), and L_{ext}=215 cm (dash-dotted), (c) LD linewidth for only first order feedback (solid) and both first and zeroth order feedback (dashed).

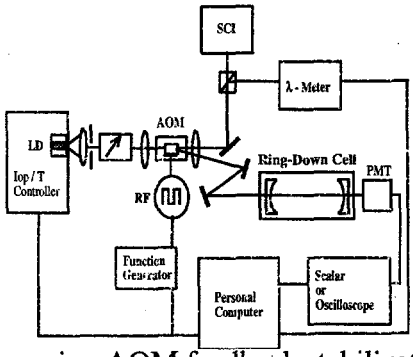


Figure 7: Laser diode CRDS setup using AOM feedback stabilization.

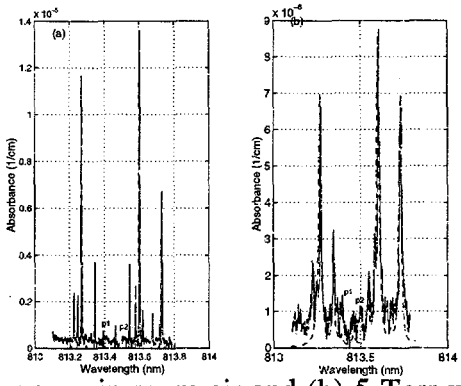


Figure 8: Spectrum of (a) water vapor in room air and (b) 5 Torr water vapor in a cell previously evacuated below 1 mTorr. Spectra based on HITRAN96 are shown as dashed lines.

CONCLUSIONS

CRDS has been applied for quantitative diagnostic study of methyl radicals in a hot-filament reactor used for diamond film synthesis. The methyl radical concentration was found to peak at several mm away from the filament surface, and is attributed to the effect of Soret diffusion. We have extended the diagnostic capabilities of our OPO laser from near-infrared studies of water vapor in harsh environments, such as flames, to mid-infrared studies of the C-H stretch in methane. This

will allow us to perform highly sensitive CRDS diagnostics of an arc-jet torch used for diamond synthesis.

Simultaneously, we have demonstrated that it is possible to not only stabilize a free-running laser diode in the presence of strong reflections from a ring-down cavity, but also control the linewidth of the laser diode. The laser diode can also be stabilized to only several MHz, if high resolution is required. We have performed CW-CRDS with ring-down repetition rates of 10-50 kHz, and have achieved a noise level of 2×10^{-8} cm^{-1}, comparable to pulsed CRDS.

ACKNOWLEDGMENTS

The authors would like to acknowledge contributions by the companion research program of Prof. C. H. Kruger, with Dr. T. G. Owano and E. H. Wahl (DOE Grant #DE-FG0388ER13957). This work has been supported by the Office of Basic Energy Research at the Department of Energy (contract #DE-FG03-92ER14303).

REFERENCES

1. P. ZALICKI and R. N. ZARE, "Cavity ring-down spectroscopy for quantitative absorption measurements," *J. Chem. Phys.* 102, 2708 (1995).

2. J. MARTIN, B. A. PALDUS, P. ZALICKI, E. H. WAHL, T. G. OWANO, J. S. HARRIS, C. H. KRUGER, and R. N. ZARE, "Cavity Ring-down Spectroscopy with Fourier-transform-limited Light Pulses," *Chem. Phys. Lett.* 258, 63 (1996).

3. P. ZALICKI, Y. MA, R. N. ZARE, E. H. WAHL, J. R. DADAMIO, T. G. OWANO and C. H. KRUGER, "Methyl radical measurement by cavity ring-down spectroscopy," *Chem. Phys. Lett.* 234, 269 (1995).

4. P. ZALICKI, Y. MA, R. N. ZARE, E. H. WAHL, T. G. OWANO, and C. H. KRUGER, "Measurement of methyl radical concentration profile in a hot-filament reactor," *Appl. Phys. Lett.* 67, 144 (1995).

5. E. H. WAHL, T. G. OWANO, C. H. KRUGER, P. ZALICKI, Y. MA, and R. N. ZARE, "Measurement of absolute CH_3 concentration in a hot-filament reactor using cavity ring-down spectroscopy [diamond CVD]," *Diamond and Related Materials* 5, 373 (1996).

6. J. XIE, J. MARTIN, B. A. PALDUS, E. H. WAHL, M. ZHAO, T. G. OWANO, C. H. KRUGER, and R. N. ZARE, "Cavity Ring-down Spectroscopic Measurements in a Flame," *Appl. Phys. Lett.* (in preparation).

7. J. MARTIN, Y. ZHOA, S. BALLE, K. BERGMANN, and M. P. FEWELL, "Visible-wavelength diode laser with weak frequency-shifted optical feedback," *Optics Communications* 112, 109 (1994).

8. B. A. PALDUS, J. S. HARRIS, J. MARTIN, J. XIE, R. N. ZARE, "Cavity Ring-Down Spectroscopy Using a Frequency-stabilized Laser Diode," *J. Appl. Phys.*, (submitted April 1997).

Final List of Participants

15th Symposium on Energy Engineering Sciences

May 14-15, 1997

Argonne National Laboratory
Argonne, Illinois

S. George Bankoff
Department of Chemical Engineering
Northwestern University
Evanston, IL 60208

Jacob Barhen
Center for Engineering Systems Advanced Research
Oak Ridge National Laboratory
P.O. Box 2008
Oak Ridge, TN 37831-6355

Tamer Basar
Coordinated Science Laboratory
University of Illinois, Urbana-Champaign
1308 West Main Street
Urbana, IL 61801-2307

Bruce S. Berger
Department of Mechanical Engineering
University of Maryland-College Park
College Park, MD 20742

Harvey W. Blanch
Department of Chemical Engineering
University of California-Berkeley
Berkeley, CA 94720

Ivan Catton
Department of Mechanical and Aerospace Engineering
University of Calfornia-Los Angeles
405 Hilgard Avenue
Los Angeles, CA 90024

Hermann Fasel
Department of Aerospace and Mechanical Engineering
University of Arizona
Building 16, Rm. 301
Tucson, AZ 85721

James R. Fincke
Department of Optical and Plasma Physics
Idaho National Engineering & Environmental Lab.
P.O. Box 1625
Idaho Falls, ID 83415-2211

Daniel Frederick
Department of Engineering Science and Mechanics
Virginia Polytechnic Institute & State University
1410 Highland Circle
Blacksburg, VA 24060

L. B. Freund
Division of Engineering
Brown University
Box D
Providence, RI 02912

Bijoy K. Ghosh
Department of Systems Science and Mathematics
Washington University
One Brookings Drive
St. Louis, MO 63130

Robert Goulard
Division of Engineering and Geosciences, ER-15
U.S. Department of Energy
Office of Basic Energy Sciences
19901 Germantown Road
Germantown, MD 20874

Takashi Hibiki
School of Nuclear Engineering
Purdue University/Kyoto University
1290 Nuclear Engineering Building
West Lafayette, IN 47907-1290

Cynthia D. Holcomb
Physical and Chemical Properties Division
National Institute of Standards & Technology
325 Broadway
Boulder, CO 80303

Mamoru Ishii
Department of Nuclear Engineering
Purdue University
1290 NUCL
West Lafayette, IN 47906

Daniel D. Joseph
Department of Aerospace Engineering
University of Minnesota
110 Union Street
Minneapolis, MN 55455

Allan N. Kaufman
Lawrence Berkeley Laboratory
MS 4/230
Berkeley, CA 94720

Gunol Kojasoy
Department of Mechanical Engineering
University of Wisconsin-Milwaukee
P.O. Box 784
Milwaukee, WI 53201

Erhard Krempl
Mechanical Engineering, Aeronautical Engineering and Mechanics
Rensselaer Polytechnic Institute
110 8th Street
Troy, NY 12180-3590

Paul A. Libby
Department of Applied Mechanics and Engineering Science
University of California-San Diego
9500 Gilman Drive
La Jolla, CA 92093-0411

Katja Lindenberg
Department of Chemistry and Biochemistry 0340
University of California-San Diego
9500 Gilman Drive
La Jolla, CA 92093-0340

Mark J. McCready
Department of Chemical Engineering
University of Notre Dame
182 Fitzpatrick Hall
Notre Dame, IN 46556

Francis C. Moon
Department of Mechanical and Aerospace Engineering
Cornell University
204 Upson Hall
Ithaca, NY 14853

Paul E. Murray
Department of Materials Joining
Idaho National Engineering & Environmental Lab.
P.O. Box 1625
Idaho Falls, ID 83415-2210

Joseph O'Gallagher
Department of Physics
The University of Chicago
Enrico Fermi Institute
5720 South Ellis Avenue
Chicago, IL 60637

Nageswara S. Rao
Center for Engineering Systems and Advanced Research
Oak Ridge National Laboratory
P.O. Box 2008, MS 6364
Oak Ridge, TN 37831-6364

Alfonso Ortega
Department of Aerospace and Mechanical Engineering
University of Arizona
Building 16, Rm. 301
Tucson, AZ 85721

Walter G. Reuter
Department of Metals and Ceramics
Lockheed Martin Idaho Technology Company
P.O. Box 1625
Idaho Falls, ID 83415-2218

Lynne E. Parker
Computer Science and Mathematics Division
Oak Ridge National Laboratory
P.O. Box 2008
Oak Ridge, TN 37831-6364

Hermann Riecke
Department of Applied Mathematics
Northwestern University
2145 Sheridan Road
Evanston, IL 60208

Tomio Y. Petrosky
Ilya Prigogine Center for Statistical Mechanics and Complex Systems
The University of Texas-Austin
RLM 7.220
Austin, TX 78712

Huseyin Sehitoglu
Department of Mechanical and Industrial Engineering
University of Illinois, Urbana-Champaign
1206 West Green Street
Urbana, IL 61801

Robert E. Price
Division of Engineering and Geosciences, ER-15
U.S. Department of Energy
Office of Basic Energy Sciences
19901 Germantown Road
Germantown, MD 20874

Herschel B. Smartt
Department of Materials Joining
Idaho National Engineering & Environmental Lab.
P.O. Box 1625
Idaho Falls, ID 83415-2210

Seth J. Putterman
Department of Physics
University of California-Los Angeles
Los Angeles, CA 90024

Vladi S. Travkin
Department of Mechanical and Aerospace Engineering
University of California-Los Angeles
48-121 Engineering IV
Box 951597
Los Angeles, CA 90095-1597

Roland Winston
Department of Physics
University of Chicago
5640 South Ellis Avenue
Chicago, IL 60637

Qiao Wu
Department of Nuclear Engineering
Purdue University
West Lafayette, IN 47907

Ying Xu
Computer Science and Mathematics Division
Oak Ridge National Laboratory
P.O. Box 2008, MS 6364
Oak Ridge, TN 37831-6364

Richard N. Zare
Department of Chemistry
Stanford University
Mudd Building
Stanford, CA 94305-5080