Proceedings of the

EIGHTEENTH SYMPOSIUM ON ENERGY ENGINEERING SCIENCES

May 15-16, 2000

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

EIGHTEENTH SYMPOSIUM ON ENERGY ENGINEERING SCIENCES

FOREWORD

This Proceedings Volume includes the technical papers that were presented during the Eighteenth Symposium on Energy Engineering Sciences on May 15-16, 2000, at Argonne National Laboratory, Argonne, Illinois. The Symposium was structured into seven technical sessions, which included 30 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 (BES), of which Engineering Sciences is a component of the Materials Sciences and Engineering Division, is responsible for the long-term, mission-oriented research in the Department. The Office has prime responsibility for establishing the basic scientific foundation upon which the Nation's future energy options will be identified, developed, and built. BES is committed to the generation of new knowledge necessary to solve present and future problems regarding energy exploration, production, conversion, and utilization, while maintaining respect for the environment.

Consistent with DOE/BES mission, the Engineering Sciences 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 to improve and extend the body of knowledge underlying current engineering practice so as to create new options for enhancing energy savings and production, prolonging the useful life of energy-related structures and equipment, and developing advanced technologies and materials processing. The program emphasis is on reducing costs through improved industrial production and performance and expanding the nation's store of fundamental knowledge for solving anticipated and unforeseen engineering problems in energy technologies.

To achieve these goals, the Engineering Sciences Program supports approximately 100 research projects covering a broad spectrum of engineering topics. The Eighteenth Symposium involved approximately one-third of the research projects currently sponsored by DOE/BES Engineering Sciences Program.

The Eighteenth Symposium was held under the joint sponsorship of the DOE Office of Basic Energy Sciences and Argonne National Laboratory (ANL). Ms. Marianne Adair and Ms. Judy Benigno of ANL Conference Services handled local arrangements. Ms. Gloria Griparis of ANL's Information and Publishing Division, Technical Communication Services was responsible for assembling these proceedings and attending to their publication.

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

Bassem F. Armaly, SC-131
Division of Materials Sciences and Engineering
Office of Basic Energy Sciences.

Proceedings of the

EIGHTEENTH SYMPOSIUM ON ENERGY ENGINEERING SCIENCES

May 15-16, 2000

Argonne National Laboratory

Argonne, Illinois

TABLE OF CONTENTS

AN OVERVIEW OF HIGH TEMPERATURE TIME DEPENDENT DAMAGE DEVELOPMENT

F. W. Brust[1], J. Oh[2], N. Katsube[2], and R. Mohan[3]

[1]Battelle, Columbus, Ohio, 43214, U. S. A.
[2]Ohio State University, Columbus, Ohio, 43214, U. S. A.
[3]Rouge Steel, Dearborn, Michigan, 48121, U. S. A.

ABSTRACT

Damage nucleation, growth, and failures of metallic structural components that operate at high temperature are overviewed. Damage nucleation usually begins with the development of small voids at a size level at the high end of the nanoscale (50 to 500 nm). These voids begin to grow via diffusion mechanisms along the grain boundaries along with dislocation creep within the grains. Voids eventually link-up to produce micro-cracks (size 2 to 20 μm). Micro-cracks then link-up to produce macro-cracks, which eventually leads to component failure. Here we overview the high temperature damage development process, especially with regard to cyclic loading, which has received little attention to date. It is seen that damage development under cyclic loading develops in a fashion quite different from the constant load situation.

INTRODUCTION

From recent studies and field experience it is now known that current engineering methods to predict the life and prevent failures of components that operate in severe high temperature environments are ineffective. Hence, an understanding of the high temperature cyclic response of these components, as well as a predictive life methodology, is very important to the DOE goal of providing safe and cheap energy to the USA. On the nanoscale level, cavitation along grain boundaries leads to isolated voids, which eventually link up and lead to a macro crack. The macro

crack then grows until it reaches a size where ultimate failure occurs. On both the micro- and macro-mechanics levels, the response under severe history dependent loading had received little attention to date. Prior work mainly focused on simple thermal and load environments. More importantly, the link between the nanoscale level where damage nucleates and the macro level, where predictions must be made has not been adequately established, especially under cyclic loading. This link is another research goal. Finally, the effect of residual stresses and porosity caused by welding on structures that operate at high temperature has not been well understood. Since failures frequently occur in the field in and near welds, it is important to extend the understanding and models to account for weld residual stress, strain, and damage effects. By learning how to manage the high temperature structural environment the goal of providing safe, cheap, and efficient energy will be improved.

HIGH TEMPERATURE DAMAGE PROGRESSION

Damage nucleation, growth, damage link-up, crack growth, and breakage is the typical progression of failure for components that operate at high temperature. Damage nucleation begins with the nucleation of a cavity at a size scale at the higher end of the nanoscale level (~50 to 300

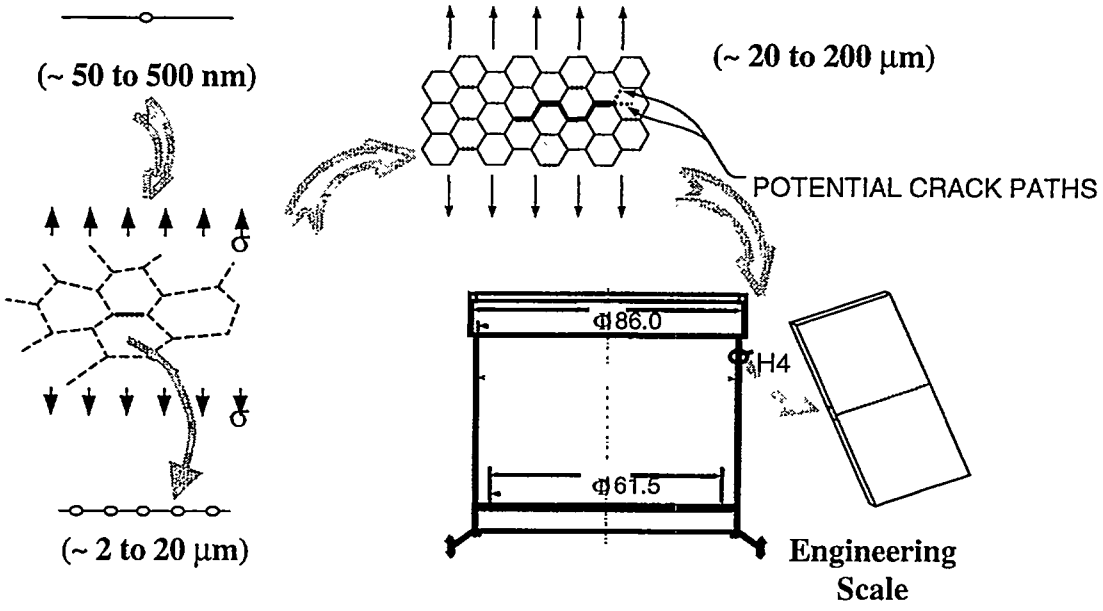

Figure 1. Scales of Creep Damage Development and Failure.

nm, depending on the material). Early in the process, such nucleation and growth phenomenon is explained by diffusion of atomic flux from the cavities to the grain boundaries, along with grain boundary sliding (to a lesser extent). As time proceeds, nonlinear viscous flow (creep) occurs, and, depending on the local stress state, eventually overrides the diffusion growth process, especially as the neighboring voids approach each other. As voids link, micro-cracks develop, link-up, and lead

to a macro-crack. Depending on the operating conditions, the macro-crack can slowly grow during component operation, or fail quickly. Often failures are catastrophic with release of large amounts of energy. In addition, as we move forward in the new millennium, higher temperature chemical processes are clearly required to increase efficiencies and reduce pollution levels

In earlier work [1-4] we focused on the understanding, control, and development of predictive methodologies to manage this type of growth under severe history dependent conditions. However, the current efforts are focused on understanding the cavitation process at the high end of the nanoscale through the grain boundary scale. Indeed, control of creep failures can only be accomplished by chemical solutions at this scale, or clever 'mechanical' solutions such as control of 'residual stresses' at this level. Since few analytical efforts have focused on cyclic time dependent cavity growth at this scale, it is the main focus of our effort. A complimentary, related effort involves the development of predictive models that can be used to control failures at the macro-scale by clever control of weld residual stresses. It is now realized that weld induced residual stresses can be a major factor in life extension.

CAVITY GROWTH BY DIFFUSION

From matter conservation and the kinetic relation between diffusive flux and chemical potential (respectively):

$$\Omega\nabla\cdot J_{GB}+\dot{\delta}=0 \qquad\qquad \Omega J_{GB}=-(D_{GB}\delta_B/kT)\nabla(\Omega\sigma_n)$$

a constitutive law for the relation between diffusive creep and field variables was developed and implemented within a finite element framework [5]. In the above, Ω is atomic volume, J is atomic flux measured in units of atoms crossing unit length per unit time, δ is the grain boundary opening (due to matter addition) with the over-riding '.' representing rate, and the rest are constants. A periodic grain boundary model was developed and the effect of elastic accommodation on grain boundary diffusion creep was studied. Figure 2 illustrates results. In Figure 2, 'a' represents void size and 'b' represents void spacing. The solution was normalized by a classical closed form expression [6], which was derived neglecting elastic effects. It is seen that, during the early transient time, transients significantly increase void growth rates. This has extremely important ramifications for cyclic loading [3]. Additional key results show that stresses change markedly ahead of the void during the transients, again having important ramifications for cyclic creep conditions. This had been suspected before but had not been proven.

CAVITY GROWTH BY VISCOUS FLOW

A model was also developed to permit the analysis of flow-induced creep of isolated and interacting voids [6]. In contrast to the observations regarding elasticity effects of diffusion

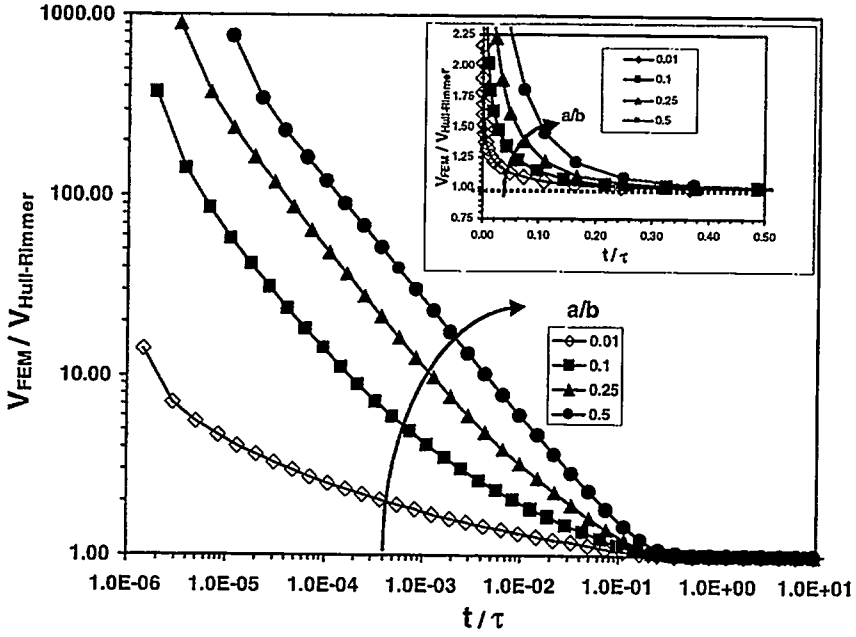

Figure 2. Effect of Elastic Accommodation On Void Growth

controlled cavity growth, creep controlled growth rates were found, somewhat surprisingly, to be only mildly influenced by elasticity, except for high initial void fractions experiencing large triaxial stress states, even under cyclic loading. However, cavity growth rates and cavity aspect ratio growth rates are strongly influenced by inclusion of large geometry effects. Though the void growth rates are somewhat higher for the large initial void volume fraction in a creeping solid exposed to very high triaxialities when elasticity is included, it is unlikely to impact the total time to failure because the growth rates are very high. Additionally, the elastic transient time during void growth is found to be quite negligible for all the cases considered.

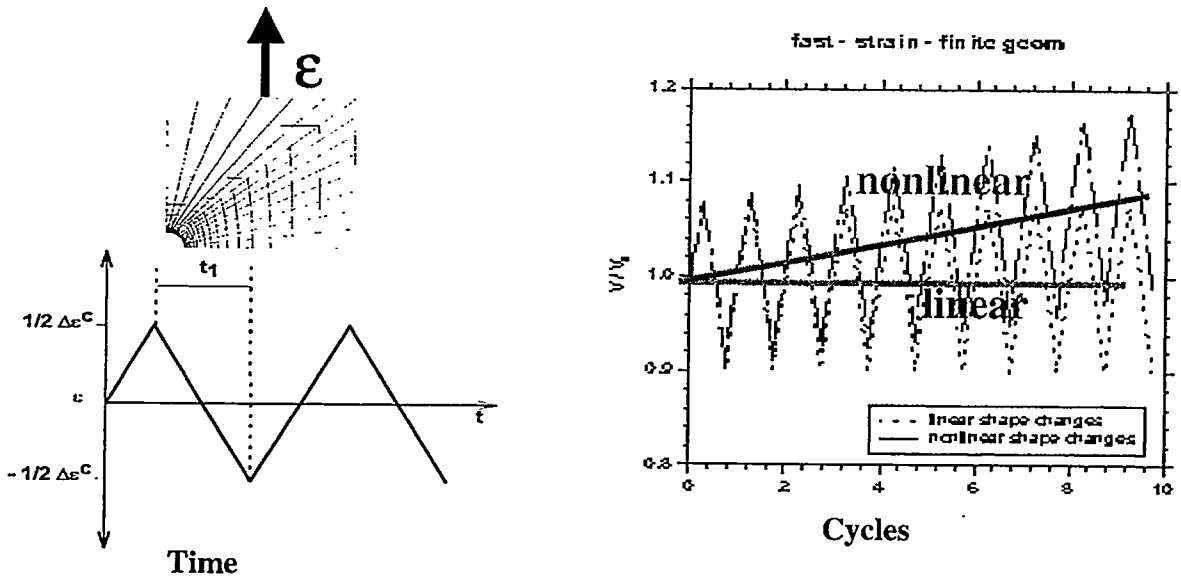

Figure 3. Void Growth Under Balanced Strain-Controlled Cyclic Creep.

Of particular interest is the fact that certain metals experience intergranular cavitation under a balanced cyclic loading condition [7]. Several attempts have been made to explain the phenomenon of intergranular void growth under balanced cyclic loading. A good discussion on these various attempts can be found in Reidel [8]. Some of our recent results are shown in Figure 3. That elastic transient effects after reversal of loading may be an important consideration in understanding this phenomenon was advocated by several investigators (see for instance the overview in [8]) in the past. The present results unequivocally demonstrate that material elasticity does not play any significant role in void growth under balanced cyclic loading if viscous flow dominates. Rather nonlinear shape changes that occur during the balanced cycling process are an important consideration in explaining this phenomenon. This is not surprising since nonlinear shape changes significantly affect the void growth and interaction even under constant stress conditions. The present calculations reveal that the cavity growth rate under balanced cycle loading is constant over the number of cycles performed. Interestingly, this observation is consistent with the experimental findings of Baker and Weertman [7], who show that cavity growth rate is constant in copper experiencing balanced cyclic loading at 678 K.

Recently, Van der Geissen and Tvergaard [9] have performed detailed numerical calculations of void growth under cyclic creep conditions using a plane strain multi-grain model. Their model accounts for interaction of creep and diffusion mechanisms as well as other complex mechanisms such as continuous cavity nucleation, grain boundary sliding, and creep constrained cavitation. In addition to considering balanced cyclic creep, they consider asymmetric cycling as well. Our discussions of these results are confined only to the balanced cycling loading under creep dominated conditions. When creep dominates cavity growth, Reference [9] shows that the damage level steadily oscillates in time around its initial level even in regions where free grain boundary sliding occurs. While this observation is clearly in contrast to our result, it is not surprising that this should be so. One obvious reason is that Van der Geissen and Tvergaard [9] use the BHS model to account for the volumetric growth rate of the cavities under creep dominated conditions. It was shown in [10] that this solution does not assume nonlinear shape changes of cavities. It is emphasized that cavity growth under balanced loading is predicted to be zero for many current models in the literature that were developed neglecting large geometry effects. Under balanced cyclic loading, for high stress triaxialities, we find that the cavities grow in an oblate fashion, favoring void interaction and link-up, in contrast to low stress triaxiality conditions. Again, this is consistent with experimental observations, which show that high constraint greatly reduces creep fatigue life. Finally, it was shown that the failure mode will depend on load type (stress versus strain control). In general, our results are in contrast to those found in the literature, which neglect large geometry effects, especially for cyclic loading at high temperature conditions. Correct trends in void growth predictions are critical for understanding the high temperature failure process and developing solutions to increase life.

COMBINED DIFFUSION AND VISCOUS FLOW

Multi-void, Multi-Grain Model Studies. The next area of study involves the consideration of multiple void growth along a grain boundary. This reduces the effect of periodic boundary assumptions used in the above work since each void along a grain boundary can grow

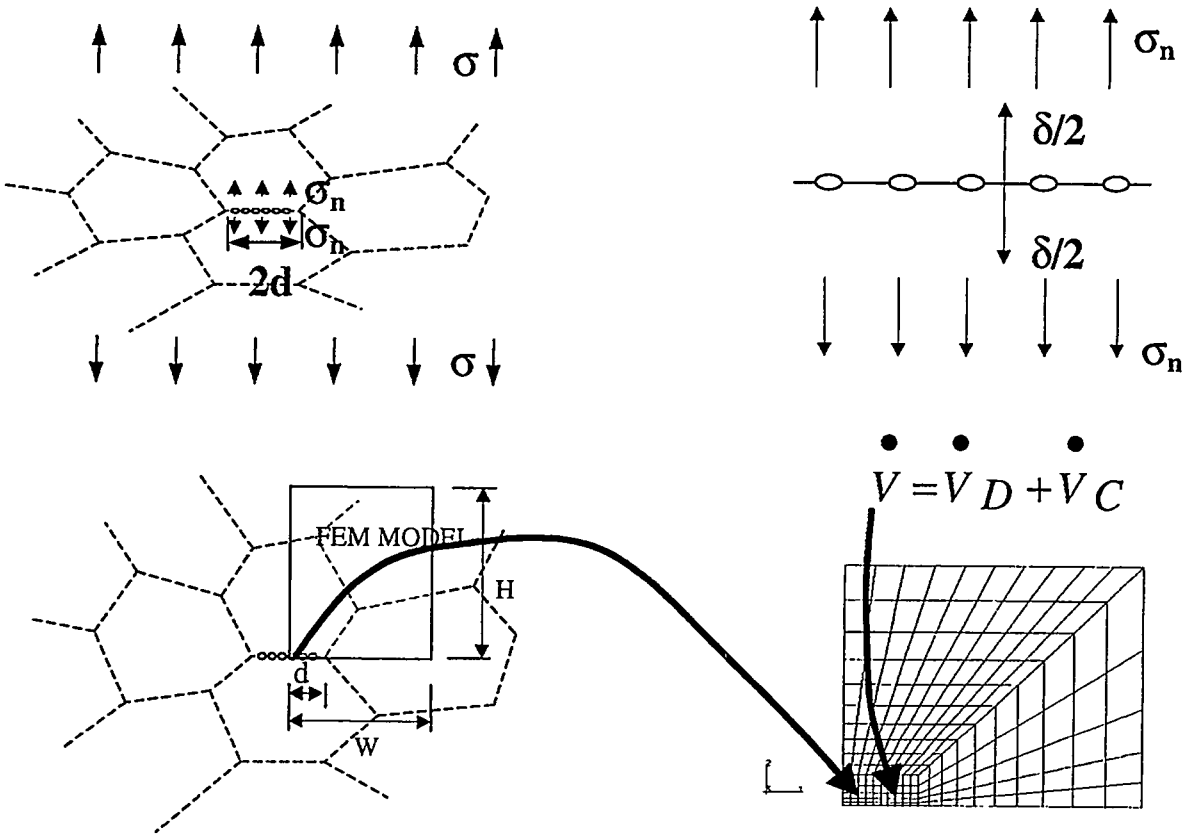

Figure 4. Multiple Void Grain Boundary Model.

independently. This work represents the link in the 2 to 20 μm scale shown in Figure 1. Figure 4 illustrates this concept. The combined effect of diffusion and viscous flow effects of a periodic array of voids within a periodic array of grains can be studied with such a model. This model represents a 'link' between the isolated void at the nanoscale discussed above and the full macro-crack studies, which were investigated in a prior phase of the program.

In Figure 4,

$$\dot{V} = \dot{V}_D + \dot{V}_C$$

represents the diffusion and creep contributions to void growth for each void along the grain boundary (see [11] for details). This model represents an extension of a very recently proposed model by Tvergaard. We have extended this model to permit proper consideration of cyclic load effects. This is done by incorporating a correct cyclic creep constitutive law (Murakami-Ohno (M-O – see [12]) rather than Norton power law creep (N)) and the inclusion of large geometry change effects [11] (which we know are important from our isolated cavity growth studies). One key result is illustrated in Figure 5 where void growth is shown as a function of time for the load spectrum shown. Important differences are clearly observed between the use of different constitutive laws. Moreover, use of small geometry changes, not shown here, shows a very low void growth rate. It is clear that the Norton creep law should be abandoned for investigations of creep fatigue growth. The evolution of cavity growth under strain controlled cycling likewise

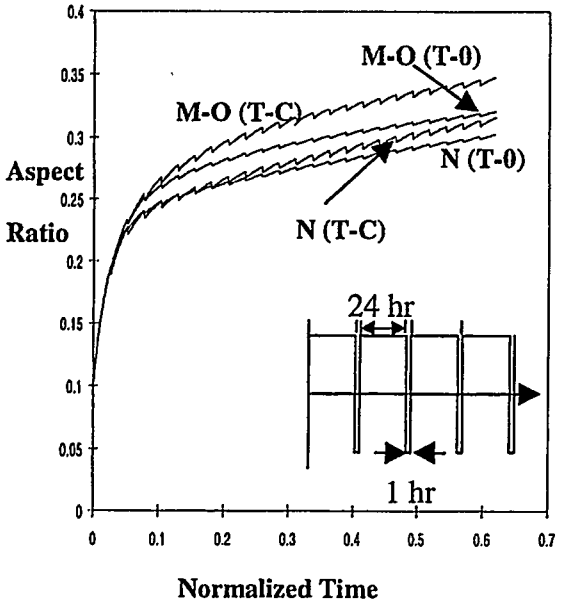

Figure 5. Cyclic Creep Growth Rates.

produces similar results. Other studies of load frequency effects and the importance of shear continue to be studied.

Welds and Residual Stress Effects. A new constitutive law that is appropriate for weld modeling has been developed (see [4]). This new law accounts for weld and base metal melting, softening, and history annihilation, among many other effects that are critical for proper weld modeling. The constitutive law makes weld analyses very rapid. The effects of residual stresses, as may occur from a weld is shown to have an important effect on cyclic creep void growth rates. It is important to note that a compressive residual stress state can greatly reduce void growth rates. The possibility of macro-scale improvement of history dependent creep lives by controlling residual stresses near welds, or from post fabrication heat treatments, is being considered. This model was used to investigate creep life extension procedures for aging power plants by using hot compression. Similar ideas will be studied in relation to nanoscale void growth reduction. Details may be found in [3], [4], and references cited therein.

Future Work. Further investigation of void nucleation and growth using a combined diffusion/creep law under both cyclic and constant loading is ongoing. Cavity growth in the presence of weld residual stresses must be studied as well. Properly including the effects of weld damage has not been performed to date. Finally, and perhaps most important, we plan to link the micro approach, which considers cavity growth at the nanoscale to crack nucleation, and the macro approach, which is used to predict final fracture. This last, difficult step is very important to permit the long-term goal of improving the safety of advanced and aging power generating equipment and other applications discussed above to be achieved. Improvement of the weld process models will continue.

ACKNOWLEDGEMENT

This work was supported by Department of Energy, Office of Basic Sciences under Grant No. DE-FG02-90ER14135. The authors thank Dr. R. Price and B. Armaly for their support.

REFERENCES

1. G. NEWAZ, B. MAJUMDAR, AND F. W. BRUST, "Thermo-Cycling Response of Quasi-Isotropic Metal Matrix Composites,", Journal of Enginering Materials and Technology,Vol. 114, pp 156-161, April, 1992.

2. BRUST, F. W., "Investigations Of High Temperature Damage And Crack Growth Under Variable Load Histories," International Journal Of Solids And Structure, Vol. 32, No. 15, pp. 2191-2218, 1995.

3. BRUST, F.W., 1999, "Classical and Emerging Fracture Mechanics Parameters for History Dependent Fracture with Application to Weld Fracture", PVP-Vol. 393, Fracture, Fatigue and Weld Residual Stress ASME.

4. F. W. BRUST, P. DONG, 2000, "Welding Residual Stresses and Effects on Fracture on Pressure Vessel and Piping Components: A Millennium Review and Beyond", To Appear in the ASME Journal of Pressure Vessel and Piping, Special Millennium Edition, 2000.

5. MOHAN, R., AND BRUST, F. W., "Effect of Elastic Accommodation on Diffusion Controlled Cavity Growth in Metals" to appear in Journal of Pressure Vessel Technology, June 2000 Special Issue.

6. MOHAN, R., AND BRUST, F. W., "An Analytical Study of Void Growth in Viscoplastic Solids", Fatigue and Fracture of Engineering Materials & Structures, Vol. 21, pp. 569-581, 1998.

7. BAKER AND WEERTMAN, J. R, 1990, *Scripta Metallurgica et Materialia*, **24**, pp. 221-234.

8. RIEDEL, H. 1987, "Mechanisms of Creep Rupture," Elsevier Applied Science, London.

9. VAN DER GEISSEN, E., AND TVERGAARD, V., 1995, Micromechanics of Intergranular Creep Failure under Cyclic Loading, *Acta Metallurgica et Materialia*, **44**, pp. 2697-2710.

10. MOHAN, R., AND BRUST, F. W., "On Void Growth in Elastic-Nonlinear Viscous SolidsUnder Creep and Cyclic Creep Conditions", to appear in Journal of Pressure Vessel Technology, June 2000 Special Issue.

11. OH, J., KATSUBE, N., AND BRUST, F. W., "Unresolved Issues With Regard to Creep and Creep Fatigue Life Prediction", ICES'2K, Proceedings of the International Conference on Engineering Sciences, August, 2000, Ed. S. N. Atluri, et al.

12. MURAKAMI, S. AND OHNO, N. (1982): "A Consitutive Equation of Creep Based on the Concept of a Creep-Hardening Surface", *Int. J. Solids Structures*, Vol.18, No.7, pp.597-609

CRACK PATTERNS DUE TO RESIDUAL STRESS IN THIN FILMS

L. B. Freund

Division of Engineering, Brown University, Providence, RI 02912 USA

V. B. Shenoy

Department of Mechanical Engineering, Indian Institute of Technology, UP 208 016 Kanpur, India

ABSTRACT

The physical system studied is a brittle elastic film bonded to an elastic substrate with different properties; a residual tensile stress is presumed to exist in the film. The focus of the study is the influence of the mismatch in elastic properties on patterns of crack formation in the film. The stress intensity factor and crack driving force for growth of a periodic array of cracks in the direction normal to the interface under two-dimensional conditions are determined for any crack depth and any mismatch in elastic parameters. It is found that, even for a relatively stiff film material, the stress intensity factor as a function of crack depth exhibits a local maximum. The driving force for crack extension in the direction parallel to the interface is then determined, and the equilibrium spacing of crack arrays is estimated for given residual stress. The results of the calculations are used as a basis for understanding the crack patterns which have been observed in GaN films on Si substrates.

INTRODUCTION

A range of developing technologies rely upon the deposition of thin films of one material on substrates of another material; examples include thermal barrier coatings, microelectronic devices and functionally graded materials. In many cases, deposition results in significant residual stress in the film. Microelectronic device technology relies heavily on the ability to fabricate devices that are made of thin films bonded epitaxially to lattice mismatched substrates. The mismatch in the lattice parameter causes the film to be severely stressed and induces several stress relaxation mechanisms which can affect the performance and structural integrity of the device. Cracking of the film is a common mode of stress relaxation when the mismatch stress is tensile and the film material is brittle. Examples of such systems include GaN and AlN films on Si substrates. The crack patterns that form in these films usually involve an array of cracks which is roughly periodic. Thouless [1] derived an expression for the crack spacing based on an energy release argument. This argument was later modified to account for the sequential formation of cracks [2]. A slightly different argument, put forth by Hutchinson and Suo [3], gives a spacing between the earlier estimates [1] and [2]. These analyses assumed identical elastic properties for the film

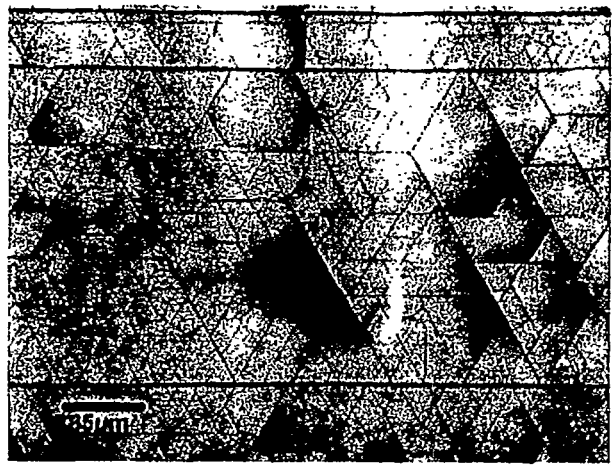

Figure 1: Crack patterns observed in GaN films on a Si substrate. The film is about 5μm thick and the shortest distance between cracks is about two to five times the film thickness.

and substrate. Beuth [4] obtained stress intensity factors for a single crack in a film as a function of the crack depth and the Dundurs [5] parameters which characterize the differences in elastic properties between film and substrate.

The present study is motivated by experimental work conducted at Brown University [6]. Films of (0001) GaN were grown on (111) Si substrates in which the triangular nets between the close-packed planes are coincident. These 5μm thick films developed cracks as shown in Figure 1.

The cracks tend to form on preferred $(1\bar{1}00)$ prismatic cleavage planes which are oriented at 120° with respect to each other due to the anisotropic nature of the surface energy in GaN. Furthermore, there are several "generations" of cracks. The first generation appears to be the deepest and they are spaced at intervals of more than ten to fifteen times the film thickness; these are seen as the thickest lines in the micrographs. Subsequent generations which form are less deep and are spaced at much smaller intervals, typically from two to five times the film thickness. The latter feature is due to the elastic interaction between members of various generations of cracks. The assumption of like elastic properties may be too drastic in the case of GaN/Si system where the elastic modulus of the film is about three times that of the substrate. The aim of this report is to extend the work of [1] and [2] to account for the mismatch in elastic constants between the film and the substrate.

THE MODEL

The film considered has thickness h and is bonded to a relatively thick substrate. The film and substrate have elastic parameters μ_1, ν_1 and μ_2, ν_2, respectively, where μ represents the shear modulus and ν the Poisson ratio. The Dundurs parameters [5] which characterize the difference in elastic properties between the film and the substrate are

$$\alpha = \frac{\Gamma(\kappa_2+1)-(\kappa_1+1)}{\Gamma(\kappa_2+1)+(\kappa_1+1)}, \qquad \beta = \frac{\Gamma(\kappa_2-1)-(\kappa_1-1)}{\Gamma(\kappa_2+1)+(\kappa_1+1)}, \tag{1}$$

where $\Gamma = \mu_1/\mu_2$ and $\kappa = 3-4\nu$. The film is assumed to have a lattice mismatch with respect to the substrate with mismatch strain ϵ_o. This mismatch strain induces a stress $\sigma_o = \bar{E}\epsilon_o$ in a uniform film, where $\bar{E} = 2\mu_1/(1-\nu_1)$ is the plane strain modulus of the film. In the uncracked configuration, the strain energy density in the film is $\sigma_o^2/2\bar{E}$.

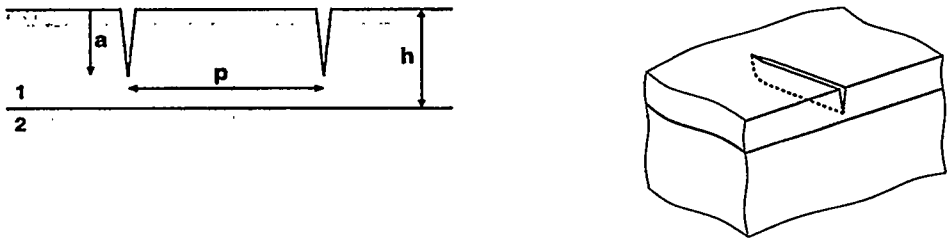

Figure 2: Geometry of a periodic array of cracks in a film bonded to a substrate (left). Inset on right shows a schematic of a channeling crack.

The strain energy developed in the film due lattice mismatch is relieved by means of crack formation. Crack patterns in the film are modeled using a periodic array of cracks. A generic configuration of the cracked film is shown in Figure 2. In such a configuration, the crack depth is a and the crack spacing is p. The task is to determine a and p, given the parameters α, β, σ_o and h.

As long as the crack depth a is less than the film thickness h, the stress field at the crack tips is described by the well known square root singularity, the strength of which is called the stress intensity factor K. When the crack touches the interface, the singularity is a more complicated type. This work does not take this complication into account and attention is restricted to the case when $a < h$. The calculation of energy release rate right up to $a = h$ is not affected by this restriction. When the film is cracked to a depth a, the amount of elastic energy released through formation of one of the cracks is given by

$$\Delta E_e = \frac{\sigma_o}{2}\int_0^a \delta(y)dy = \frac{\pi\sigma_o^2 ha}{\bar{E}} G^*(\frac{a}{h}, \frac{p}{h}, \alpha, \beta) \tag{2}$$

where y is a coordinate along the crack face with the crack tip as the origin and $\delta(y)$ is the crack opening profile over $0 < y < a$. On the other hand, some energy goes into creation of new surfaces and this can be expressed as

$$\Delta E_f = \gamma a = \frac{\pi\sigma_o^2 ha}{\bar{E}}\gamma^* \tag{3}$$

for one crack, where γ is the fracture energy per unit area. Thus, the total change in energy due to formation of a single crack is

$$\Delta E = -\Delta E_e + \Delta E_f = \frac{\pi\sigma_o^2 ha}{\bar{E}}(\gamma^* - G^*). \tag{4}$$

The total free energy of the system per unit interface area in the cracked configuration is

$$U = \frac{1}{2}\mathcal{E}_o h + \frac{1}{p}(-\Delta E_e + \Delta E_f) = \mathcal{E}_o h\left[\frac{1}{2} + \pi\frac{a}{p}(\gamma^* - G^*)\right]. \tag{5}$$

The objective is to determine a and p for a given γ^*.

Several ideas have been proposed to obtain p as a function of γ^*. Thouless [1] obtained this spacing for $a = h$ using the criterion that $\Delta E \leq 0$, that is, the crack spacing is determined by the greatest lower bound on the value of p for which there is a reduction in the free energy. This criterion assumes that all cracks nucleate and propagate simultaneously. Thouless et al. [2] modified this argument to account for the sequential formation of cracks. In this case, the crack

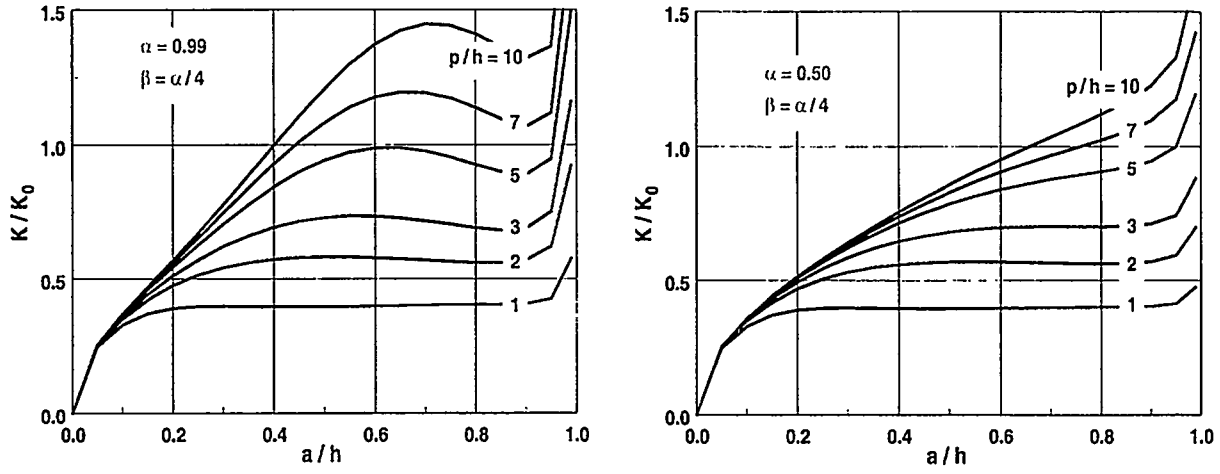

Figure 3: Stress intensity factor normalized by $K_o = \sigma_o\sqrt{\pi h}$ as a function of crack depth a/h for various values of crack spacing p/h and Dundurs parameter $\alpha = 0.99$ and 0.50.

spacing is taken to be that which *minimizes* the free energy U. This gives a spacing larger than the former case and is in better agreement with experiments. Hutchinson and Suo [3] used a different argument in that they consider the lateral propagation or *channeling* of cracks across the film and suggest that the driving force for self-similar lateral propagation or channeling determines the crack spacing. The driving force for channeling can be written as

$$\mathcal{G}_{ss} = \frac{1}{a}\Delta E_e = \frac{\pi\sigma_o^2 h}{\bar{E}} G^*. \tag{6}$$

Thus, the quantity G^* represents the nondimensional driving force for crack channeling. The channeling argument gives a spacing intermediate to the two criteria of [1] and [2].

The objective here is to report results of calculation of $G^*(a/h, p/h, \alpha, \beta)$ which implies values of driving force for channeling as well. The equilibrium crack spacing is obtained by finding the value which minimizes the free energy U. G^* has been determined using a singular integral equation formulation based on the concept of a continuous distributions of dislocations, each of infinitesimal strength [7].

REPRESENTATIVE RESULTS

In this section, the results for stress intensity factor, driving force for channeling, and equilibrium crack spacing are summarized. It is found that the results are relatively insensitive to changes in the value of β and, therefore, results are presented for various values of α but only for $\beta = \alpha/4$.

Stress intensity factors

Figure 3 shows representative plots with $\alpha = 0.50$ and 0.99 of the normalized stress intensity factor under two-dimensional conditions as a function of the crack depth a/h, for various values of crack spacing. The most significant feature to note in these graphs, as well as others included in [9], is the existence of a local maximum in the stress intensity factor as a function of crack depth when $\alpha > 0$. As the crack approaches the interface, however, the stress intensity factor

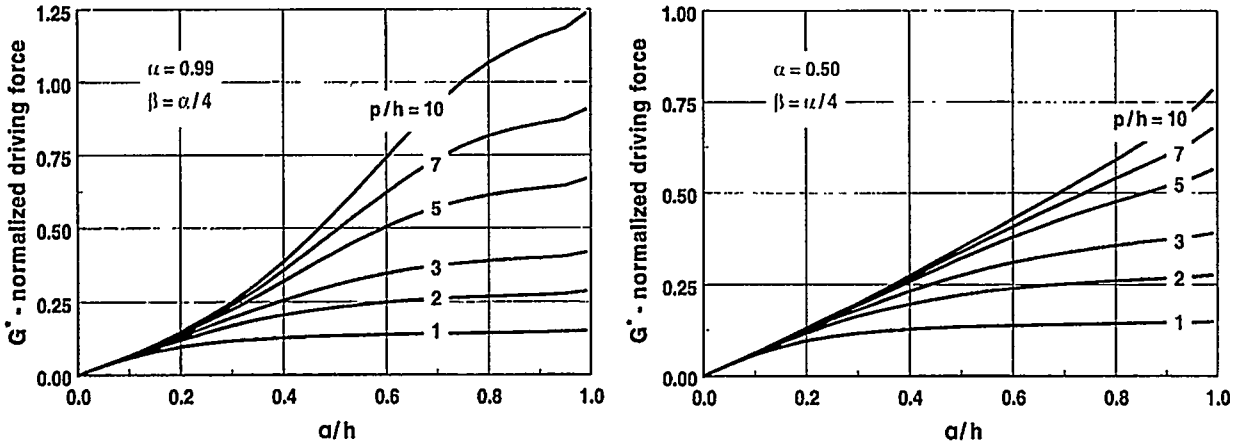

Figure 4: Normalized driving force for channeling as a function of crack depth a/h for various values of crack spacing p/h and Dundurs parameter $\alpha = 0.99$ and 0.50.

invariably increases and it becomes indefinitely large as $a \to h$. As the crack spacing increases, the local maximum in the stress intensity factor vanishes. Also, the crack depth at which the local maximum occurs increases with increasing crack spacing until the maximum disappears. No such maxima are observed for an isolated crack, that is, when $p/h \to \infty$. When the substrate is stiffer than the film ($\alpha < 0$), the stress intensity factor exhibits a local maximum, and it then approaches zero as the crack edges approach the interface. This is similar to the result of [4] for a single crack.

Driving force for channeling

Figure 4 shows representative plots with $\alpha = 0.50$ and 0.99 of the normalized driving force for channeling as a function of the crack depth a/h for various values of crack spacing The most important feature in these results is that, as the crack spacing p/h decreases, the driving force decreases. For crack spacing p/h between 1 and 3 the driving force initially increases with crack depth, and subsequently attains an almost constant value for larger crack depths. In cases where the substrate is stiffer than the film, the driving force has a shallow maximum as the crack approaches the interface. This effect is also observed by Beuth [4] for a single crack.

Equilibrium crack spacing

The equilibrium crack spacing is obtained by finding the crack spacing that minimizes the free energy. In fact, both the crack depth and spacing are to be determined by the minimization of free energy. It turns out from the calculations that the free energy is always minimized when the crack approaches the interface, and therefore the crack spacing p/h that minimizes U as defined in (5) is determined with a/h set equal to 0.99. The results of the calculation shown in Figure 5 provide equilibrium crack spacing as a function of σ^*, the normalized measure of mismatch stress. For a given value of σ^*, the crack spacing for a relatively stiff film is larger than that for a relatively compliant film. It is noted that the equilibrium curves for $\alpha = -0.75$ and -0.99 end abruptly at about $\sigma^* = 1.8$. This is due to the fact that, in this case, the minimum value of energy of the cracked film is larger than for an uncracked film for normalized stress less than σ^*. In other words, cracking is not possible for smaller values of σ^*. It is also clear that this critical value of stress, below which fracture cannot occur, decreases with increasing α.

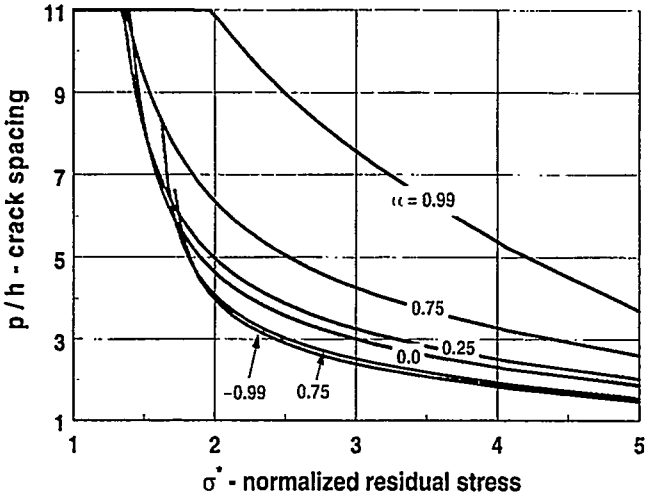

Figure 5: Equilibrium crack spacing versus normalized mismatch stress $\sigma^* = \sigma_o\sqrt{\pi h/\bar{E}\gamma}$.

Crack opening

In plan view photographs of cracked films, cracks are seen as dark lines, as illustrated in Figure 1. It seems reasonable to associate the thickness of these lines with the amount of crack opening on the film surface. This opening can be calculated from the solution of the integral equation for dislocation density. Figure 6 shows plots of $\bar{E}\delta(a)/\sigma_o h$ versus a/h. The most interesting feature to be observed in this figure is that, when α is positive and very close to 1, normalized displacement has a local maximum at some value of crack depth a/h. Thus, for very stiff films, thicker lines on the micrograph do not necessarily imply deeper cracks! This feature, although present, is not significant at lower α.

DISCUSSION

The foregoing results indicate that the elastic mismatch between the film and the substrate can significantly influence behavior such as that of the stress intensity factor as a function of crack depth for a fully formed crack array. The local maximum which arises in the stress intensity factor as a function of crack depth may be understood by the following arguments. The stress intensity at the tip of each crack is affected by the presence of all other cracks in the array. The presence of all other cracks except the one of interest may be taken into account by considering their influence on the film. The main effect of the presence of the cracks is that the film effective modulus is lowered. The reduced effective modulus $\bar{E}_{eff}$ can be estimated to be

$$\frac{\bar{E}_{eff}}{\bar{E}} = 1 - \frac{2\pi a}{p} G^*(\frac{a}{h}, \frac{p}{h}, \alpha, \beta). \tag{7}$$

It is noted that this approximate formula [2, 3] neglects the effect of the substrate. It is clear that, as the crack depth increases, the effective modulus falls. Furthermore, the fall is larger for a larger G^*. It is clear from Figure 4 that, for a given geometry of cracks, G^* is larger for larger α and, therefore, this effect must be greater for stiffer films. These approximate arguments are in

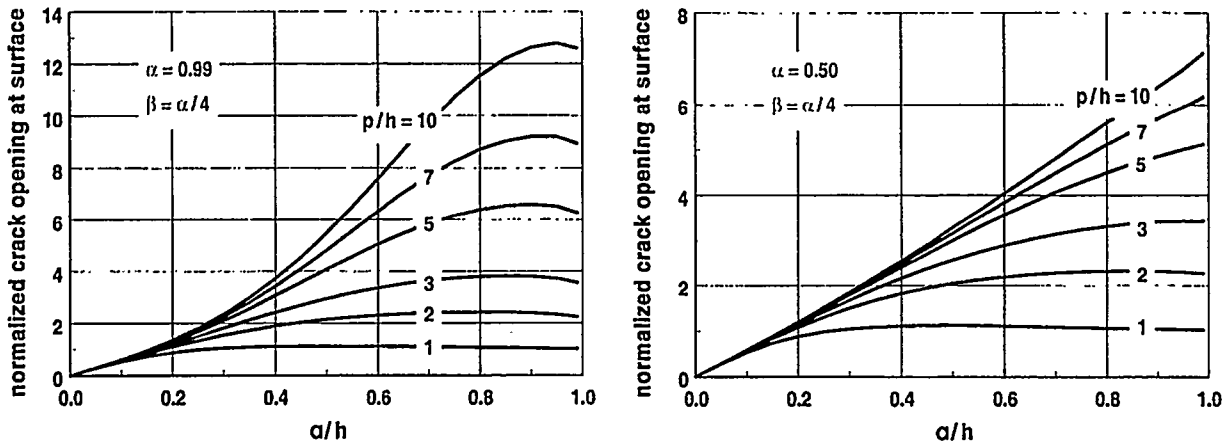

Figure 6: Normalized crack opening $\bar{E}\delta(a)/\sigma_o h$ versus crack depth a/h for various values of crack spacing p/h and Dundurs parameter $\alpha = 0.99$ and 0.50.

agreement with the calculations. It is, however, to be noted that as the crack depth approaches the film thickness, the assumption used to derive (7) breaks down in that the role of the substrate becomes increasingly significant.

Turning now to equilibrium crack spacing, it is clear that very stiff films will have cracks spaced further apart. Also, the least normalized stress σ^* for which fracture is possible decreases with increasing stiffness of the film. Thus, cracks are more easily (in nondimensional terms) nucleated in stiff films but their density is smaller. On the other hand, more compliant films require larger σ^* for the formation of cracks. After this stress is attained, the crack density is higher than for stiffer films.

The features observed in the experimental photograph in Figure 1 may be explained as follows. The threefold symmetry occurs due to the fact that the surface energies of these three crystallographically equivalent planes is the lowest, so they are the preferred fracture planes. The stresses in this system are generated by a combination of lattice mismatch and mismatch in thermal expansion coefficient. As the sample is cooled from its growth temperature, the tensile stress in the film increases. When the critical stress for the formation of cracks is attained, the first generation of cracks forms; these are more than ten to fifteen film thicknesses apart and they are as deep as the film thickness. As the film cools further, select new generations of cracks channel across the film in the gaps between individual cracks in the previous generation, in an attempt to attain the equilibrium crack density. When the crack spacing reaches about two or five times the film thickness, the driving force for channeling does not increase significantly with crack depth, and in addition, the stress intensity factor falls with increasing crack depth. Thus, the last generation of cracks that forms is not as deep as the film thickness. These arguments are meant to provide only a qualitative understanding of the cracking patterns. The quantitative details of these patterns are obviously affected by the biaxial state of stress in the film and by the three dimensional nature of the cracks which is not taken into account in the present analysis.

CONCLUSIONS

This report is aimed at extending previous work on crack patterns in brittle thin films to account for the mismatch in elastic properties between the film and substrate. The main conclusions

of this investigation are:

- The stress intensity factors at the edges of a periodic array of cracks under two-dimensional plane strain conditions shows a local maximum for small enough values of the periodic spacing of cracks. This effect is stronger when the film is stiff relative to the substrate.
- The driving force for channeling cracks falls with increasing crack density. For small crack spacing, the driving force for channeling increases initially with crack depth, but attains an almost constant value for greater depth.
- The equilibrium crack spacing is larger for a relatively stiffer film at a given value of the nondimensional mismatch stress. It is easier to nucleate cracks in a more stiff film (in nondimensional terms) than in a more compliant film. More compliant films allow for cracks only at larger values of the nondimensional mismatch stress, with a larger crack density.

REFERENCES

1. THOULESS, M. D., "Crack spacing in brittle films on elastic substrates," *Journal of the American Ceramic Society* **73**, 2144 (1990).

2. THOULESS, M. D., OLSSON, E. and GUPTA, A., "Cracking of brittle films on elastic substrates," *Acta Metallurgica et Materialia* **40**, 1287 (1992) .

3. HUTCHINSON, J. W. and SUO, Z., "Mixed mode cracking in layered materials," *Advances in Applied Mechanics* **29**, 63 (1992).

4. BEUTH, J. R., "Cracking of thin bonded films in residual tension," *International Journal of Solids and Structures* **29**, 1657 (1992).

5. DUNDURS, J., "Edge-bonded orthogonal elastic wedges," *Journal of Applied Mechanics* **36**, 650 (1969).

6. STEVENS, K. S., OHTANI, A., SCHWARTZMAN, A. F. and BERESFORD, R., "Growth of group III nitrides on Si(lll) by plasma-assisted molecular beam epitaxy," *Journal of Vacuum Science and Technology* **812**, 1186 (1994) .

7. FREUND, L. B. and KIM, K. S., "Spiral cracking around a strained cylindrical inclusion in a brittle material and implications for vias in integrated circuits, Materials Research Society Symposium Proceedings 226, 291 (1991).

8. SUO, Z., "Singularities interacting with interfaces and cracks," *International Journal of Solids and Structures* **25**, 1133 (1989).

9. SHENOY, V. B., SCHWARTZMAN, A. F. and FREUND, L. B., "Crack Patterns in Brittle Thin Films," *International Journal of Fracture*, to appear.

OBSERVATION OF CLEAVAGE FRACTURE AFTER SUBSTANTIAL DIMPLE RUPTURE IN ASTM A710 STEEL

W. G. Reuter and W. R. Lloyd

Idaho National Engineering and Environmental Laboratory
P.O. Box 1625, MS 2218
Idaho Falls, ID 83415-2218

ABSTRACT

A major concern often arising in structural integrity predictions is the possibility that low-energy brittle fracture could result as a consequence of cleavage either under normal operating or design accident conditions. This can be especially troublesome when the leak-before-break (LBB) approach shows an additional safety margin of the design. For LBB to be applicable, the fracture process must remain ductile (dimple rupture), and not change to cleavage. The American Society for Mechanical Engineers Boiler and Pressure Vessel Code (Code) provides guidelines for avoiding cleavage fracture for Code-accepted materials. Experimental results for a non-Code steel are provided, and show that cleavage may occur for a thickness under16 mm (where the code suggests it will not) after stable crack growth (Δa) of up to 20 mm. This work is still in progress; test results are provided along with possible reasons for the mode transition, but complete explanations are still being developed.

INTRODUCTION

Cleavage fracture during crack growth initiation generally results in sudden, catastrophic structural failure, and must be avoided! Cleavage fracture after stable ductile crack growth is also a concern, but this transition has only been observed in limited cases where crack growth was less than 2 mm.[1,2] However, the possibility of a sudden fracture after stable crack growth has significant practical concerns. Steel structures, including pressure vessels, should certainly incorporate a fracture-resistant design. If cracks should appear and grow, the mechanism should be hole growth (ductile) rather than cleavage (brittle) to limit crack velocities and fragmentation. The need to reduce chances of cleavage fracture led to development of design procedures to prevent unexpected, sudden failure of structural components. Welding Research Council Bulletin 175[3] provides these procedures, based on the requirement that the minimum operating temperature will exceed the material's "transition temperature." This corresponds to a fracture mechanism change from cleavage (low absorbed energy) to ductile hole growth (high absorbed energy). The approach requires drop weight tests (per ASTM E208-69), when geometrically possible, to obtain the transition temperature (T_{NDT}), and Charpy V-notch (CVN) impact tests when the material thickness exceeds 16 mm.

NB-2332 of the Code[4] provides guidelines for the minimum operating temperature relative to CVN impact results. Minimum specified CVN values are required at test temperatures less than or equal to the lowest desired material service temperature. Thin materials (<16 mm) have no requirement, and

intermediate thickness have a minimum CVN lateral expansion requirement. Plates thicker than 63.5 mm have a minimum absolute operating temperature (RT_{NDT} + 56°C) specified. Thus, the ability to design structural components to prevent non-ductile fracture appears to be established simply by following these requirements.

The Leak-Before-Break (LBB) approach to fracture-resistant design was originally developed by Irwin[5] to establish the required fracture toughness to prevent crack growth when a material was loaded to the yield strength. In the late 1970s, LBB was extended (required fracture toughness such that a through-thickness crack, twice as long as the material thickness, will not extend unstably – corrosion-enhanced growth excluded). This suggested that, if a crack penetrated the wall thickness of a vessel pressurized with a fluid, and the crack length was less than twice the wall thickness, then the vessel would simply leak, hence, "leak-before-break." Further, such a vessel would not experience catastrophic failure until additional crack growth occurred. During the time that the crack leaks, it is assumed that the leak will be detected and remedial measures performed before the crack reaches a critical size. Therefore, by following this hypothesis, it would be possible to predict conditions where LBB would apply. In those cases, LBB can be a safety mechanism of last resort to detect a significant crack before unstable fracture occurs. The LBB concept of pressure vessel safety is now used throughout the world.

This paper provides experimental results from surface crack test specimens (simulating the fracture behavior of structural components). Cleavage fracture occurred in the experiments when the Code sugests it will not. Through careful examination and critical analysis of these results, we hope to better understand factors controlling cleavage fracture following stable ductile tearing. This will help provide solutions to practical problems associated with applying LBB concepts to structural safety.

MATERIAL CHARACTERIZATION

The A710 steel used in this study is not explicitly qualified by the Code. But, the A710 steel showed the unexpected cleavage transition behavior of interest, and a large database of material properties and experimental data already exist that will aid in analyzing the phenomenon. Therefore, for purposes of comparison and analysis, the intent of the Code relating to material properties was applied.

The thickness of the A710 steel plate was 31.8 mm. The specimens containing surface cracks were either 6.4 or 12.7 mm thick, and they were removed from the central plate thickness. The chemistry of the A710 steel is: 0.05 C, 0.47 Mn, 0.010 P, 0.004 S, 0.25 Si, 0.74 Cr, 0.85 Ni, 0.21 Mo, 1.20 Cu, and 0.038 Cb with an ASTM grain size of 8. ASTM E208,[6] P-3 NDT tests were performed on the 31.8 mm thick steel plate. The results showed that T_{NDT} is -18°C (0°F) for plate C12188. CVN impact tests were performed per ASTM E23[7] with the specimens oriented in the transverse (T) direction.

A number of specimens containing surface cracks with a/t ranging from 0.15 to 0.85 and a/2c ranging from 0.1 to 0.5 were tested. Data collection included: stop-action photographs at the front and back surfaces; acoustic emission; applied force, crosshead displacement, and crack mouth opening displacement (CMOD); and the applied force when the growing crack penetrated the opposite surface. (penetration was identified using a rubber air bulb held to the back surface by vacuum; it fell off when air passed through the opening at crack penetration, releasing the vacuum.)

TEST RESULTS FOR SURFACE CRACKED SPECIMENS

We have observed that surface cracks grow predominantly through the thickness direction, with little or no growth in the plate width (2c) direction. Once the crack penetrates the back surface, it then

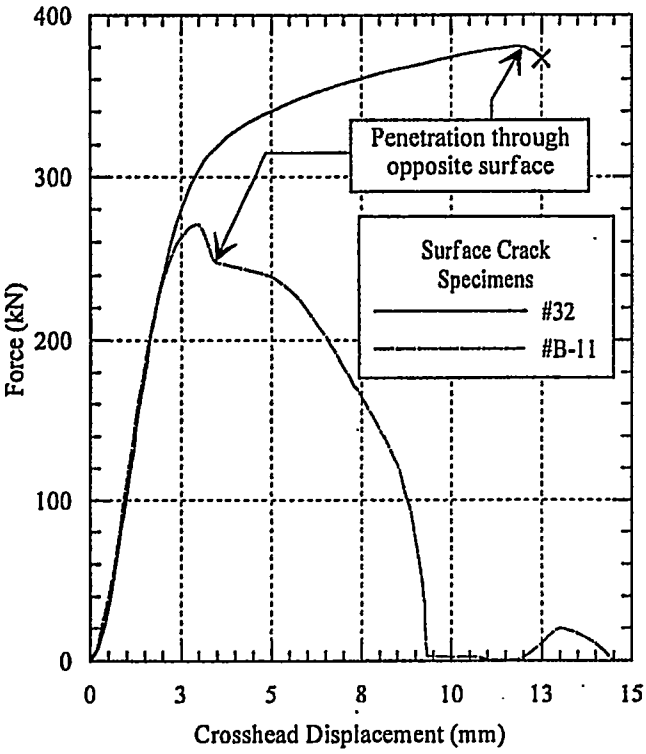

Figure 2. Test record for two surface crack specimens.

grows until the length at the back surface is the same as the length at the front surface. The crack growth to this point is essentially dimple rupture occurring in the plane of the original fatigue precrack. As the test progresses, the specimen configuration is analogous to a middle crack [M(T)] specimen. The crack growth then transforms to a single- or double-slant fracture. Stable crack growth continues until either the crack tip reaches the specimen edge, or sudden cleavage fracture results in catastrophic specimen failure. Figure 1 shows force vs. crosshead records for Specimen 32 (experienced cleavage fracture at both crack tips) and for Specimen B-11 (did not experience cleavage). Figure 2 shows the fracture surfaces of specimens (#15, #B-42) that experienced cleavage fracture. It is apparent that the amount of stable crack growth preceding catastrophic failure varied considerably (ranging from a few mm, up to 20 mm), and that the transition from ductile fracture to cleavage was abrupt. The fracture surfaces of these specimens were examined by light microscopy, and with SEM, and only a few islands of cleavage were detected behind the main ductile/cleavage boundary. These isolated islands were small (~30 μm) and occurred infrequently. Figure 2 suggests that there is almost a straight line of transition separating dimple rupture from cleavage.

Recently, another series of specimens (B-24, B-29, B-43, C-6, C-23, and E-14) were tested. Of the six specimens, four experienced cleavage (B-29, B-43, C-23, and E-14). Figure 3 shows the fracture surface for Specimens B-29 and B-43.

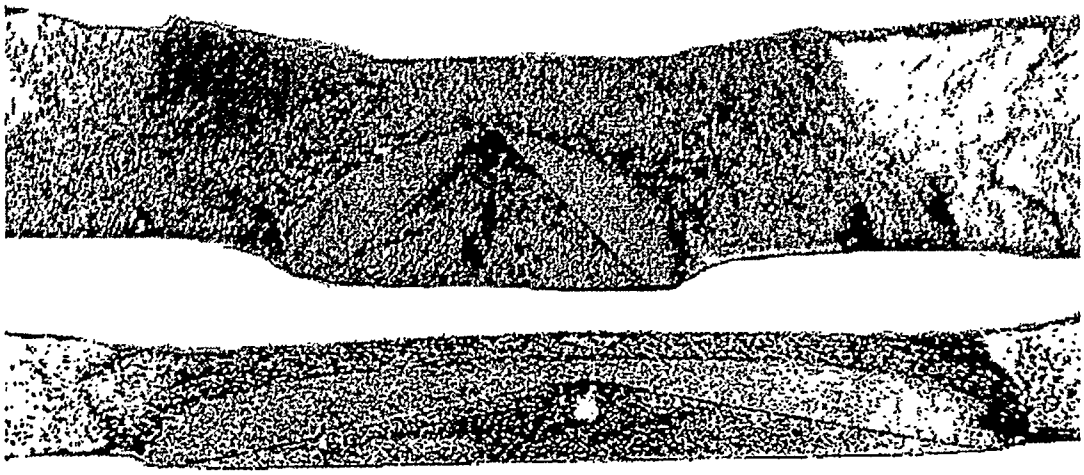

Figure 1. Fracture surfaces of specimens #B-42 (above, 12.7 mm thick) and #15 (below, 6.4 mm thick) tested in mid-1980s. Transition from ductile (darker) to cleavage (brighter) fracture is visible. Note chevron shape of transition boundary at left on #15.

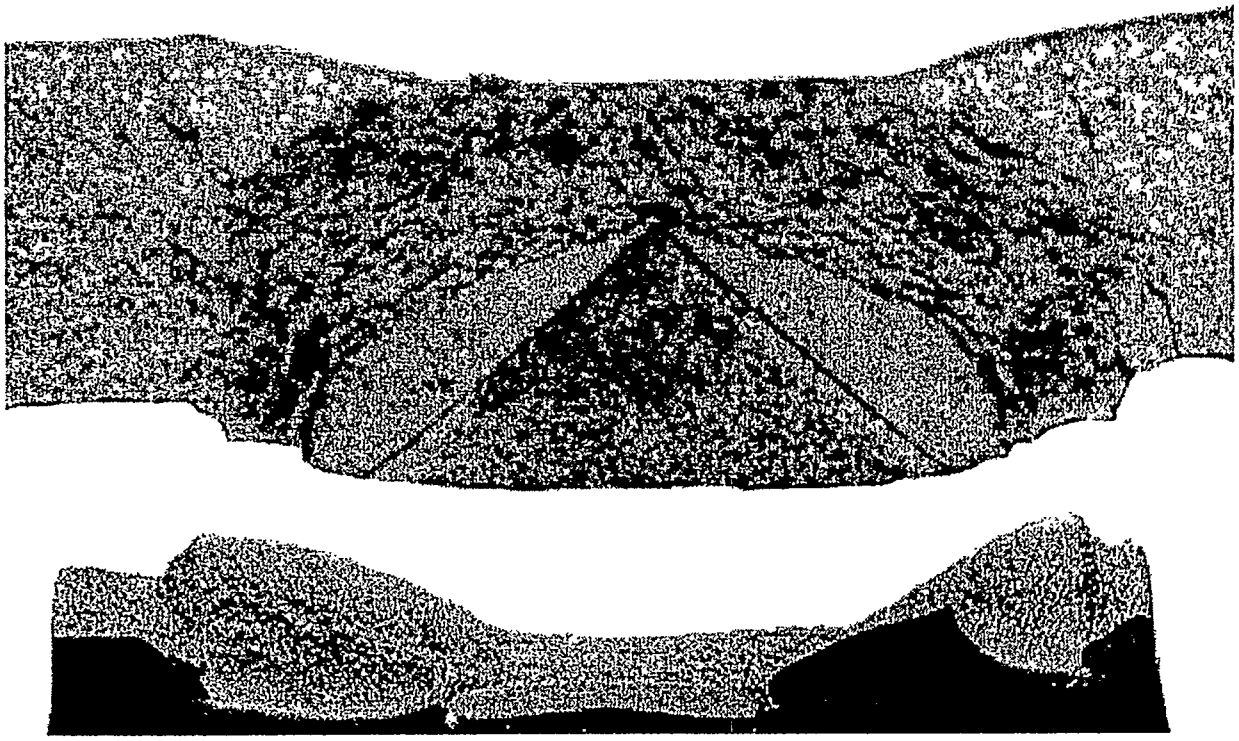

Figure 3. Specimen B-43 (above) and B-29 (below). Darker, crescent-shaped region extending to back surface on B-43 is ductile fracture region. Cleavage on B-29 is isolated to far left and right.

DISCUSSION

CVN Results: T_{NDT} is -18°C (0°F) for plate C12188 of A710 steel. The steel's 0.89 mm lateral expansion at 24°C easily exceeds the Code requirement of 0.64 mm for materials up to 38.1 mm thick. Therefore, the test results at 24°C determined the minimum "operating" temperature. At 24°C the CVN impact energy results for the transverse-oriented (L-T) specimens were 62, 62, and 87 J with corresponding lateral expansion of 0.94, 0.94, and 1.24 mm. Therefore, applying the "intent" of the Code, the minimum operating temperature was set to be 24°C. Reference 3 notes that energy absorption (higher measured CVN energy) may be increased by increasing material yield strength or ductility. Of these, ductility (associated with CVN lateral expansion) is a better indicator of fracture toughness change. For this material, the measured lateral expansion of 0.94 mm exceeds the Code-required values, and is a positive indicator that the material meets the minimum Code-inferred fracture toughness. Another item of possible concern is that the CVN specimens had only 10% shear fracture area at 24°C. However, this parameter is not considered in the ASME Code.

Surface Cracked Specimens: Figures 2 and 3 clearly show that cleavage did not occur until after the surface cracks had penetrated the back surface and were growing in the plate width (2c) direction. In many instances, the cleavage fracture initiated from 45 deg ductile, slant fracture, e.g. #B-29. In one of the fracture surfaces studied extensively in the SEM (Figure 2, B-42), there is a distinct, uninterrupted boundary separating ductile and cleavage fracture zones. There are only a few islands of cleavage in the dominantly ductile region, and they are located close to the transition boundary. It is commonly accepted that the probability of cleavage fracture initiation is modeled by a Weibull-type continuous probability function. The model uses a critical cleavage stress and a representative material length as its governing parameters for a particular material. The probability of cleavage fracture initiation, at any location in the

volume being considered, for a given crack boundary and stress state in that same volume, is calculated through volume integration of principal stresses near the crack boundary. For a given material, a longer crack boundary (specimen/crack size effect) or higher principal stresses (material strength/hardening influence and applied loading) can increase the probability of cleavage fracture initiation. This model is limited to predicting the probability that cleavage fracture will initiate at any instant at some point within the volume of material considered. The so-called weakest link model of macroscopic cleavage fracture assumes that an individual initiation will always progress to a macroscopic fracture event. However, if cleavage fracture does initiate, whether or not it will continue beyond the grain in which it initiates is not an explicit model output, since the material stress state has changed. In addition, the model does not explicitly predict the likelihood that cleavage fracture will initiate simultaneously at multiple locations in the volume under consideration.

Gerberich et al.[8] extend this probablistic cleavage initiation concept to macroscopic cleavage crack growth. They suggest that sustained cleavage crack growth (following initiation) may very likely be the result of multiple cleavage initiation events occurring almost simultaneously. The various distinct cleavage islands then grow together to sustain the macroscopic cleavage fracture process. In effect, they acknowledge the validity of the Weibull model for cleavage initiation, but reject the weakest-link assumption for continuing macroscopic fracture. Gerberich's suggested crack growth process was strongly supported by INEEL's surface crack test results. For example, extensive examination of the fracture surface of specimen B-42 showed "river patterns" (chevron-shaped features) on the cleavage region pointing back to numerous initiation sites along the ductile/cleavage fracture boundary. Other specimens show isolated islands of cleavage fracture embedded within the ductile fracture region. These are sites where cleavage initiated, but most likely arrested due to insufficient driving energy. Certainly, the weakest link assumption provides a useful boundary to application of the Weibull PDF where cleavage fracture is to be avoided. However, it is a rather severe restriction when it comes to fracture process modeling. The real crack growth process observed in experiments is not allowed by the "weakest link" stipulation.

It may be possible to extend the Weibull probability model to estimate the likelihood of a macroscopic cleavage fracture event arising from essentially simultaneous multiple initiations. But, the complexities of the process make the job enormous. The fracture process volume must be discretized to small volume elements, and the associated stresses determined using FEM. Then, an iterative probability simulation considering all possible outcomes must be performed. We hope that the INEEL microtopography fracture process analysis system will provide additional information on crack boundary locations at various points in the fracture process. This will eliminate a very tedious and still very difficult 3-D plastic crack growth modeling process, and reduce some uncertainty in the model associated with predicting crack growth. Rather than calculating crack growth increments based on some estimated criteria (which may not match the actual specimen behavior), incremental crack boundary positions can be specified based on experimental data. We also hope to establish at what point (crack boundary position, applied remote load) some of the isolated cleavage islands were generated. This will be accomplished through the correlation of engineering test data (force, displacement, COD, etc.), acoustic emission monitoring data, electric potential crack growth monitoring data, and microtopography-determined crack boundary locations.

At the present time, test data exists from a variety of A710 steel surface crack specimen geometries. Some specimens exhibited sudden catastrophic failure after little ductile crack growth, others after a large amount of ductile growth, while others experienced plastic instability with no cleavage at all. Various test parameters, such as test machine and specimen compliance, crack growth rate, specimen thickness, and initial surface flaw configurations were examined for correlation with the crack growth behavior. No correlation with observed crack growth behavior was found.

Specimen E-14 (a recent surface crack test) exhibited a double 45 deg slant fracture with a central flat region that was nominally triangular in shape, with mostly ductile fracture appearance. Point sources

of specular light reflection from small cleavage facets were observed in this flat region. However, considerable ductile crack growth occurred before the transition to cleavage. Specimen B-24, also tested recently, exhibited a single 45 deg slant fracture. Light reflection from a section that went completely through the thickness (see Figure 4) shows that cleavage fracture occurred in this region. These types of cleavage regions were not observed in the first series of surface crack tests performed in the mid-1980s. We presently have no guess why these regions of cleavage did not cause catastrophic failure of the test specimens. We also do not have a qualified explanation for the difference in fracture behavior of recent tests compared to those performed some 15 years ago.

Our best hypothesis at present includes two inter-related factors for the surface crack specimens. First, crack tip constraint varies significantly during the fracture process. Second, plastic deformation of the material far ahead of the original crack (pre-straining) creates an effective yield strength increase before the fracture process zone reaches those locations (also assisting in elevation of local constraint). Reuter et al.[9] show that differences in constraint exist between specimens containing surface cracks and SE(B) specimens. In all surface crack specimens (that had some cleavage) that have been examined, the cleavage events initiated in the central portion of the specimen thickness. Constraint is known to be higher there than nearer the specimen surface. Further, the maximum local stress intensity, K (or local J-integral), obtained from FEM analysis methods for surface cracks, does not occur at a corresponding location. Crack initiation studies indicate that local maximum driving force, $K^{(9)}$ (or $J^{(2)}$), alone do not dictate the crack initiation location, but that some (yet to be defined) critical combination of crack driving parameter and constraint measure (e.g. K and T or J and Q)[9,2] control the fracture initiation process.

We also noted that cleavage fracture did not occur until the surface crack penetrates the opposing surface. Reduced constraint on the crack perimeter due to plastic flow to the opposing surface is a likely cause. The critical stress level required to initiate cleavage cannot, therefore, be reached. Once the crack penetrates the thickness and is growing parallel to the plate surface, the local constraint probably increases due to several factors[1] until the critical condition for cleavage initiation can be reached. Green and Hundy[10] note that pre-straining a material increases its yield strength, causing the ductile-

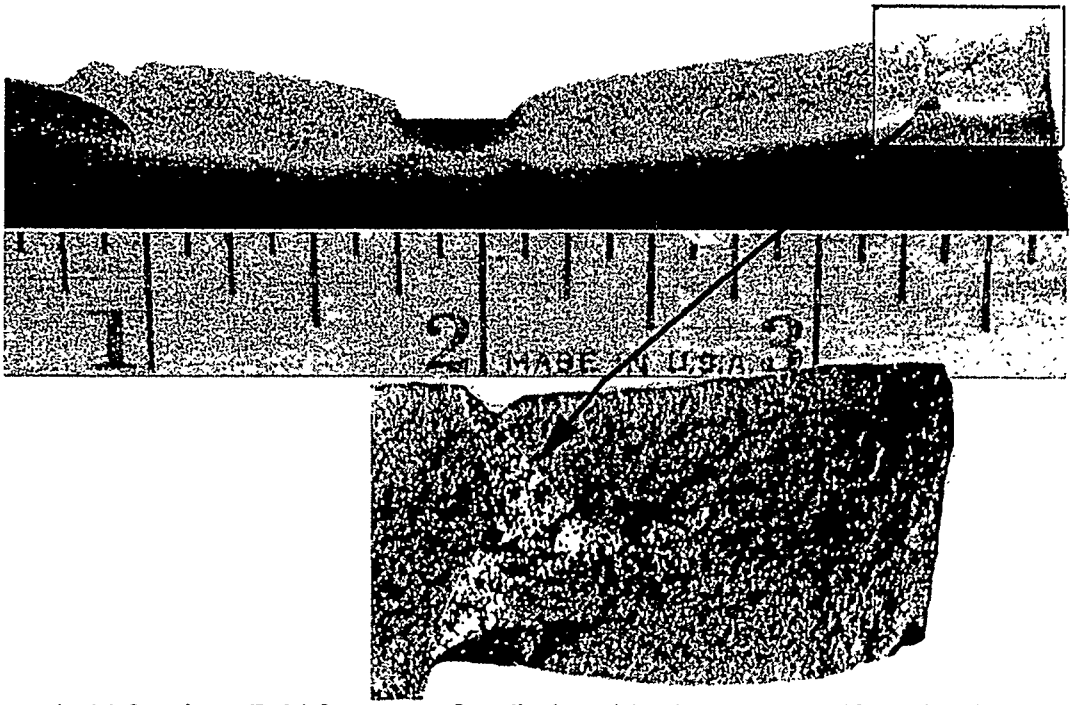

Figure 4. (a) Specimen B-24 fracture surface (inch scale). (b) Inset, magnified, showing band of cleavage traversing the specimen thickness.

brittle transition temperature to increase. At a given temperature, this will increase the probability of brittle cleavage fracture occurring. M(T) plate specimens of the same A710 steel (tested in the mid-1980s) do not undergo as much material pre-straining as the surface crack specimens, and they never had a cleavage fracture. This infers that the level of material pre-straining (in the surface crack specimens) plays some role in the ductile-to-cleavage fracture mode change. Dodds et al.[2] also note that crack tip meandering (out of the nominal crack plane) associated with ductile crack growth has a substantial effect on the stress fields at the crack tip.

The difference in behavior (observed cleavage regions) between specimens tested recently and those tested about 15 years ago may be explained by a change in the material's sensitivity to constraint. The percentage of cleavage fractures experienced in the recent test series appears to be about twice that of the earlier tests. This difference may be real which would mean that embrittlement occurred while the test specimens have been in storage at nominally 24°C. This may be due to a room temperature aging phenomenon, which we feel is unlikely, but that is yet unproven. To evaluate this possibility, the grip section (only elastically strained) of some of the recently tested specimens were machined into CVN specimens. They were impact tested, and lateral expansion exceeded 1.6 mm at 24°C. Comparison of recent CVN test results with original material certification data (1981) suggests that: T_{NDT} is unchanged; the transition region slope is now steeper (narrower temperature range); and fracture energy (CVN) is elevated. These results, obtained just a few days ago, are contrary to what was expected based on the recent surface crack tests. Because of these conflicting results, we have no explanation for the change in observed behavior of the surface crack specimens. We are planning additional tests and analyses to unravel the mystery of this apparent material change.

At this time, we can draw the following conclusions:

(1) Catastrophic cleavage fracture can occur after a surface crack has penetrated the opposing surface when the plate thickness is less than 16 mm, even after ductile crack growth (by hole growth/dimple rupture) has progressed over 20 mm;

(2) Catastrophic cleavage fracture can occur even when the test temperature equals a temperature where the Code requirement for minimum lateral expansion in a CVN test is exceeded;

(3) Comparison of Charpy impact (CVN) test results, from the same heat of material but measured 18 years apart, show substantial differences. The more recent results suggest cleavage fracture should not occur, while the surface crack tests result in more observed cleavage;

(4) We cannot explain the sudden transition from ductile fracture to cleavage that occurs in the surface crack specimens, but recommendations for future work to study the phenomenon are as follows:

RECOMMENDED ACTIONS

(1) Calculate (via FEM) the stress fields, and corresponding constraint, at the front of the growing cracks for C(T), SE(B), and surface cracked (SC) specimens. Experimental methods will be used to establish incremental crack front positions, CTOD, and CTOA for input to the FEM;

(2) Test additional surface crack specimens (crack geometries and plate sizes) to replicate all configurations that were tested in the mid-1980s for comparative purposes;

(3) Tests of middle crack plate [M(T)] specimens for comparison with recent surface crack test results, and with earlier M(T) and surface crack results;

(4) Perform tensile tests of the A710 steel over a range of temperatures for comparison with original tensile data (1981);

(5) Use tensile or notched bend specimens to measure the critical cleavage stress of the A710 in its present condition; and

(6) Continue SEM studies to identify structure features responsible for initiation of cleavage fractures.

These tests and analyses will help us understand the substantial differences between the earlier (1981) and recent (2000) CVN test results, and the role these differences play in the sudden transition from ductile fracture to cleavage fracture.

ACKNOWLEDGEMENTS

This work was supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, Engineering Research, Under DOE Idaho Operations Office Contract DE-AC07-99ID13727.

REFERENCES

(1) Faleskog, J. (1995), Effect of Local Constraint Along Three-Dimensional Crack Fronts – A Numerical And Experimental Investigation, **J. Mech. Phys. Solids, 43**(3), pp. 447-493.

(2) Dodds, R. H., Jr., M. Tang, and T. L. Anderson (1993), *Effects of Prior Ductile Tearing on Cleavage Fracture Toughness in the Transition Region*, **UILU-ENG-93-2014**, Department of Civil Engineering, University of Illinois, Urbana, Illinois, November 1993.

(3) PVRC Ad Hoc Group on Toughness Requirements (1972), *PVRC Recommendations on Toughness Requirements for Ferritic Materials*, **WRC Bulletin 175**, August 1972.

(4) NB-2330, Section III, Division 1-NB (1998), **ASME Boiler and Pressure Vessel Code**, 1998 Ed.

(5) Irwin, G. R. (1964), *Structural Aspects of Brittle Fracture*, **Applied Mat'ls Res, 3**, April 1964, p. 65.

(6) ASTM E208-91 (1993), *Standard Test Method for Conducting Drop-Weight Test to Determine Nil-Ductability Transition Temperatures of Ferrite Steels*, **1993 Annual Book of ASTM Standards, Volume 03.01**, American Society for Testing Materials, Philadelphia, PA, pp. 374-385.

(7) ASTM E23-93 (1993), Standard Test Methods for Notched Bar Impact Testing of Metallic Materials, ibid., pp. 206-226.

(8) Gerberich, W. W., S.-H. Chen, C.-S. Lee, and T. Levine (1987), *Brittle Fracture: Weakest Link or Process Zone Controlled?*, **Metallurgical Transactions A, 18A**, November 1987, pp. 1861-1875.

(9) Reuter, W. G., J. C. Newman, Jr., J. Skinner, and M. Mear (2000), Use of K_{Ic} and Constraint to Predict Location and Load Corresponding to Initiation of Crack Growth for Specimens Containing Surface Cracks, **32nd National Symposium on Fatigue and Fracture** (accepted for publication), West Conshohocken, PA, June 2000.

(10) Green, A. P. and B. B. Hundy (1956), *Initial Plastic Yielding in Notch Bend Tests*, **J. Mech. Phys. Solids, 4**, pp. 128-144.

CAVITY RING-DOWN SPECTROSCOPY AS A PLASMA DIAGNOSTIC: APPLICATIONS TO DIAMOND FILM GROWTH

U. Lommatzsch [1], E.H. Wahl [2], C.H. Kruger [2], R.N. Zare [1]

[1] *Department of Chemistry, Stanford University, Stanford, CA 94305, USA*
[2] *Department of Mechanical Engineering, Stanford University, Stanford, CA 94305, USA*

ABSTRACT

Cavity ring-down spectroscopy is a highly sensitive technique for absorption measurements and is used here for concentration and temperature measurements in a CVD reactor. Results are reported for the CH and the CH_3 radical in a hot-filament reactor for the synthesis of diamond thin films. The spatially resolved concentration measurements indicate different formation mechanisms for both radicals. The temperature measurements show a large temperature drop as the distance increases from the filament. The large gradients of concentration and temperature are typical for a system far from equilibrium and seem to be an important condition for achieving diamond growth under thermodynamically unfavorable conditions.

INTRODUCTION

The extreme hardness, the very high thermal conductivity and the high chemical resistance of diamond have led to large interest in diamond thin films for technological applications [1, 2]. The outstanding physical properties of diamond are the consequence of its structure and the strength of the bonding between the carbon atoms. Diamond thin films are used in such different areas as wear-resistant coatings for drilling tools, as thermal management devices in microelectronics, as sensors in harsh environments, and as optical materials. Synthetic diamond was first made by a high-pressure synthesis under conditions for which diamond is the thermodynamically stable modification of carbon. Diamond synthesis at low pressures and temperatures for which diamond is metastable with respect to graphite was first achieved in 1953. Based on this work several groups developed in the early 1980s the synthesis of thin film diamond by chemical vapor deposition (CVD). Diamond synthesis by CVD is accomplished in three steps: (1) activation of a mix of reactant gases to create reactive species in the gas phase;

(2) transport of those species to the substrate coupled with additional gas-phase reactions; and (3) deposition of carbon-containing species on the substrate surface and additional surface reactions resulting in diamond formation.

Activation is achieved thermally with a hot filament (HFCVD), by microwave plasmas, by RF plasmas or in flames. In a HFCVD system methane and hydrogen flow over a hot wire at a typical temperature of 2500 K. The activation includes the dissociation of the molecular hydrogen and the subsequent formation of various C_xH_y species. Diamond formation occurs on a substrate located about 1 cm from the filament. The main precursor(s) in the diamond growth is (are) still not identified, but CH_3 and C_2H_2 seem to be the most likely candidates [1, 2]. Despite much effort, widespread applications of diamond thin films are still limited owing to its high production cost. A rational attempt to improve film quality and growth rates requires the understanding of the reaction mechanisms and the identification of the precursor(s) in the diamond growth. An identification of the precursor would allow for a specific optimization of the growth conditions. Further scientific interest in diamond CVD synthesis is stimulated from the need to explain the growth of a material under metastable conditions.

In this study, cavity ring-down spectroscopy (CRDS) [3, 4] is used as a gas-phase diagnostic in a diamond HFCVD reactor. CRDS is a novel method for laser-based absorption measurements of species that either weakly absorb or exist at low concentration. The high sensitivity ($\alpha_{min} = 10^{-9}$ cm^{-1}), the simplicity of quantification without the need of a calibration procedure, the applicability to harsh environments and the spatial resolution make this technique especially useful for this purpose. In CRDS the intensity of a light pulse circulating in an optical cavity containing the absorber(s) is measured. In its simplest implementation a short light pulse from a pulsed laser is injected into a high-finesse linear resonator built from two highly reflective mirrors. Owing to mirror losses and sample absorption an exponential decay of the light intensity I with a characteristic time constant τ, the ring-down time constant, is observed.

$$I(t) = I_0 \exp(-t/\tau) \tag{1}$$

The ring-down time constant is inversely proportional to losses of the empty resonator including mirror losses and scattering and also to the absorption of gaseous species present within the resonator. By measuring the decay constants τ and τ_0 for a cavity with and without a sample present, respectively, the sample absorbance A is determined (see eq. 2, for $t_r = 2l/c$ where c is the speed of light and l is the distance between both mirrors).

$$A = \frac{t_r}{2}\left(\frac{1}{\tau} - \frac{1}{\tau_0}\right) \tag{2}$$

The number density N can then be extracted with the help of Beer's law (eq. 3) from the sample absorbance A when the absorption cross section σ and the path length l are known.

$$A = \sigma \cdot N \cdot l \tag{3}$$

The high sensitivity of CRDS results from the insensitivity to shot-to-shot fluctuations in laser intensity and the long absorption path length.

In this report it is demonstrated how CRDS can be used to determine the concentrations of short-lived radicals and their spatial distributions in a HFCVD reactor. From the line intensities in the absorption spectra of those radicals the gas temperature in the reactor is derived.

EXPERIMENTAL

Experimental details can be found in Refs. [5, 6]. For completeness a short description is given here. The CVD reactor is maintained at a pressure of 20 Torr and is filled with a mixture of 1% methane in hydrogen at a flow rate of 100 sccm. A resistively heated filament made out of tungsten with a length of 20 mm and a diameter of 0.2 mm is used. The molybdenum substrate is positioned at a filament distance of 7 mm. The optical cavity consists out of two mirrors with a maximum reflectivity R = 99.993 % at 430 nm that are separated by 65 cm. They are attached to the reactor by means of flexible bellows, allowing the CVD reactor to be moved independently with respect to the light path in the resonator. Thus spatially resolved measurement become possible. An optical parametric oscillator (OPO) pumped by a Nd:YAG laser with a repetition rate of 10 Hz is used as the light source. Before introducing the light into the optical cavity some optics are used to match the laser beam profile to the TEM_{00} mode of the cavity. The light exiting the cavity is detected by a photomultiplier (PMT) whose output is processed by a digital oscilloscope and a personal computer to extract the ring-down time constants.

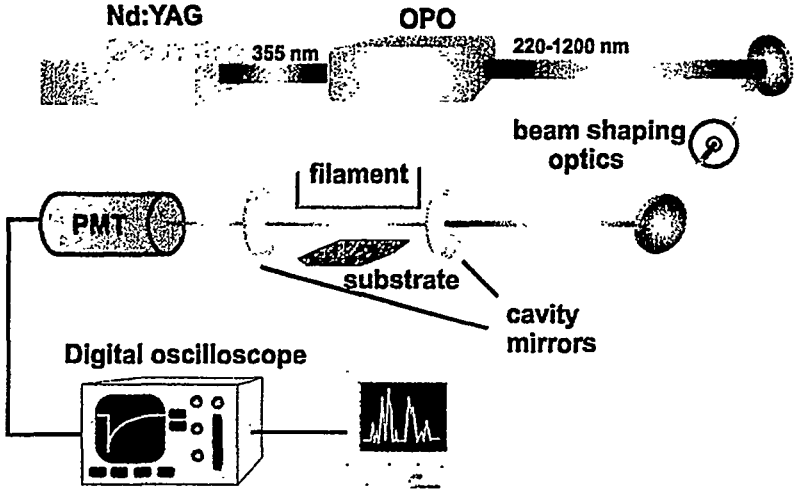

Figure 1. Schematic of the experimental setup. The CVD reactor with the filament and substrate inside can be translated to measure the ring-down signal at different distances from the filament.

RESULTS and DISCUSSION

CH Absorption Spectrum

The absorption spectrum of a molecule can be measured by CRDS when the change in ring-down time versus wavelength is recorded. As an example Fig. 2 shows the Q-branch members of the CH A-X band around 430 nm. Also shown in Fig. 2 is a simulated spectrum that is in excellent agreement with the experimental data. The Λ-type doubling and spin-orbit coupling can be clearly resolved owing to the small bandwidth of the OPO.

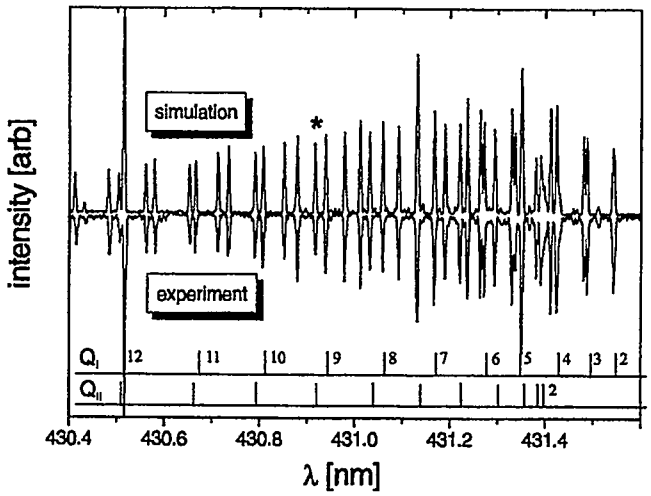

Figure 2. Experimental (top) and simulated (bottom) part of the absorption spectrum of CH A-X (0,0) band. The line marked with an asterisk is used to evaluate concentrations.

Spatial Distribution of Radical Species

Fig. 3 compares the concentration distributions of the CH_3 and the CH radical in our HFCVD reactor. The concentration profiles differ in two ways. First, the CH_3 concentration on the order of 10^{13} molecules/cm^3 is two orders of magnitude larger than the CH concentration. Secondly, both profiles differ also in their qualitative trend. While for CH_3 a concentration maximum occurs at a distance of 2 mm from the filament, the CH concentration declines steadily with increasing distance from the filament. Two conclusions can be drawn from those results related to diamond growth. The absolute concentration values for both radicals make it unlikely that CH is important in the diamond growth and confirm the view that CH_3 could be an important precursor in diamond growth. The differences in spatial distribution indicate different formation mechanisms for both radicals. While the maximum of the CH_3 concentration excludes a filament production mechanism, the CH profile resembles those of the H radical, a species that is formed heterogeneously at the filament [7].

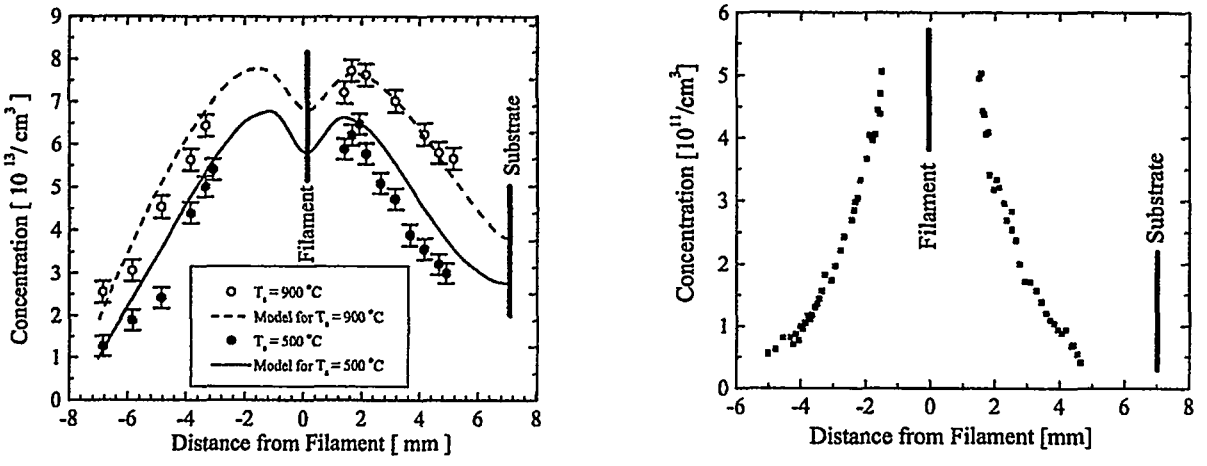

Figure 3. Concentration of CH_3 (left) and CH (right) at a filament temperature of 2500 K and a substrate temperature of ~ 1200 K.

Gas temperature in the CVD reactor

From the line intensities in the absorption spectrum the rotational temperature of a species can be derived from a Boltzmann plot. Fig. 4 displays the temperature of the CH radical at different locations in our reactor. Remarkable is the large temperature difference of ~ 1000 K between the filament and the gas phase. This behavior was already observed by Langmuir [8] in 1927 and can be related to the breakdown in the continuum energy conduction theory at low pressures. The comparison of the CH temperature profile with published profiles for the H atom [7] reveals a high correspondence. It can be thus concluded that most gas-phase species in the reactor have the same temperature, which is a behavior that is expected from the large number of collisions in the gas phase at a pressure of 20 Torr. Therefore the temperature profile in Fig. 4 can be used to represent the temperature in the reactor in theoretical simulations.

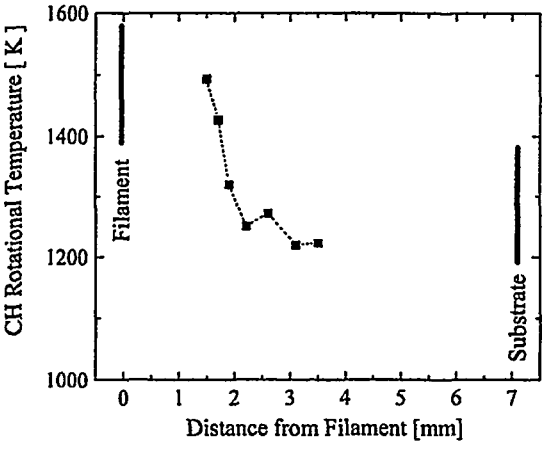

Figure 4. The temperature of the CH radical. Transitions with N=5-15 were included in the Boltzmann Plot. Filament temperature is 2500 K.

From the large gradients in concentration and temperature it follows that the whole CVD system is far from a thermodynamic equilibrium state. This non-equilibrium state is an important prerequisite for achieving diamond growth under thermodynamically unstable conditions.

ACKNOWLEDGMENTS

Uwe Lommatzsch gratefully acknowledges support from the Deutsche Forschungsgemeinschaft. This work was supported by the Engineering Research Program of the Office of Basic Energy Sciences at the Department of Energy.

REFERENCES

1. M. N. R. Ashfold, P. W. May, C. A. Rego, and N. M. Everitt, "Thin-Film Diamond By Chemical-Vapor-Deposition Methods", *Chem. Soc. Rev.* 23, 21 (1994).

2. D.G. Goodwin and J.E. Butler, in: M.A. Prelas, G. Popovici, L.K. Bigelow (Eds.), *Handbook of industrial diamonds and diamond films*, Marcel Dekker, New York, p. 527 (1998).

3. A. O'Keefe and D.A.G. Deacon, "Cavity ring-down optical spectrometer for absorption measurements using pulsed laser sources", *Rev. Sci. Instr.* 59, 2544 (1988).

4. K.W. Busch and M.A. Busch (Eds.), *Cavity-Ringdown Spectroscopy*, American Chemical Society, Washington (1999).

5. E.H. Wahl, T.G. Owano, C.H. Kruger, U. Lommatzsch, D. Aderhold and R.N. Zare in: J.C. Angus, W.D. Brown, J.P. Dismukes, M.D. Drory, A. Gicquel, A. Grill, R.H. Hauge, H. Kawarada, C.P. Klages, R.L. Opila, A. Paoletti, Y. Sato, K.E. Spear, and B.V. Spitsyn (Eds.) Diamond Materials VI, *Proc. Elelectrochem. Soc.* PV99-32 (1999), "Spatially resolved measurements of CH concentration and temperature in a hot filament CVD reactor" The Electrochemical Society, Pennington (in press).

6. U. Lommatzsch, E.H. Wahl, T.G. Owano, C.H. Kruger and R.N. Zare, "Spatial Concentration and Temperature Distribution of CH Radicals Formed in a Diamond Thin Film Hot-Filament Reactor", *Chem. Phys. Lett.* 320 (2000) 339.

7. S. A. Redman, C. Chung, K. N. Rosser, and M. N. R. Ashfold, "Resonance enhanced multiphoton ionisation probing of H atoms in a hot filament chemical vapour deposition reactor", *Phys. Chem. Chem. Phys.* 1 (1999) 1415.

8. I. Langmuir, "The dissociation of hydrogen into atoms", *J. Am. Chem. Soc.* 37, 417 (1915).

LIGHT SCATTERING MEASUREMENTS OF THERMAL DIFFUSIVITY FOR REFERENCE STANDARDS

Chris Muzny and Richard Perkins

Chemical Science and Technology Laboratory
National Institute of Standards and Technology
Boulder, Colorado 80303

ABSTRACT

Measurements of the thermal diffusivity of 1,1,1,2-tetrafluoroethane (R134a) are reported for the saturated liquid and vapor at temperatures from 293 to 374 K. These data were obtained by dynamic light scattering experiments on a very pure sample of R134a. This sample has been well characterized and was used previously for international round-robin measurements of the thermal conductivity and viscosity organized by the IUPAC Subcommittee on Transport Properties. The thermal diffusivity data can be used to calculate values for the thermal conductivity with values for the density and specific heat from the equation of state recommended by the International Energy Agency Annex 18 for R134a. Good agreement is found between thermal conductivity data obtained directly by transient hot-wire measurements on this same round-robin sample.

INTRODUCTION

Accurate knowledge of the thermophysical properties of fluids is required to optimally design gas transmission pipelines, thermomechanical systems such as power plants and refrigeration units, and chemical process plants. Internationally accepted values for the thermophysical properties of fluids are required to calibrate and use mass flow meters in pipelines as a basis for custody transfer agreements and to calibrate instruments used to measure thermophysical properties. The thermal diffusivity is a key thermophysical property that characterizes transient heat transfer through a medium. The thermal diffusivity a is defined as

$$a = \frac{\lambda}{\rho C_p}, \tag{1}$$

where λ is the thermal conductivity, ρ is the density, and C_p is the isobaric specific heat. The thermal diffusivity characterizes the ratio of energy transport to the volumetric energy storage of a medium. Eq. (1) shows that the thermal diffusivity is related to three thermophysical properties that can be independently measured. NIST has a unique capability to measure all of these properties with low uncertainty over a wide range of conditions (30 K to over 700 K, at pressures to 70 MPa) due to long-term support from the U.S. Department of Energy. Comparison of independent measurements of each of these properties through Eq. (1) provides a verification of the consistency of all of the measurements.

Both C_p and λ are divergent at the gas-liquid critical point. This divergence can be characterized by critical exponents, with C_p diverging about twice as quickly as λ so that the thermal diffusivity approaches zero at the critical point. Direct measurements of thermal conductivity require a temperature gradient that has the potential to drive convection of the fluid, making the measured heat flux higher than that due to pure conduction. Since the fluid is very compressible in the critical region, convection can be driven by extremely small temperature gradients making accurate thermal conductivity measurements nearly impossible. Additionally, a temperature gradient introduces heat transfer by thermal radiation. The thermal radiation contribution to the measured heat flux increases in proportion to the absolute temperature cubed. Most fluids are not transparent to infrared radiation but instead have complex absorption and emission spectra that are a function of both temperature and density, so that analysis of the contribution of thermal radiation during a thermal conductivity measurement becomes extremely difficult. Dynamic light scattering does not require a macroscopic temperature gradient to measure thermal diffusivity so it has significant advantages over other techniques used to measure thermal conductivity and thermal diffusivity in the critical region (fluid convection) and at elevated temperatures (thermal radiation).

The refrigerant 1,1,1,2-tetrafluoroethane (R134a) was selected for the present measurements of thermal diffusivity by dynamic light scattering. The gas-liquid critical point of R134a is located at a temperature of 374.21 K and a density of 511.9 $kg \cdot m^{-3}$. R134a is the most widely studied and the best characterized of the new alternative refrigerants. It is used as the reference fluid in corresponding states models that allow the prediction of the thermophysical properties of mixtures of alternative refrigerants. For this reason, it is necessary to know the properties of R134a with the lowest uncertainty possible. However, early measurements of the viscosity, thermal conductivity, and thermal diffusivity of alternative refrigerants, such as R134a, were found to have inconsistencies of as much as 30 to 40 % [1]. These large discrepancies were observed between laboratories that historically provided reliable data for nonpolar fluids. R134a is a relatively polar fluid with very good solvent properties. It easily picks up water and ionic solutes if not handled properly. Moderate electrical conductivity is observed in some samples that increases as a function of the water concentration. In 1992, the IUPAC Subcommittee on Transport Properties organized a round-robin study of these transport properties to try to resolve these discrepancies [1,2].

EXPERIMENTAL

The sample of 1,1,1,2-tetrafluoroethane (R134a) used in the present study was prepared for the IUPAC round robin on transport properties. It is from one of nine cylinders that were filled from a single high-purity sample, using procedures for pharmaceutical materials to insure the cleanliness of the cylinders before use. The purity of the sample was verified by gas chromatography for organic compounds and by the Karl-Fischer test for water. The purity of the batch was greater than 99.9 %, and the principal impurity was R134 at a concentration of 850 ppm. Water was present at a concentration of 6 ppm. NIST has made extensive measurements of the thermal conductivity of this round-robin sample

with two transient hot-wire instruments [1-3]. All the laboratories that participated in the IUPAC study agreed to analyze their data using the equation of state recommended by Annex 18 of the International Energy Agency for R134a [4,5]. This equation of state is also used to calculate the fluid density and specific heat as a function of the measured temperature and pressure in the present analysis.

The dynamic light scattering technique (DLS) is used to probe the microscopic entropy fluctuations of a transparent sample of pure fluid that remains in macroscopic equilibrium. For the thermal diffusivity, the broadening of the central Rayleigh line of the scattering spectrum is probed. At conditions away from the critical point, the scattering signal is a homodyne composite of the scattering from the sample containment windows and the fluid. As the fluid scattering becomes more intense near the critical point, the composite scattering signal is dominated by scattering from the fluid and may be treated as a self-beating correlation function. For the homodyne case, the intensity correlation function is of the form

$$G(\tau) = A + B \exp\left(\frac{-\tau}{\tau_c}\right), \tag{2}$$

where A, B, and τ_c are experimental parameters determined from a fit of the correlation function. The thermal diffusivity a is calculated using

$$a = \frac{1}{q^2 \tau_c}. \tag{3}$$

For scattering of coherent laser light of wavelength λ_0 at low scattering angles in the fluid θ_s,

$$q = \frac{4\pi n}{\lambda_0} \sin\left(\frac{\theta_s}{2}\right) \cong \frac{2\pi n \theta_s}{\lambda_0} = \frac{2\pi \theta_e}{\lambda_0}, \tag{4}$$

so the refractive index n of the fluid of interest is not required since it is the scattering angle θ_e, external to the vessel window, that is measured. The error associated with the low angle approximation is of the order of θ_e^3.

The sample of interest is contained in a cylindrical pressure vessel with highly polished quartz windows on each end. The quartz windows are sealed with polyimide o-rings that are compressed with threaded closures. The entire vessel is made from black anodized aluminum to minimized surface reflections and is designed for pressures to 35 MPa at 600 K. The fluid enters the cell from the bottom to minimize convection in the critical region at elevated temperatures. The fluid vessel is held tightly in an aluminum thermostat block with temperatures controlled with either circulating fluid or electrical heaters. The entire thermostat is well insulated to minimize temperature gradients and isolated from the optical bench with a temperature controlled shield plate. The temperature of the cell is measured with a reference platinum resistance thermometer (PRT) with an uncertainty of 5 mK. The calibration of the PRT is verified periodically with a water triple point cell. The pressure of the sample is measured with a quartz pressure transducer with a range from 0 to 21 MPa with an uncertainty of 0.01 % full scale (0.002 MPa). The calibration of the pressure transducer is verified periodically relative to a gas dead-weight gauge.

A schematic of the optical cell is shown in Fig. 1. The light source is a 1 W argon ion laser operating at a wavelength λ_0 of 514.5 nm. This laser is locked to a single longitudinal mode and the beam intensity is externally stabilized and focused prior to entering the scattering cell. The cell is maintained at

a fixed angle θ_e relative to the incident laser beam. This angle is determined accurately by the distance between the incident beam and its reflection from the vessel window at a known distance from the window. The scattered light passes through two apertures to minimize scattered light from the windows and detected by 2 cross-correlated photomultiplier detectors located normal to the cell windows. The use of two photomultiplier detectors eliminates errors due to afterpulsing. The signals from the photomultiplier detectors are collected and analyzed with a log-time correlator with a minimum time resolution 25 ns with up to 400 channels. From 4 to 20 replications with a duration of 5 minutes were obtained for each data point reported here.

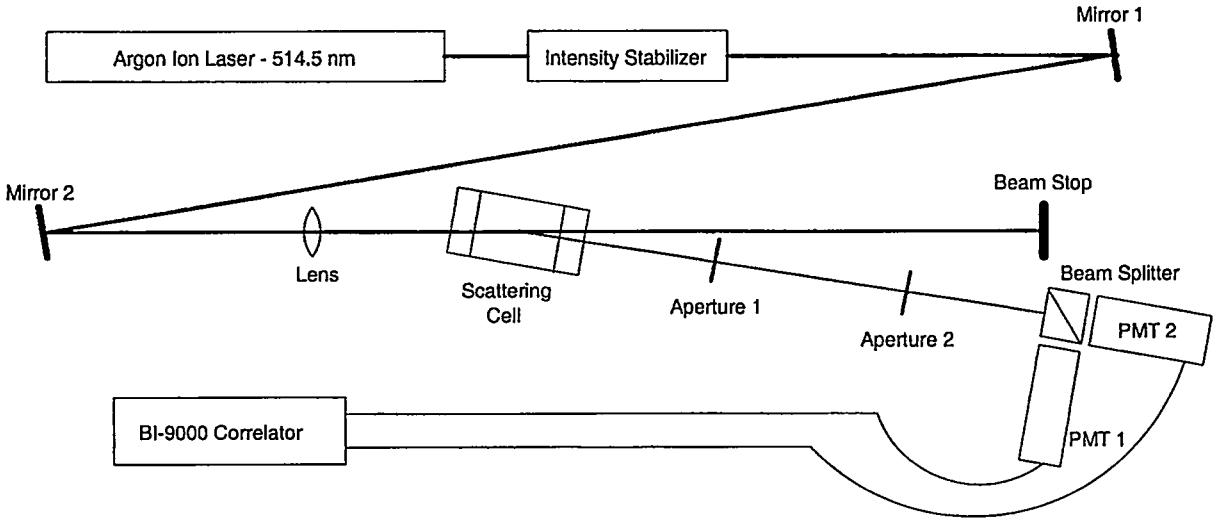

Figure 1. Schematic of the dynamic light scattering apparatus to measure the thermal diffusivity of fluids. The scattering cell operates at temperatures from 273 K to 500 K with pressures to 35 MPa.

The measurement vessel and its associated pressure system were evacuated and filled from the vapor phase of the round-robin sample of R134a. Three filling cycles were made prior to measurements to purge adsorbed compounds. The volume of the pressure system was varied with a hand piston pump to adjust the position of the liquid-vapor interface relative to the laser beam in the optical cell. The measured thermal diffusivity is provided in Table 1 and shown in Fig. 2 for saturated liquid and vapor at temperatures from 294 to 374 K. All temperatures are reported based on the ITS90 temperature scale.

DISCUSSION

The reference correlation of Krauss et al. [6] is selected as the basis for comparing the present measurements with other data obtained on this same round-robin sample of R134a [3,7-10]. This correlation is valid at temperatures from 240 to 410 K with densities to 1500 $kg \cdot m^{-3}$. The estimated uncertainty of this correlation is 5 %, increasing to 10 % in the near critical region. The critical point is located at a temperature of 374.274 K (ITS90) and a density of 515.25 $kg \cdot m^{-3}$. This correlation includes a crossover equation of state and is currently the best representation of the thermal diffusivity of R134a. Evidence of residual particle scattering (Stokes-Einstein diffusion) was observed in the liquid phase R134a at the lower temperatures where the scattering intensity was very low. Good agreement is found between the present measurements and the DLS measurements of Kraft and Leipertz [7] on the same round-robin sample of R134a.

Table 1

Thermal diffusivity of the round robin sample of R134a in the liquid (l) and vapor (v) phases.

T_l (K)	a_l (m^2s^{-1})	T_v (K)	a_v (m^2s^{-1})
294.478	4.72E-08		
297.927	4.94E-08		
306.750	4.69E-08		
316.558	4.60E-08		
326.589	4.41E-08		
336.279	4.34E-08		
346.130	2.97E-08	346.042	1.07E-07
355.894	2.86E-08		
365.726	1.59E-08	365.719	2.86E-08
372.712	5.14E-09	372.696	5.82E-09
373.174	3.88E-09	373.230	3.96E-09

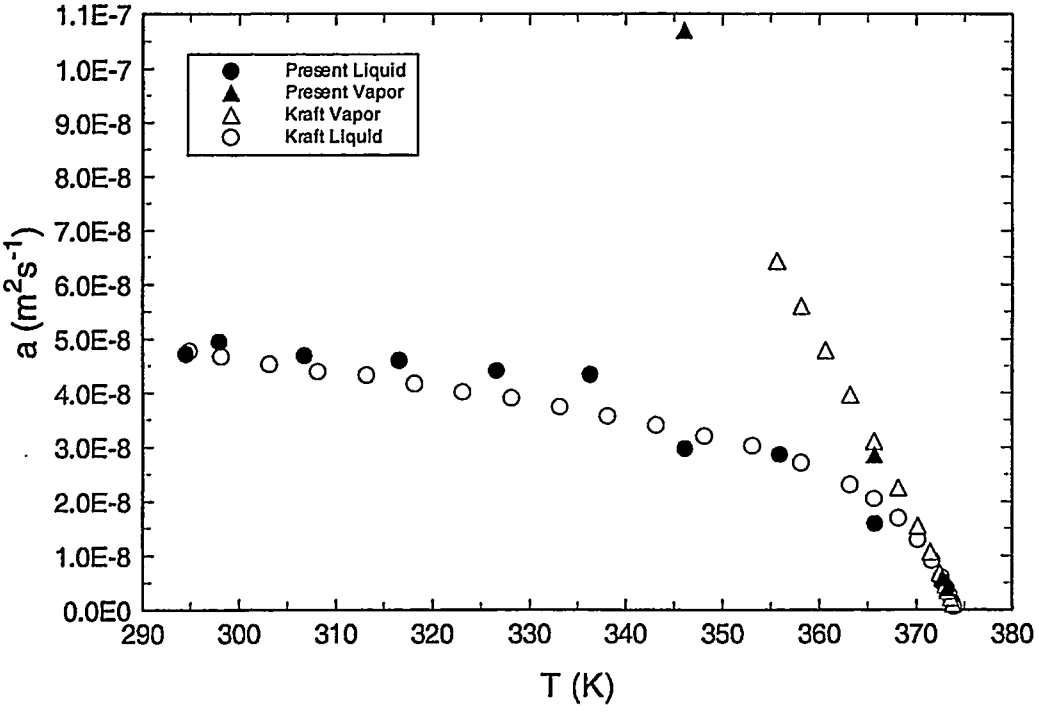

Figure 2. Thermal diffusivity of the round-robin sample of R134a near saturation as a function of temperature.

Saturated Vapor

The thermal diffusivity data were used to obtain estimates for thermal conductivity using Eq. (1) with density and specific heat values calculated using the Krauss et al. [6] correlation. These thermal conductivity estimates can then be compared with direct measurements of the thermal conductivity that were made on the vapor phase of the same sample of R134a. Deviations of these results from the Krauss et al. [6] correlation are shown in Fig. 3 for the saturated vapor phase. Good agreement is found between the present results and the results of the Kraft et al. [7] measurements of thermal diffusivity as well as the direct thermal conductivity measurements on the same sample. There are no systematic trends in the present data. The relatively large scattering in the present measurements is attributed to light scattering associated with the Stokes-Einstein diffusion of small particles that were observed in the sample. Good measurements were obtained at lower temperatures (lower intensity of scattered light) than those reported by Kraft et al. [7]. This demonstrates the high quality of the present optical setup.

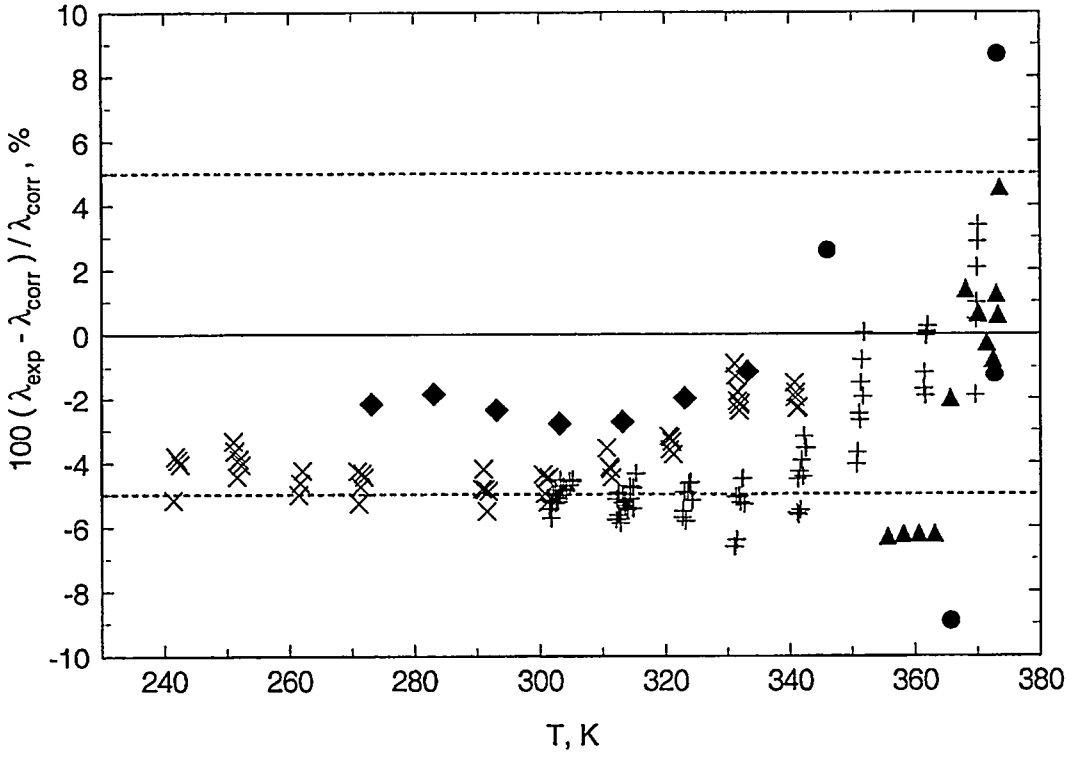

Figure 3. Deviations of the measurements of the thermal conductivity of vapor R134a along the saturation line from the correlation of Krauss et al. [6]. Round-robin sample: ◆, Assael et al. [8]; ×, Perkins [3] (Tr-Pt); +, Perkins [3] (Tr-Ta); ▲, Kraft and Leipertz [7]; ●, present results.

Saturated Liquid

Deviations of these results from the Krauss et al. [6] correlation are shown in Fig. 4 for the saturated liquid phase. Good agreement is found between the present results and the results of the Kraft et al. [7] measurements of thermal diffusivity as well as the direct thermal conductivity measurements on

the same sample. There are no systematic trends in the present data. The relatively large scatter in the present measurements is again attributed to the presence of particles in the sample and the transition from the homodyne to the self-beating scattering signal at temperatures from 320 to 340 K.

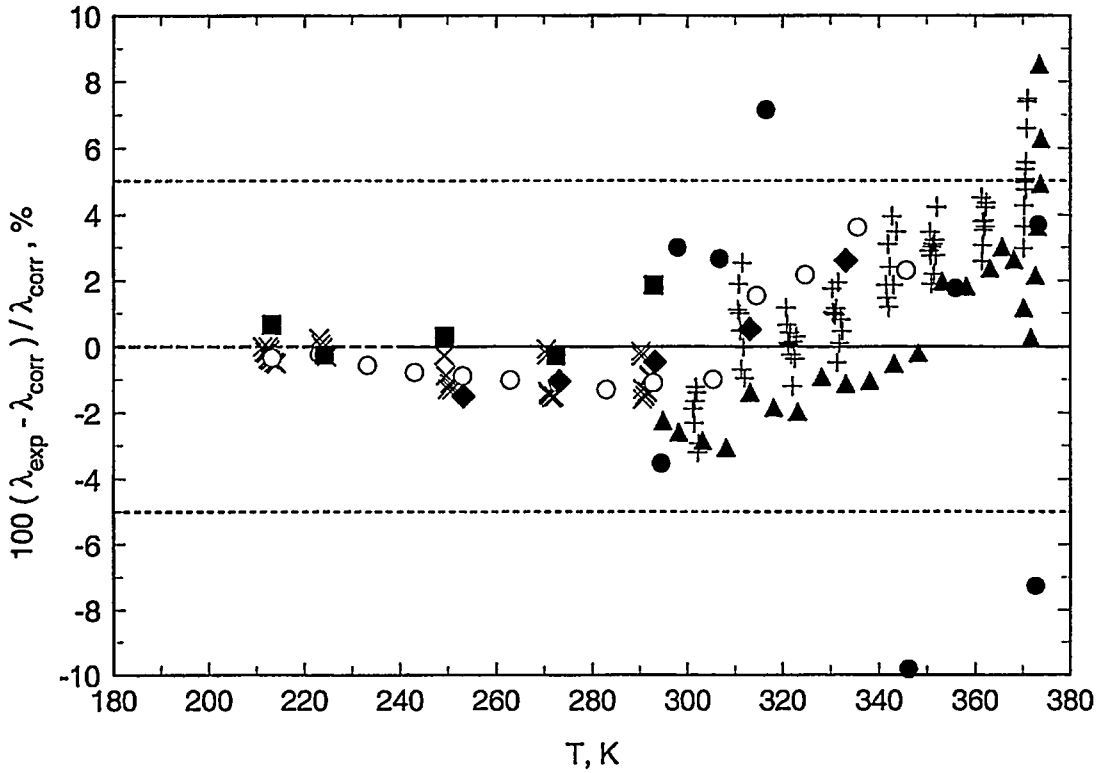

Figure 4. Deviations of the measurements of the thermal conductivity of liquid R134a along the saturation line from the correlation of Krauss et al. [2]. Round-robin sample: ◆, Assael et al. [8]; ×, Perkins [3] (Tr-Pt); +, Perkins [3] (Tr-Ta); ■, Gurova et al. [9]; ○, Nagasaka [10]; ▲, Kraft and Leipertz [7]; ●, present results.

FUTURE WORK

The uncertainty of future thermal diffusivity measurements will be reduced by careful cleaning of the optical cell and improved filtration of all samples. Measurements will be made at multiple scattering angles to eliminate uncertainties in the transition of the scattering signal from homodyne to self beating. The measurements will be validated over an extended range of temperature with both R134a and toluene. Thermal diffusivity measurements will be made on toluene over a wide range of temperatures (178 to 592.8 K) in conjunction with an international round robin on density calibration standards.

Since the dynamic light scattering technique does not require a temperature gradient, the thermal diffusivity measurements allow verification that independent measurements of thermal conductivity, density, and specific heat are consistent with each other at high temperatures where thermal radiation effects become very significant.

The dynamic light scattering technique will be extended to simultaneous measurements of the thermal diffusivity and mass diffusivity of key mixture systems that are common in the chemical and energy industries. Selected measurements will be made on well-defined mixtures to develop and validate predictive models for these mixture properties. These validated models will allow the design of highly optimized chemical and energy processes and the assessment of the environmental fate of chemical waste.

ACKNOWLEDGEMENT

The financial support of the United States Department of Energy, Division of Engineering and Geosciences, Office of Energy Sciences, is gratefully acknowledged.

REFERENCES

1. M.J. ASSAEL, Y. NAGASAKA, C.A. NIETO DE CASTRO, R.A. PERKINS, K. STRÖM, E. VOGEL, and W.A. WAKEHAM, "Transport Property Measurements on the IUPAC Sample of 1,1,1,2-Tetrafluoroethane (R134a)," *Int. J. Thermophys.* 21, 1 (2000).

2. M.J. ASSAEL, Y. NAGASAKA, C.A. NIETO DE CASTRO, R.A. PERKINS, K. STRÖM, E. VOGEL, and W.A. WAKEHAM, "Status of the Round Robin on the Transport Properties of R134a," *Int. J. Thermophys.* 16, 63 (1995).

3. R.A. PERKINS, J. HOWLEY, M.L.V. RAMIRES, A.N. GUROVA, and L. CUSCO, "Experimental Thermal Conductivity Values for the IUPAC Round-Robin Sample of 1,1,1,2-Tetrafluoroethane (R134a)," National Institute of Standards and Technology, NISTIR, 2000, in press.

4. R. TILLNER-ROTH and H.D. BAEHR, "An International Standard Formulation for the Thermodynamic Properties of 1,1,1,2-Tetrafluoroethane (HFC-134a) for Temperatures from 170 K to 455 K and Pressures up to 70 MPa," *J. Phys Chem Ref. Data* 23, 657 (1994).

5. S.G. PENONCELLO, R.T. JACOBSEN, K.M. DE REUCK, A.E. ELHASSAN, R.C. WILLIAMS, and E.W. LEMMON, "The Selection of International Standards for the Thermodynamic Properties of HFC-134a and HCFC-123," *Int. J. Thermophys.* 16, 781 (1995).

6. R. KRAUSS, J. LUETTMER-STRATHMANN, J.V. SENGERS, and K. STEPHAN, "Transport Properties of R134a," *Int. J. Thermophys.* 14, 951 (1993).

7. K. KRAFT and A. LEIPERTZ, "Thermal Conductivity and Sound Velocity of Round-Robin R134a," *Fluid Phase Equil.* 125, 245 (1996).

8. M.J. ASSAEL, N.A. MALAMATARIS, and L. KARAGIANNIDIS, "Measurements of the Thermal Conductivity of Refrigerants in the Vapor Phase," *Int. J. Thermophys.* 18, 341 (1997).

9. A.N. GUROVA, U.V. MARDOLCAR, and C.A. NIETO DE CASTRO, "The Thermal Conductivity of Liquid 1,1,1,2-Tetrafluoroethane (HFC 134a)," *Int. J. Thermophys.* 18, 1077 (1997).

10. Y. NAGASAKA, Personal Communication, Keio University, Yokohama, Japan.

COHERENCE AND SYNCHRONIZATION IN ARRAYS OF CLASS B LASERS

Y. Braiman, G. Bitton, H. K. Liu, V. Protopopescu, L. Zhang, and J. Barhen

Center for Engineering Science Advanced Research
Computer Science and Mathematics Division
Oak Ridge National Laboratory
Oak Ridge, TN 37831-6355

ABSTRACT

We report preliminary results on the feasibility of a compact source of high power and high intensity coherent radiation emitted from an array of semiconductor diode lasers via the injection of a controlled electromagnetic field into the cavity of each laser. The new source is expected to generate coherent radiation of the order of magnitude of 10 Watts and to spawn a new laser technology of compact high power devices.

INTRODUCTION

In recent years there has been a dramatic increase in demands for high power, high intensity diffraction limited beams. High power compact coherent sources are extremely useful in many engineering applications. For this reason, phase-locked arrays of diode lasers have been studied extensively over the last 15 years [1,2]. Such devices have been built to achieve high coherent radiation for applications such as space communication, blue-light generation via frequency doubling, optical interconnects, parallel optical signal processing, high-speed, high-resolution laser printing, and end-pumping solid-state lasers. Conventional, narrow-stripe ($< 3 - 4$ μm wide), single-mode lasers provide, at most, 100 mW reliably [1], since they are limited by the optical power density at the laser facet. For reliable operation at watt ranges, broad-area laser (BAL) arrays with large-aperture (≥ 100 μm in width) are necessary. However, such BAL arrays usually exhibit multi-lateral-spatial modes and the output beams are mutually incoherent. Thus, the challenge has been to obtain single-mode operation from large-aperture devices, and maintain stable, coherent behavior to high power levels.

Laser arrays provide an intriguing class of nonlinear dynamical systems with many degrees of freedom. Of particular interest is the emergence of mutual synchronized behavior where all the

elements execute in-phase oscillations. This phenomena is very important in a variety of engineering, physical, and biological systems, yet our current theoretical understanding of the subject is far from being complete. Stability of the in-phase dynamics in laser arrays was recently theoretically studied for both solid state [3-6] and semiconductor [7-9] lasers. It is well documented [10] that the most common behavior in laser arrays is, indeed, the anti-phase behavior where the phases of adjacent lasers differ by π. Therefore, an external forcing is required to induce stable in-phase dynamics.

To overcome the antiphasing tendency and to maintain stable and coherent operation of the array, one possible technique is to inject a controlled electromagnetic field into the cavity of each laser. This field will synchronize the array and control chaos if it arises. Some of the challenges associated with the successful implementation of this idea are:

- Achieve effective *uniform* injection into each laser with a moderate power single-mode laser.
- Phase lock the array (though lasers are almost identical, the desired in-phase state is unstable for a broad range of parameters) and maintain the coherence.

Injection locking has been successfully used to obtain single mode emission in high power diode lasers or laser arrays [11]. The general method is to inject an external beam from a master single frequency laser into the cavity of the slave laser (see Fig. 1). The incident angle can be adjusted to stimulate a specific mode, which gives high coherent output power. An alternative approach is to feedback part of the output beam through grating, etalon, or phase conjugate mirror.

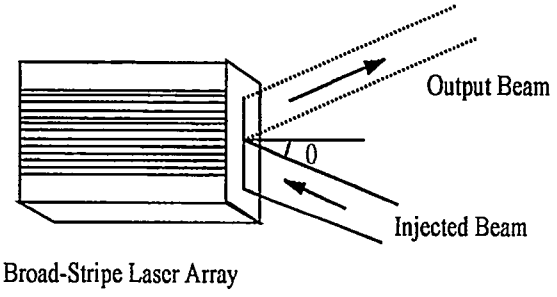

Figure 1 Schematic diagram of ordinary injection locking.

A free-running broad area diode-laser array usually generates output with a broad spectrum (~ 3 nm). Its far-field intensity distribution exhibits large divergence. When injection-locking is introduced, the spectral distribution will follow the injection laser. As a result, both the spatial and spectral power density are improved [11,12]. The injection-locking efficiency relies on the injected power, which may range from 0.1% to 10% of the output power depending on the injection structure. For closely spaced and internally coupled laser arrays, uniform or even single stripe injection can achieve a satisfactory effect at low current. However, for broad stripe and widely separated high power laser arrays, the lateral mode structure is more complicated. High power operation also brought robust conditions for maintaining coherent operations. Maximum injection effect can be achieved via individually controlled mode matching. Individually controlled phase modulation gives us another freedom to fine-tune the injections. To our knowledge, these techniques have not been used on broad stripe laser arrays.

To achieve phase locking, couplings among diode laser emitters are required. To phase lock broad stripe laser arrays, external coupling might provide the only viable approach. Global couplings [13,14] can provide automatic phase locking via gain control. However the stability region in global

coupling is small compared with nearest neighbor (series) couplings. With larger drive current, the system can be driven into chaos easily. With nearest neighbor coupling more external controls can be applied, thus helping us to collect information on phase locking high brightness diode laser arrays.

Here we present preliminary results on the theoretical analysis and experimental implementation of the idea discussed above.

PHASE MODEL ANALYSIS FOR TWO-LASER ARRAY

Solid state [3-6] and semiconductor [7-9] lasers are considered class B lasers and are described by similar dynamical equations. For solid state lasers, under certain dynamical conditions [3,15], the main features of the full dynamics are adequately captured within a simplified description, called the phase model. Here we present the analysis of the entrainment of coupled solid state lasers over a large range of injection fields [15] (work on semiconductor laser arrays is in progress). In particular, we elucidate a newly observed dynamical behavior of the total output intensity, namely, strongly nonmonotonic growth as the function of the injection strength [4].

Our starting point is the system of equations describing the dynamics of two evanescently coupled lasers, where the polarization is adiabatically eliminated [3-6]:

$$\dot{E}_j = (G_j - \alpha_j + i\delta_j)E_j + \kappa(E_{j+1} + E_{j-1}) + E_e(t)$$

$$\dot{G}_j = \frac{\tau_e}{\tau_f}[p_j - (1+|E_j|^2)G_j], \qquad j = 1,2 \tag{1.1}$$

The variables E_j and G_j are the dimensionless complex electric field and gain for the *jth* laser. All times and frequencies are scaled relative to the cavity round trip time, τ_c, and τ_f is the fluorescence time of the laser medium; α_j and p_j are the dimensionless cavity decay and pump rates for the *jth* laser respectively, κ is the evanescent coupling constant between the two lasers, and $E_e(t)$ is the slowly varying amplitude of the external field which drives each laser. Eqs. (1.1) are written in a frame rotating with frequency ω_e, at which the external field has a non-zero Fourier component. This frequency is tuned to minimize the detuning from the cavity resonances. In practice, the output power emitted from an array depends on the tuning of external field to the cavities [16]. The detuning $\delta_j \approx \omega_e - \omega_{cj}$, where ω_{cj} is the cavity resonance frequency for laser *j*. For solid state lasers, the latter dynamic contribution to the detuning is generally ignored. A variation in detuning amongst the lasers would result from a variation in cavity lengths for the laser elements. However, we have in mind a single cavity containing the array.

In the following, we allow for a small spread in detunings as a way to test the robustness of the entrainment mechanism to a physically reasonable parameter spread. We assume $\alpha_j = \alpha$, $p_j = p$, $p > \alpha$. Substituting $E_j(t) = \sqrt{I_j(t)}\exp(i\phi_j(t))$, where $I_j(t)$ and $\phi_j(t)$ are the intensity and the phase of laser *j* and assuming $E_e(t) = \sqrt{I_e}$ to be a constant field, the model equations for two lasers reads:

$$\dot{I}_j = 2(G_j - \alpha)I_j + 2\kappa\sqrt{I_1 I_2}\cos(\phi_2 - \phi_1) + 2\sqrt{I_e I_j}\cos\phi_j$$

$$\dot{\phi}_j = \delta_j + (-1)^j \kappa \frac{\sqrt{I_1 I_2}}{I_j}\sin(\phi_1 - \phi_2) - \sqrt{\frac{I_e}{I_j}}\sin\phi_j \qquad (1.2)$$

$$\dot{G}_j = \frac{\tau_c}{\tau_f}(p - G_j - G_j I_j)$$

Eqs. (1.2) have been studied theoretically for N coupled lasers [4] and the condition for full entrainment has been derived. This condition assumes small deviations in detunings and small coupling. We denote the dimensionless amplitude of the injected field by $A_e = \sqrt{I_e/I}$ where $I = p/\alpha - 1$. Ideally, to entrain an array of N identical lasers requires injected field amplitude $A_{entr} = 4\kappa$, or $A_{entr} = 4\kappa\sqrt{I}$. The functional form of the total output intensity may significantly depend on the parameters of the array (such as detunings and the coupling constant).

In Fig. 2 we show the normalized total intensity $I_{tot} = (|E_1|^2 + |E_2|^2)/4I$ of two coupled lasers as a function of A_e. The injected field frequency approximately corresponds to the average of frequencies of each laser, thus it is tuned to minimize the detunings from the cavity resonances.

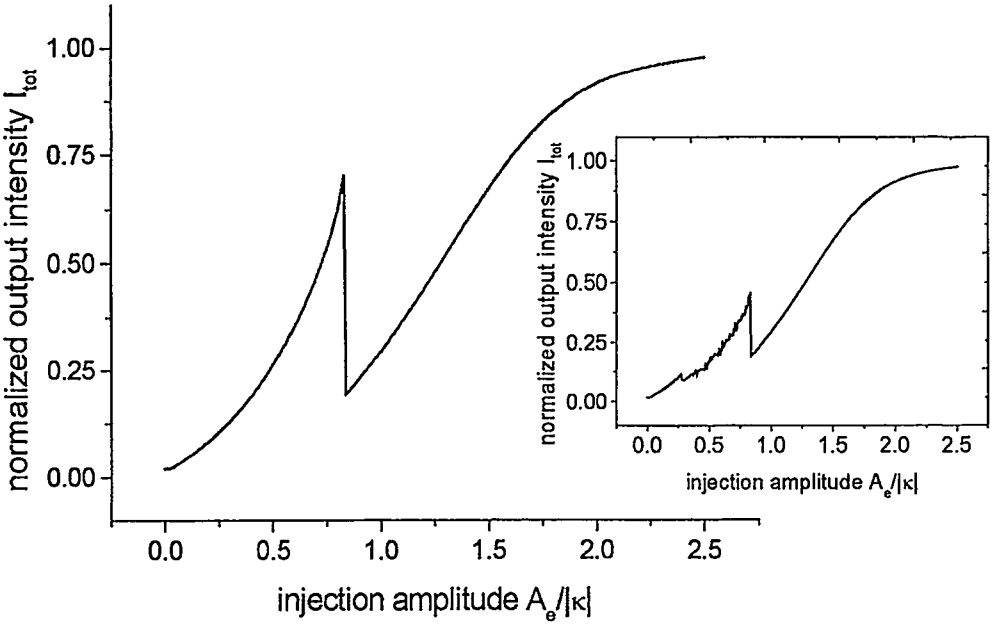

Figure 2. The normalized total intensity, $I_{tot} = (|E_1|^2 + |E_2|^2)/4I$, as a function of the strength of the dimensionless injected field, $A_e = \sqrt{I_e/I}$. An inset shows the averaged normalized total intensity, I_{tot}, as a function of A_e.

We continuously change the strength of the injected field to mimic an experiment where the injected field is gradually increased. Initially, the total intensity grows with the injected field. When the injection strength reaches the critical amplitude, A_c, the total intensity drops discontinuously to a significantly lower level. We notice that, just below A_c, the total output intensity of the array is about 70% of the maximum intensity (at full entrainment), but requires only about 20% of the entrainment injected field, E_{entr}. We estimate that if we apply a different set of initial conditions, the probability to obtain qualitatively very similar behavior, as demonstrated in the Fig. 2, is in the vicinity of 60%. Our estimation is based on simulating a sample of 500 realizations of distinct initial conditions. In the inset of Fig. 2 we present the *averaged* value of the normalized total intensity of two coupled lasers versus the dimensionless amplitude of the external field $A_e = \sqrt{I_e / I}$. The curve is obtained by numerically solving Eqs. (1.2) for two coupled lasers and averaging over 500 realizations of the initial conditions.

A characteristic feature of independent solid state lasers (i.e. without coupling and external field) is that, for any initial data their intensities and gains relax to a stationary state $(I, G) = (p/\alpha - 1, \alpha)$, i.e. the amplitudes $|I_j - I|$ and $|G_j - G|$ decay to zero. Numerical experiments [4], using physically realistic parameter values, show similar transient behavior of intensities and gains in the full laser array system, where both coupling and excitation terms are present. Once these transients have decayed, it turns out that the dynamics of the phases no longer depend on intensities. This motivates, at least at a heuristic level, the use of the phase equations in Eqs. (1.2), with $I_j = I$, as an approximation model to the full system. It turns out that the phase equations retain the essential features of the dynamics and can be used to explain the nonmonotonic behavior displayed by the solution of the complete system (1.2) [15].

To better understand the dynamics of laser array, phase models are widely used. Phase model is a very powerful tool to study dynamics of coupled lasers [3] and laser arrays [4]. It is used mostly in analysis of solid state lasers (rather than in semiconductor lasers) in situation where the fluctuations of the intensities and gains of each laser are small. The phase equations in (1.2) provide a significantly reduced description which captures nevertheless the essential dynamics, including the sudden drop in output intensity depicted in Fig. 2, as we now show. The frequency of the external field is tuned to minimize the detunings from the cavity resonance, thus we may assume $(\delta_1 + \delta_2) \approx 0$. This assumption allows us to reduce the dimensionality of the parameter space and carry out the (simplified) analysis of the dynamics and of the fixed points in the plane $(\delta_1 - \delta_2, \kappa)$. On this plane, we use experimentally suggested parameters for weakly coupled solid state Nd:YAG lasers [17], therefore we look in the rectangle $\kappa \in (-0.5, -3)$ and $\delta \equiv \delta_1 - \delta_2 \in (0, 1)$.

With these assumptions, the stationary form of the phase equations in Eq. (1.2) reads:

$$\begin{aligned} &\sin\phi_1 + \sin\phi_2 = 0 \\ &\delta_1 - \delta_2 + 2\kappa(\sin(\phi_2 - \phi_1)) - A_e(\sin\phi_2 - \sin\phi_1) = 0. \end{aligned} \tag{1.3}$$

The first equation in (1.3) implies that either a): $\phi_2 - \phi_1 = (2m+1)\pi$ or b): $\phi_1 + \phi_2 = 2\pi m$, where m is an integer. Solutions of class (a) imply $\sin(\phi_2 - \phi_1) = 0$, yielding $\sin\phi_1 = \delta_1 / A_e$, $\sin\phi_2 = \delta_2 / A_e$ and $\sin(\phi_1 - \phi_2) = \sin(\sin^{-1}(\delta_1 / A_e) - \sin^{-1}(\delta_2 / A_e)) \neq 0$, i.e. inconsistency. Hence, the

only possibility is the class (b) of solutions which, in turn, can be divided in two sub-classes: m even and m odd. For m even, the second equation in (1.4) becomes:

$$f(x) \equiv -\delta - 2\kappa \sin x - 2A_e \sin\frac{x}{2} = 0 \tag{1.4}$$

where we substituted $\delta = \delta_1 - \delta_2$, and $x = \phi_2 - \phi_1$.

For small values of A_e, this equation has two solutions, one stable and one unstable. By increasing the strength of the injected field A_e, a saddle-node bifurcation occurs at a critical value, A_c. For $A_e > A_c$, Eq. (1.4) has no real solution. To determine A_c, we solve the system $f(x) = -\delta - 2\kappa \sin x_c - A_c \sin x_c/2 = 0$, and $f'(x) = -2\kappa \cos x_c - A_c \cos x_c/2 = 0$. Making the substitution $\tan x_c/2 = z$ and eliminating A_c we obtain a cubic equation for z that admits the explicit solution:

$$z = (-\frac{q}{2} + \sqrt{D})^{1/3} + (-\frac{q}{2} - \sqrt{D})^{1/3} - \frac{\delta}{12\kappa} \tag{1.5}$$

where $D = (p/3)^3 + (q/2)^2$, $p = -\delta^2/48\kappa^2$, and $q = \delta/4\kappa + \delta^3/864\kappa^3$. Then up to higher order terms, $z_c = (\delta/-12\kappa)^{1/3} + \delta/12\kappa$ and the critical amplitude where the jump in intensity occurs is [15]:

$$A_c = -2\kappa \frac{1 - z^2}{\sqrt{1 + z^2}}. \tag{1.6}$$

For m odd, the second equation in (1.3) reads:

$$g(x) \equiv -\delta - 2\kappa \sin x + 2A_e \sin\frac{x}{2} = 0. \tag{1.7}$$

A similar analysis shows that this equation has two solutions, one unstable close to $x \simeq 0$ and one stable, close to $x \simeq \pi$. Thus, at small values of the amplitude of the injected field A_e, the system has two stable solutions, one close to $\pi/2$ that solves the Eq. (1.4) and one close to π that solves Eq. (1.7). Since the total output intensity is given by $I_{tot} = 4\cos^2(x/2)$, one solution has high intensity, while the other one has low intensity. Each of these stable solutions has a basin of attraction and the selection of the solution depends, of course, on the initial conditions. When $A_e = A_c$, the high-intensity solution disappears at the saddle-node point and for higher values of A_e only the low intensity solution remains. We obtained an excellent fit between the numerical and analytical expressions. The numerical simulations and theoretical analysis for larger arrays is in progress.

EXPERIMENTAL SETUP

The objective of the experiment is to investigate the feasibility of extracting high coherent power from a semiconductor laser array using optical injection of self and/or external optical fields. The experimental setup whose conceptual scheme is shown in Figure 3 is of master-slave type. For the

master laser we have chosen a single mode tunable diode laser. For the slave laser we have chosen a high power (30W) broad area diode laser bar composed of 19 different non-coupled emitters. The key point in this setup is the ability to access light rays associated with each emitter separately. This setup enables us to measure the intensity and relative phase of each emitter separately. This information will then be used to decide how to modify the system parameters so a high power coherent radiation can be achieved. The information from the emitters' radiation profile will be used as control measure for modifying the amplitudes and relative phases of the injected optical field to each emitter separately and also to change the magnitude of the coupling between the different emitters.

The detailed experimental setup is as follows: The master laser used for injection is a single mode tunable diode laser (DL100). The laser line-width is 1 MHz and maximum output power of up to 120 mW. To avoid optical feedback into the master laser cavity, a Double Faraday Isolator (DLI from LINOS Photonics) with more then 60 dB isolation is used. Part of the beam is then sampled by a beam splitter (BS1) to be used as reference beam for the coherence measurements and the rest of the beam passes through a computer-generated hologram (CGH) optics (Rochester Photonics Corp). Illuminating the CGH with one beam produces an array of 21 equal spots. The size of those spots can be manipulated by the size of the spot in the entrance to the CGH. The distance between the CGH and the collimating optics (collimator 2) determines the distance between the spots. The collimated beams can be phase or amplitude modulated by a spatial light modulator (SLM1). The individually modulated injection beams are then passed through two beam splitters (BS2 and BS3) and a micro-lens array, and enter the cavity of the diode laser bar as a seeding beams. The Diode laser bar (B1-81-20c-19-30-A from Coherent) is composed of 19 emitters with 500μm spacing between centers and 1 × 150 μm^2 emitting surface. The output from the laser array is collimated by a micro-lens array with astigmatism compensation. The micro-lens array is coated with a low reflection coating, which will provide nearest neighbors coupling between the different emitters in the diode laser array. The coupling strength can be adjusted by changing the reflectivity of the coating and the distance between the laser array and the micro-lens array. The output beam from the diode laser bar are then past trough another SLM (SLM2) which can be used to select certain beams from the total output of the diode laser bar for measurements. The measurement system consists of infrared CCD cameras for near-field and far-field pattern observation; Mach-Zender interferometer for coherence measurement; scanning Fabry-Perot for mode structure measurement; and a monochromator for wavelength measurement.

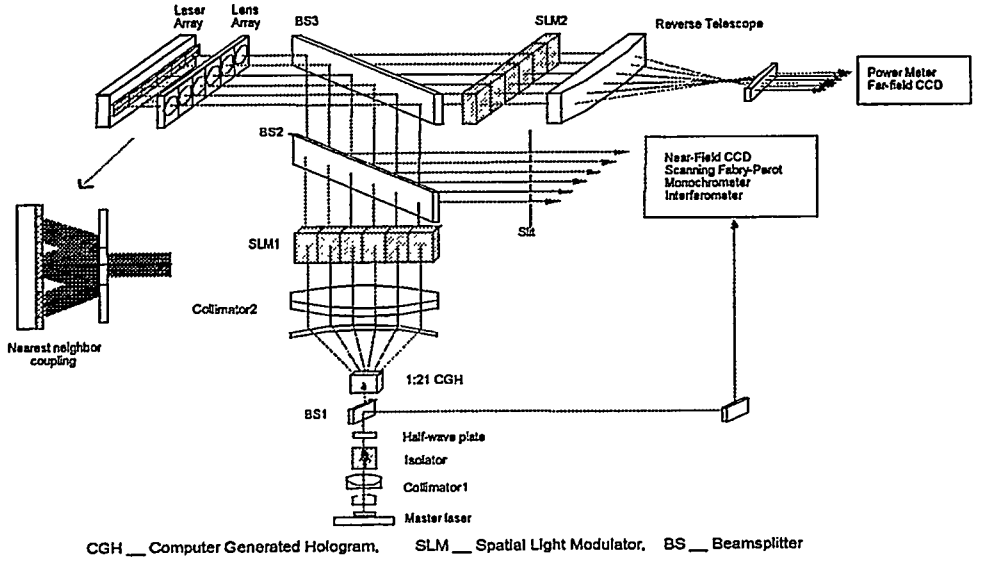

Figure 3 Schematic diagram of the experimental arrangement

The experimental setup suggested in Figure 3 was essentially realized (Figure 4). The measuring instruments were installed including the software and the hardware needed for the computer control over the instruments and the experiment. The main part of the measuring equipment include, CCD cameras for characterizing the far field and near field patterns of the emission from the diode laser bar, and a monochromator (From CVI) with resolution of 0.07nm for spectral characterization of the diode laser bar output. For higher spectral resolution a Fabry-Perot spectrum analyzer with 2GHz free spectral range and 100 MHz resolution is used.

External optical injection into all the emitter cavities by the master laser was achieved. The optical characteristics of the free running diode laser bar were measured. Without injection of external optical field or injection of optical feedback, the emissions from the different emitters are mutually incoherent and the beam divergence is large (beam divergence of about 10×35 degrees) as expected. Preliminary results of the influence optical feedback using partial reflective mirrors suggest that an in-phase coherent radiation mode can be induced in the diode laser bar.

One of the most important parameters that may affect the injection locking is the thermal stability. When the environment temperature changes, the resonant wavelength of the laser changes. When this change is larger than the longitudinal mode spacing $\Delta\lambda_{FP}$, mode hopping occurs and injection locking fails. The cooling system and the materials used for thermal contact were chosen as to avoid such a failure. The diode laser bar produces about 80 watt of heat power, while our cooling system (from "Neslab) is able to remove 500 watt of heat power. The cooling system seems to be efficient enough even when the diode laser bar is operated at full power.

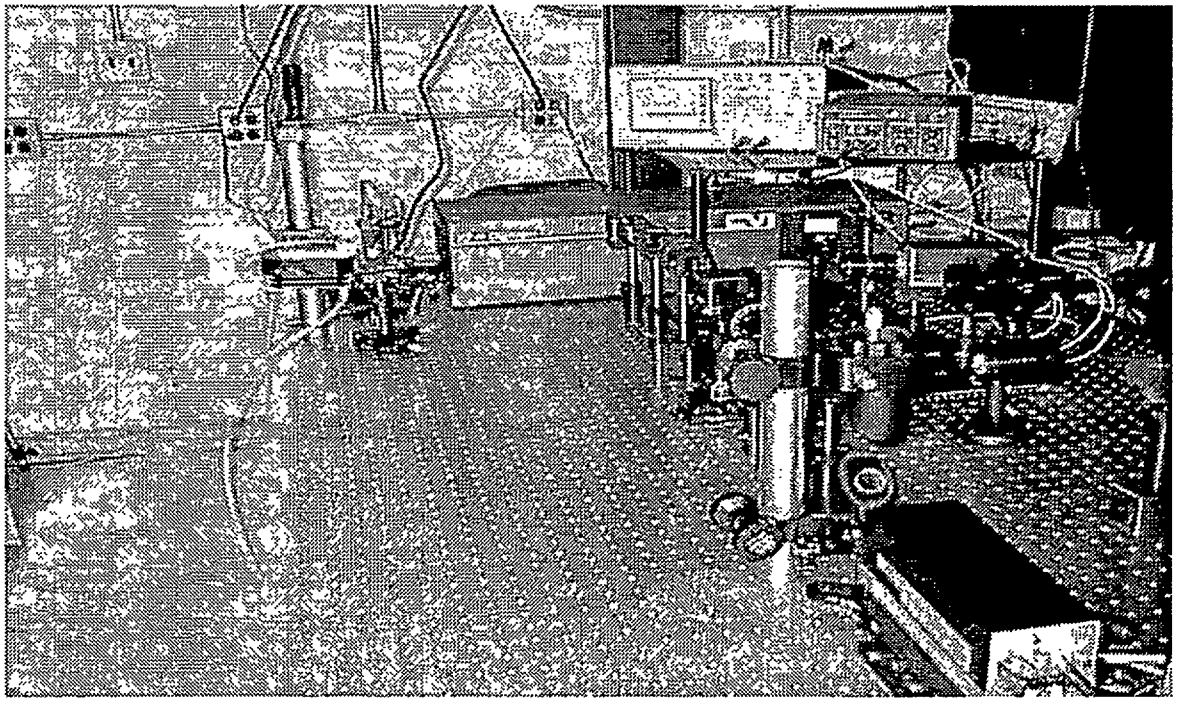

Figure 4: Actual Experimental Setup

SUMMARY

We have described preliminary work on the feasibility of a compact source of high power and high intensity coherent radiation emitted from the array of semiconductor lasers. Theoretically, we have studied how the total output intensity of two coupled lasers depends on the injection strength and elucidated mechanism of nonmonotonic behavior. In particular, we notice that, just below A_c, the total output intensity of the array is about 70% of the maximum intensity (at full entrainment), but requires only about 20% of the entrainment injected field, E_{entr}. This result is of great experimental importance, since obtaining high power injection in to laser arrays is not a simple task. In addition, by reducing injection strength, the input power for device operation is reduced, and we will explore this phenomena for larger laser arrays. Experimentally, we have essentially implemented the proposed design and we are in the process of fine-tuning the apparatus.

In this paper we discussed in detail the in-phase synchronization of laser arrays. Besides the in-phase synchronization, other modes of synchronized behavior are of great interest for variety of application, as well as for better understanding the dynamics of laser arrays. We are planning to pursue studies on periodic and chaotic [18] synchronization between lasers in arrays, as well as between two distinct arrays.

ACKNOWLEDGMENT

This research was partially sponsored by the Engineering Research Program of the Office of Basic Energy Sciences, U.S. DOE, by the Laboratory Directed Research and Development Program of Oak Ridge National Laboratory (ORNL), managed by UT-Battelle, LLC for the U. S. DOE under Contract No. DE-AC05-00OR22725, and by the Office of Naval Research.

REFERENCES

1. D. BOTEZ and D. R. SCIFRES, *Diode Laser Arrays*, Cambridge University Press, (1994).
2. N. W. CARLSON, *Monolitic Diode-Laser Arrays*, Springer-Verlag (1994).
3. A. I. KHIBNIK, Y. BRAIMAN, T. A. B. KENNEDY, K. WIESENFELD, *Phase Model Analysis of Two Lasers with Injected Field*, Physica D **111**, 295 (1998).
4. Y. BRAIMAN, T. A. B. KENNEDY, K. WIESENFELD, and A. I. KHIBNIK, *Entrainment of Solid-State Laser Arrays*, Phys. Rev. A **52**, 1500, (1995).
5. M. SILBER, L. FABINY, and K. WIESENFELD, *Stability Results for In-Phase and Splay-States of Solid State Laser Arrays*, J. Opt. Soc. Am. B **10**, 1121 (1993).
6. H. ADACHIHARA, O. HESS, R. INDIK and J. V. MOLONEY, *Semiconductor-Laser Array Dynamics-Numerical Simulations on Multistripe Index-Guided Lasers,* J. Opt. Soc. Am. B **10**, 496 (1993).
7. S. S. WANG and H. G. WINFUL, *Dynamics of Phase-Locked Semiconductor-Laser Arrays*, Appl. Phys. Lett. **52**, 1774 (1988).
8. R.-D. LI and T. ERNEUX, *Preferential Instability in Arrays of Coupled Lasers* Phys. Rev. A **46**, 4252 (1992).
9. J. MERCIER and M. McCALL, *Stability and Dynamics of an Injection-Locked Semiconductor Laser Array,* Optics Communications **138**, 200 (1997).

10. T. FISHMAN and A. HARDY, *Injection Locking Analysis of Vertical-Cavity Laser Arrays,* J. Opt. Soc. Am. B **16**, 38 (1999).
11. L. GOLDBERG, H. F. TAYLOR, J. F. WELLER, and D. R. SCIFRES, *Injection Locking of Coupled-Strip Diode Laser Arrays,* Applied Physics Letters **46**, 236 (1985).
12. B. BEIER, J.-P. MEYN, R. KNAPPE, K.-J. BOLLER, G. HUBER, and R. WALLENSTEIN, *A 180 mW Nd:$LaSc_3(BO_3)_4$ Single-Frequency TEM_{00} Microchip Laser Pumped by an Injection-Locked Diode-Laser Array,* Appl. Phys. B **58**, 381 (1994).
13. J. R. LEGER, G.J. SWANSON, and W.B. VELDKAMP, *Coherent Beam Addition of GaAlAs Lasers by Binary Phase Gratings*, Appl. Phys. Lett. **48**, 1240 (1986).
14. J. R. LEGER, M.L. SCOTT, and W.B. VELDKAMP, *Coherent Addition of AlGaAs Lasers Using Microlenses and Diffractive Coupling*, Phys. Lett. **52**, 1771 (1988).
15. A. KHIBNIK, Y. BRAIMAN, V. PROTOPOPESCU, T. A. B. KENNEDY, and K. WIESENFELD, *Amplitude Dropout in Coupled Lasers*, submitted to Phys. Rev. A (May, 2000).
16. M. K. CHUN, L. GOLDBERG, and J. F. WELLER, *Injection-Beam Parameter Optimization of an Injection-Locked Diode-Laser Array,* Opt. Lett. **14**, 272 (1989).
17. L. FABINY, P. COLET, R. ROY, and D. LENSTRA, *Coherence and Phase Dynamics of Spatially Coupled Solid State Lasers,* Phys. Rev. A **47**, 4287 (1993).
18. R. ROY and K. S. THORNBURG, *Experimental Synchronization of Chaotic Lasers*, Phys. Rev. Lett. **72**, 2009 (1994).

Experimental Tests of Radiative Transfer Incorporating Statistical Optics Using Blackbody Sources

Yu Sun, Roland Winston, and Joseph J. O'Gallagher
Enrico Fermi Institute, University of Chicago, Chicago, Illinoise 60637
Keith A. Snail
Naval Research Laboratory, 4555 Overlook Avenue, SW, Washington, DC 20375-5320

Abstract

In this symposium we present two experiments using blackbody sources that show departures from radiative transfer theory based on geometrical optics but are in accord with the predictions of statistical optics with the inclusion of measurement.

In the first experiment, an infrared camera is used to scan an edge illuminated by a blackbody source. Diffraction effects resulting in negative values and oscillations are predicted using Walther's second definition of generalized radiance. With the inclusion of measurement a prediction that is independent of the choice of definition used, does not go negative, and has no oscillations is found to agree well with experiment.

In the second experiment, an infrared camera is used to scan distant square blackbody sources. The experimental conditions are such that we have partial coherence at the entrance plane of the camera. Important parameters are identified showing when large departures from geometrical optics will take place. It will be seen that under these conditions, the radiance level is significantly lower than predicted by geometrical optics.

1 Introduction

Radiative transfer is an important subject with applications in astronomy and astrophysics, particle beam physics, medical physics, and machine vision. Yet it is still grounded in geometrical optics. Central to the subject is the quantity radiance and the equation of transfer. The equation of transfer is an analogue of the Boltzmann equation in statistical mechanics with radiance serving as the distribution. The distribution however is a power distribution in phase space rather than a probability distribution. As a distribution radiance can be used to calculate the energy density, the energy flux density, and the radiation pressure as a function of position.

Attempts have been made in constructing the foundation of radiative transfer by defining a generalized radiance based on the two-point correlation function of the wavefield. Major contributions along this line of research have been made by A. Walther [1] [2] and Wolf and his school (see [3] for further references). The equation of transfer, Boltzmann equation for light, would then result from Helmholtz equation and the definition for generalized radiance. This program, however, is mired in questions about the choices of definitions.

In a soon to be published paper, we propose a formalism of radiative transfer including statistical optics that is independent of the phase space representation used. We use the Wigner representation in this paper for the reason that we want a representation for which transformations of distributions based upon geometrical optics in phase-space (classical Hamiltonian mechanics) would serve as a good approximation. The most important ingredient of the theory is measurement. The output signal of an instrument is posited to be described by the expression [4]:

$$Q = \int \frac{d^2\mathbf{r}_\perp d^2\mathbf{k}_\perp}{(2\pi)^2} W(\mathbf{r}_\perp, \mathbf{k}_\perp; \omega, z) M(\mathbf{r}_\perp, \mathbf{k}_\perp) \, . \tag{1}$$

Here, $W(\mathbf{r}_\perp, \mathbf{k}_\perp; \omega, z)$ is the distibution in ray phase-space of the wavefield at the entrance plane of the instrument, and $M(\mathbf{r}_\perp, \mathbf{k}_\perp)$ is the distribution for the instrument. The conducted experiments presented in this paper test how well this expression agrees with the actual registered signal.

2 Theoretical expression

The scans, done by internal mechanisms of the cameras, are angle scans. The signal curve is then posited to be given by

$$Q(\phi) = \int \frac{d^2\mathbf{r}_\perp d^2\mathbf{k}_\perp}{(2\pi)^2} W(\mathbf{r}_\perp, \mathbf{k}_\perp; \omega, z) M(\mathbf{r}_\perp, \mathbf{k}_\perp; \phi) \ , \tag{2}$$

where ϕ is the angle position of the instrument.

In arriving at the theoretical expression for the signal curve, we first need an expression for the distribution of the wavefield at the entrance plane of the camera. The detailed calculation of the blackbody Wigner function at the souce plane for both an edge and a square source can be done analytically using Kirchhoff's approximation and the sinc correlation for a blackbody. The result for an edge is:

$$W(\mathbf{r}_\perp, \mathbf{k}_\perp; \omega, 0) = \begin{cases} \dfrac{B_\nu(T)}{\pi\sqrt{1-b^2}} [\mathrm{Si}(2k_0 x(1-b)) + \mathrm{Si}(2k_0 x(1+b))] & : \ x \geq 0 \\ 0 & : \ x < 0 \end{cases} \tag{3}$$

where $b = |\mathbf{k}_\perp / k_0|$, $k_0 = 2\pi/\lambda$, T is the source temperature, and Si is the sine integral. The function B is given by:

$$B_\nu(T) = \frac{2h\nu^3}{c^2} \frac{1}{e^{h\nu/kT} - 1} \tag{4}$$

The instantaneous field of view (ifov), the field of view per pixel (see figure (1)), of the cameras used are on the order of $1mrad$ (we will use the symble θ_0 for the ifov in the rest of this paper). We can therefore use the approximation,

$$b \simeq 0 \ , \tag{5}$$

in equation (3),

$$W(\mathbf{r}_\perp, \mathbf{k}_\perp; \omega, 0) = \begin{cases} \dfrac{2B_\nu(T)}{\pi} \mathrm{Si}(2k_0 x) & : \ x \geq 0 \\ 0 & : \ x < 0 \end{cases} \tag{6}$$

As $\lambda \longrightarrow 0$ this reduces to the classical expression.

The fact that θ_0 is small also allows us to use the paraxial approximation in evolving the blackbody distribution. For a distance z from the source plane, the Wigner function, valid up to the Fresnel diffraction regime, correponds to a shearing in phase space [5].

$$W(\mathbf{r}_\perp, \mathbf{k}_\perp; \omega, z) = W(\mathbf{r}_\perp - \frac{\mathbf{k}_\perp}{k_0} z, \mathbf{k}_\perp; \omega, 0) \ . \tag{7}$$

For the instrument, we use a simple model: a detector placed at the focal point of a lens. The relevent dimensions are given in the figure (1). The Wigner distribution for this instrument in 2-d phase space is derived in [4] and is given by

$$I(x,k) = \frac{1}{\pi}\mathrm{Si}\left[\frac{N\pi}{2}(1+\frac{k}{\kappa})(1-\left|\frac{x}{a}\right|)\right] + \frac{1}{\pi}\mathrm{Si}\left[\frac{N\pi}{2}(1-\frac{k}{\kappa})(1-\left|\frac{x}{a}\right|)\right] \ . \tag{8}$$

In this expression,

$$\kappa = \frac{\theta_0 k_0}{2} \ , \tag{9}$$

and N is,

$$N = \frac{d\theta_0 k_0}{2\pi} = \frac{d\theta_0}{\lambda} \ ; \quad d = 2a. \tag{10}$$

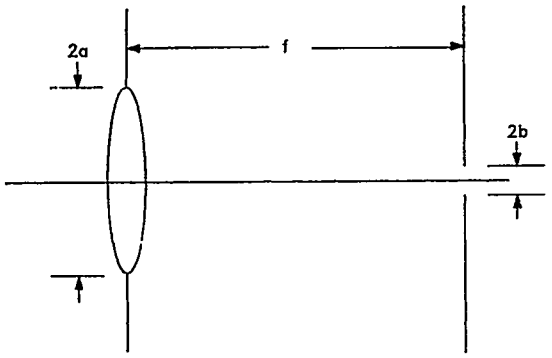

Figure 1: Model for camera. The focal length of the lens is f, and the pixel size is $2b$. $\theta_0 = 2b/f$

Of great importance is the value of N, which is an estimate of the number of phase-space cells occupied by the instrument (in 2-d phase space a cell has an area of 2π), or as a rough estimate of the number of spatial modes the instrument accepts. From equation (10), we see that if the detector of the camera lies with its edges at the first zeros of the Airy diffraction pattern then $N = 2.44$.

For large N, $N >> 1$, $I(x,k)$ goes over into the classical expression for the instrument. To see this, use the large x limit of $\mathrm{Si}(x)$,

$$\lim_{x\to\pm\infty} \mathrm{Si}(x) = \frac{\pi}{2}\,\mathrm{sgn}(x) = \pi\left[\Theta(x) - \frac{1}{2}\right] \tag{11}$$

We then have for $N >> 1$

$$I(x,k) = I_c(x,k)\ , \tag{12}$$

where

$$\begin{aligned} I_c(x,k) &= \Theta\left[\left(1+\frac{k}{\kappa}\right)\left(1-\left|\frac{x}{a}\right|\right)\right] + \Theta\left[\left(1-\frac{k}{\kappa}\right)\left(1-\left|\frac{x}{a}\right|\right)\right] - 1 \\ &= \begin{cases} 1 & : \ |x| \le a,\ |k| \le \kappa \\ 0 & : \ \text{otherwise} \end{cases} \end{aligned} \tag{13}$$

is the classical window function. It selects the phase space cells over which one wants to take the average of a quantity.

We will use the expression,

$$M(\mathbf{x}_\perp, \mathbf{k}_\perp) = I(x,k_x)I(y,k_y)\ , \tag{14}$$

as the 4-d phase space Wigner distribution of the instrument. Physically, we are modelling our instrument as having a square aperature instead of a circular aperature.

To arrive at the signal curve we need $M(\mathbf{r}_\perp, \mathbf{k}_\perp; \phi)$. We will model the angle scan as displacements of the instrument distribution in the direction transverse to the optic axis.

$$M(\mathbf{r}_\perp, \mathbf{k}_\perp; \phi) = M(x - z\,\phi, y, \mathbf{k}_\perp) \tag{15}$$

In performing the integration, the rapid oscillations of the blackbody distribution can be integrated out under the right circumstances by the much slower oscillations of the instrument distribution. This occurs when

$$\frac{2k_0 a}{N} \gg 1\ , \tag{16}$$

and

$$\frac{2z\kappa}{N} \gg 1\ . \tag{17}$$

Under the conditions of our experiments, $N \simeq 1$ and $\lambda \ll a,\ z\theta$. Therefore, when performing the integration, we can basically approximate the blackbody Wigner function as a classical distribution. For the edge, the distribution at the source plane is

$$W(\mathbf{r}_\perp, \mathbf{k}_\perp;\, \omega, 0) = B_\nu(T)\, \Theta(x) \tag{18}$$

For a square source of diminsions D X D, the distribution at the source plane is

$$W(\mathbf{r}_\perp, \mathbf{k}_\perp;\, \omega, 0) = B_\nu(T)\ [\Theta(x + D/2) - \Theta(x - D/2)][\Theta(y + D/2) - \Theta(y - D/2)] \tag{19}$$

In this sense, how non-classical a distibution is depends on the other distribution which it is integrating against. By making this approximation, we reduce the 4-d phase-space integration, equation (2), to 2-d phase-space integrations.

For the case of scanning an edge we have the expression:

$$Q(\phi) = B_\nu(T) N \int \frac{dxdk}{2\pi} I(x - z\,\phi, k)\, \Theta(x - \frac{k}{k_0} z)\ . \tag{20}$$

For the case of scanning a distant square blackbody source with diminsions D X D, we have the expression:

$$Q(\phi) = B_\nu(T) Q_y \int \frac{dxdk}{2\pi} I(x - z\,\phi, k)\ \left[\Theta(x + D/2 - \frac{k}{k_0} z) - \Theta(x - D/2 - \frac{k}{k_0} z)\right]\ , \tag{21}$$

where

$$Q_y = \int \frac{dydk}{2\pi} I(y, k)\ \left[\Theta(y + D/2 - \frac{k}{k_0} z) - \Theta(y - D/2 - \frac{k}{k_0} z)\right]\ . \tag{22}$$

In the conducted experiments, we aim to see departures of the measured signal from the predictions of geometrical optics but are in agreement with our formalism. The differences in the two predictions can be traced to the fact that $I(x, k)$ occupies a larger region in phase space than $I_c(x, k)$. This is a consequence of accepting k values larger than the restriction placed by geometrical optics. This characteristic becomes more pronounce as N becomes smaller.

In the edge scan experiment, $N << 1$. For the scans of distant sources, $N \simeq 1$ and $N \simeq 2$. Most instruments have $N \gtrsim 1$. Using $N << 1$ may not be the recommended thing to do in practice, but our intentions are to verify the theory.

3 Edge scan

Different aluminum aperture plates were placed in front of the camera lens to control the value of N. The camera performs an angle scan. The output signal of the camera is sent to a computer which performs a ten frame averaging.

In comparing experiment with theory, the following steps were taken to circumvent the problem of obtaining absolute power levels. The plateau for the theory curve as well as the curve for geometrical optics was raised to coincide with the plateau of the data. The background level was taken as the zero level. The theoretical curve plotted against the experimental data points is shown in figure (2) for a particular case.

It should be noted that the use of the classical blackbody distribution, i.e. no negative values or oscillations, in equations (18) and (20) is a result of using the Wigner basis. If one had used the Kirkwood basis the real part of the blackbody distribution for small k values would be [6], neglecting the $B_\nu(T)$ factor,

$$W(\mathbf{x}_\perp, \mathbf{k}_\perp; \omega, z) = [\mathrm{C}(u) + \mathrm{S}(u) + 1]/2 \tag{23}$$

with u given by

$$u = \sqrt{k_0/\pi z}\,(x - \left|\frac{\mathbf{k}_\perp}{k_0}\right| z)\ . \tag{24}$$

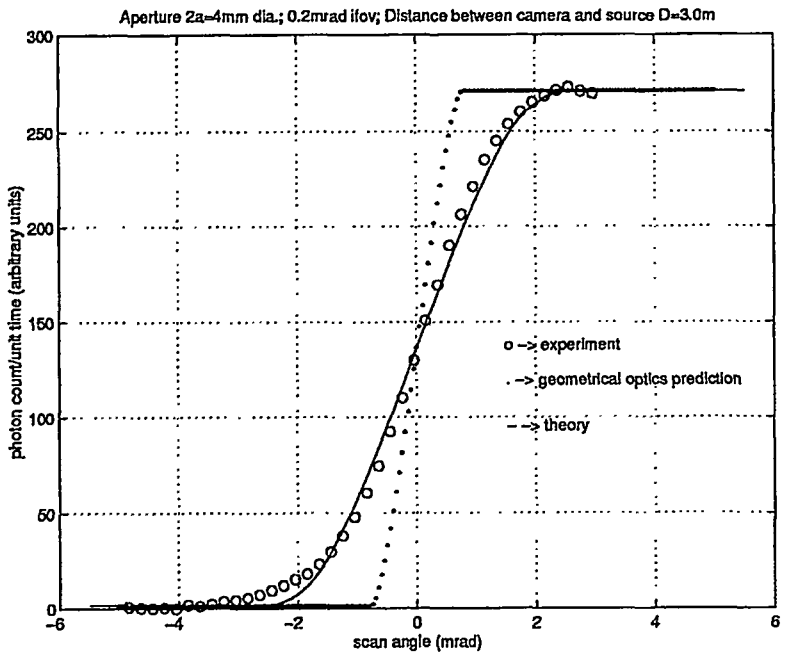

Figure 2: Edge scan. Open circles are data points. The dotted curve is the geometrical optics prediction, and the solid curve is the theoretical prediction.

$C(u)$ and $S(u)$ are Fresnel integrals. This expression is also valid up to the Fresnel diffraction regime. The plot is shown in figure (3). Because of the $\sqrt{k_0/\pi z}$ factor, this distribution does not correspond to a shearing in phase space upon propagation. From the figure one sees that oscillations and negative values are prominent. The distribution becomes less classical as z increases [7]. If we had chosen to calculate Q using the Kirkwood basis for large z values we would not have been able to use the classical distribution for the wavefield in the integration. However, since our formalism is representation independent, the result would have been the same as using the Wigner representation, although we would have to use the Kirkwood representation for I. The role of I in the Kirkwood representation is thus to smooth out the oscillations and to give positive and real values for the predicted power.

4 Partial coherence for distant sources

In this experiment, an infrared camera is used to scan square blackbody sources whose diminsions and distances from the camera are chosen for the appearance of partial coherence instead of having simple plane wave diffraction at the entrance of the camera.

The condition to be satisfied is for the transverse coherence length of the radiation field, l_s, to be smaller than the diameter of the camera lens, d,

$$l_s = \frac{\lambda}{\theta_s} \lesssim d\,, \tag{25}$$

where θ_s is the angle subtended by the source at the entrance of the camera. From above, a useful parameter characterizing the source is

$$N_s = \frac{d\theta_s}{\lambda} \gtrsim 1\,. \tag{26}$$

Assuming that at the entrance plane of the camera, the camera aperture is completely filled by the wavefield. N_s then provides an estimate of the number of cells in phase space, the number of spatial modes of the wavefield, reaching the camera. Using the definition for N, equation (10), we can express

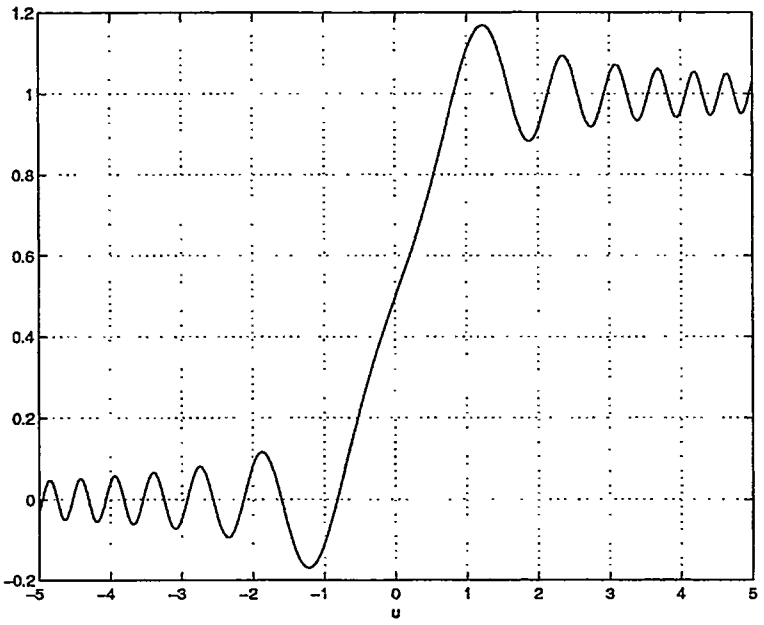

Figure 3: Plot of $(C(u)+S(u)+1)/2$. $C(u)$ and $S(u)$ are Fresnel cos and sin integrals. $u = \sqrt{\frac{k_0}{\pi z}}\,(x - \left|\frac{k_\perp}{k_0}\right| z)$

N_s as

$$N_s = N\,\frac{\theta_s}{\theta_0}\ . \tag{27}$$

The source size, D, is related to θ_s by

$$\theta_s = \frac{D}{z}\ . \tag{28}$$

A convenient scale to measure the source size is the "footprint", sometimes known as the beam size of the instrument. The footprint is defined as

$$\mathrm{FP} = d + z\theta_0\ . \tag{29}$$

In using the footprint, we are in a sense propagating the instrument function in the negative z-direction instead of propagating the correlation in the positive z-direction. Let

$$F = \frac{D}{FP}\ . \tag{30}$$

We use F and N_s as the parameters replacing z and D. They are related by

$$z = \frac{FN\lambda}{\theta_0^2\left(\frac{N_s}{N} - F\right)}\ , \tag{31}$$

$$D = \frac{FN_s\lambda}{\theta_0\left(\frac{N_s}{N} - F\right)}\ . \tag{32}$$

The three parameters, N, N_s, and F, allow us to scale the problem. The scaling is a particular type of cannonical transformation. N and N_s are invariant under a cannonical transformation where the x dimension is stretched (shrinked) and the k dimension is shrinked (stretched). F and N_s then set the D and z values.

Knowing that $I(x,k)$ accepts k values larger than the restriction placed by geometrical optics, the cross section of the beam at large distances will be larger than FP. The experiment consist of looking for the drop in radiance level as z is increased, or N_s is decreased, for fixed F and N. This of course requires us to vary D when N_s is varied. The prediction of geometrical optics is that the radiance level should be the same for $F \geq 1$. For large distances, the power level stops decreasing with increasing distance. From equation (31), this corresponds to $N_s \simeq FN$.

The experiments were conducted using the following camera characteristics: the aperature diameter after being stopped down is 3/8 in. The field of view per pixel is 1.5mrad. $\lambda = 9.3\mu m$ and $\Delta\lambda = \pm 0.75\mu m$. As our model of the instrument assumes a square camera aperature, we have to decide whether to use an inscribed square or a circumscribed square of a circle for the camera aperature in our model. After some trial and error, the theoretical curves for the inscribed square aperature agree better with experiment. For our camera characteristics $N = 1.54$ for the circumscribed square and $N = 1.09$ for the inscribed square.

The source size is controlled by placing aluminum masks of thickness $1/16^{"}$ and with square aperature of dimensions D X D over the aperature of the blackbody sources set at $T = 600°C$ for $F \simeq 1$ and $T = 100°C$ for $F \gtrsim 2$.

The maximum power level reaching the detector as predicted using geometrical optics is unchanged as z is changed, while the two theories agree at short distances. The power level reaching the detector as a fraction of the maximum power level predicted using geometrical optics can be measured by first measuring the maximum power for a close distance, $z \simeq 1m$, and using it as the maximum level for geometrical optics at all distances. We will call the close distance measurement our calibration measurement.

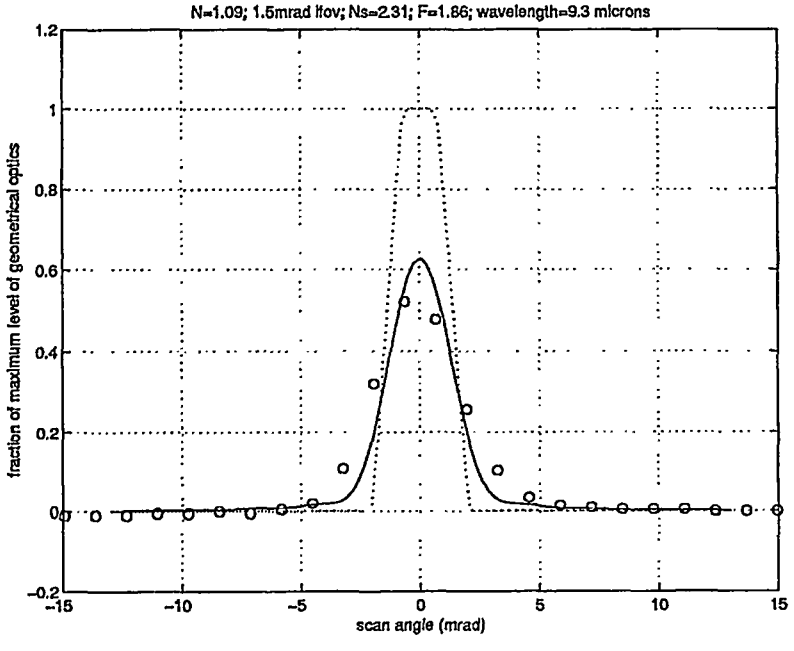

Figure 4: Radiance Drop. Open circles are data points. The dashed curve is the geometrical optics prediction, and the solid curve is the theoretical prediction.

The contribution of the background to the power reaching the detector is not negligible. In comparing experiment with theory, the power reaching the detector is measured from that due to the background, i.e. treating it as the zero level. This is also done for the calibration measurement. The difference between the maximum measured value and the background in the calibration measurement is taken as the maximum level for geometrical optics. The ratio between the measured value minus the background level and the maximum level for geometrical optics is compared with theory. This protocol in processing the data for comparison with theory allows us to sidestep the problem of obtaining absolute power and

temperature values. In taking these steps, we have essentially gotten rid of the $B_\nu(T)$ factor in front of equation (21).

A plot of theory versus experiment is shown in figure (4). The global behavior for the height of the radiance level when $N = 1.43$ is shown in figure (5).

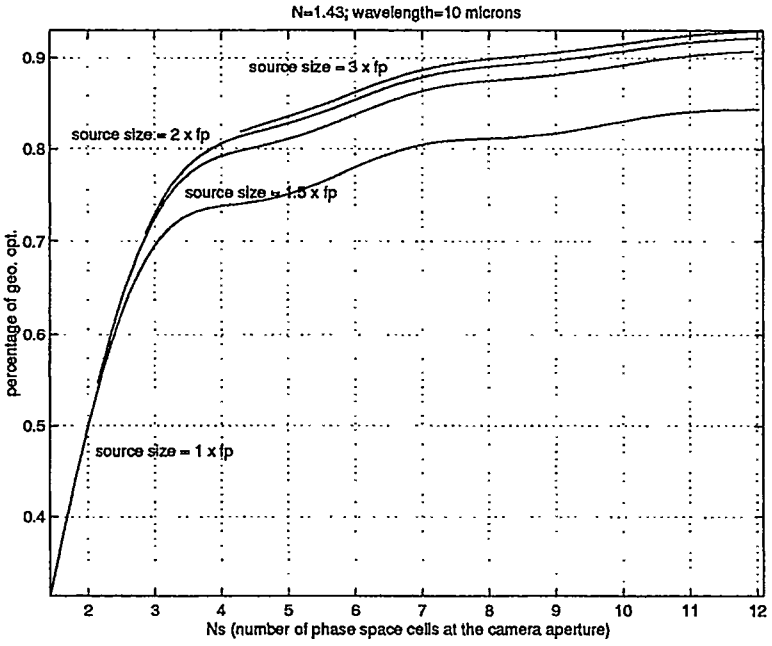

Figure 5: Global behavior. Plot of the theoretical predicted height as a fraction of the geometrical optics predicted height as a function of N_s for different F values.

5 Acknowledgements

We have greatly benefited from discussions with Robert Littlejohn. This work was supported in part by the U.S. Department of Energy, Engineering Research Program of the Office of Basic Energy Science.

References

[1] A. Walther, "Radiometry and coherence," J. Opt. Soc. Am. 58, 1256-1259 (1968).

[2] A. Walther, "Radiometry and coherence," J. Opt. Soc. Am. 63, 1622-1623 (1973).

[3] Mandel and Wolf, *Optical Coherence and Quantum Optics*, Cambridge University Press (1995).

[4] R.G. Littlejohn and R. Winston, "Generalized radiance and measurement," J. Opt. Soc. Am. A. Vol. 12, No. 12(1995) 2736-2743.

[5] R.G. Littlejohn, "The Semiclassical Evolution of Wave Packets," Physics Reports, June (1986).

[6] R. Winston and X. Ning, "Generalized radiance of uniform Lambertian sources," J. Opt. Soc. Am. A. Vol. 5, April(1988) 516-619.

[7] R.G. Littlejohn and R. Winston, "Corrections to classical radiometry," J. Opt. Soc. Am. A. **10**, 2024-2037 (1993).

LUBRICATED TRANSPORT

D.D. Joseph
Aerospace Engineering and Mechanics
University of Minnesota
Minneapolis, MN 55454

Lubricated transport of viscous liquids and solids are discussed. Distinctions are made between lubrication of core annular flows in which water is continuously added and emulsions in which water is present. Further distinctions are between oil in water emulsions, in which lubrication occurs because oil particles migrate away from the wall, and water in oil emulsions that self-lubricate due to the breaking of the emulsion near the wall at high shear stress.

SOME FLOW TYPES

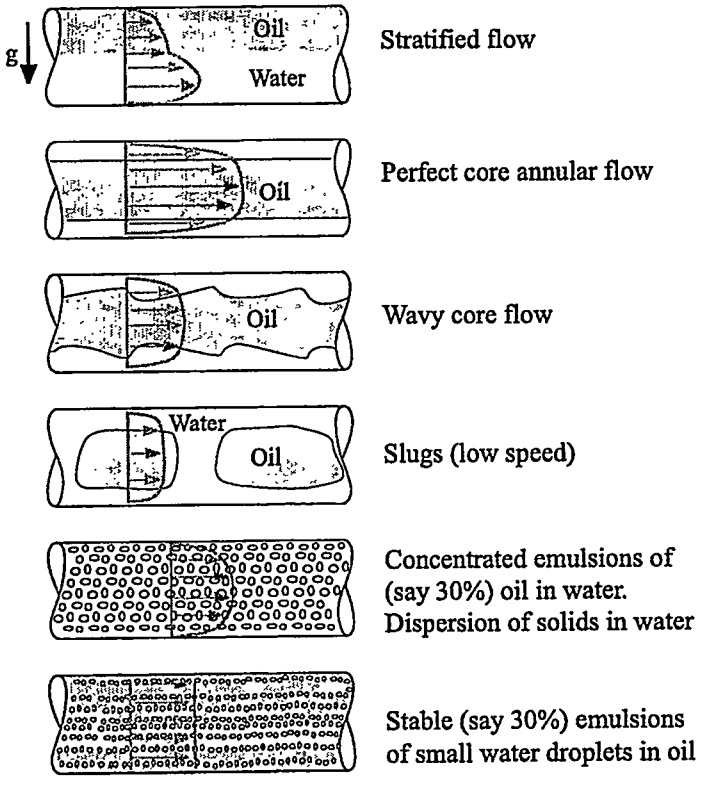

Figure 1. Some Flow Types

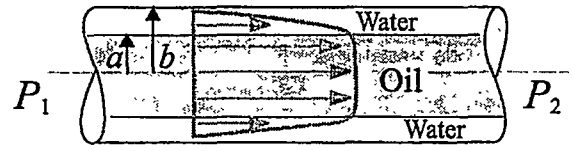

Figure 2. Ideal Core-annular Flow

Oil moves as a rigid body impelled forward by the pressure $\Delta P = P_1 - P_2$ and resisted by the shear stress in the water. Choose *a/b* to maximize the total volume flux of oil and water (o/w) for a given ΔP. This problem has a solution $a \neq 0$.

You can transport very viscous oil in water more cheaply than water alone. You get drag reductions of the order

$$\eta_{oil} / \eta_{water} = 1000 \frac{1}{100} \approx 105$$

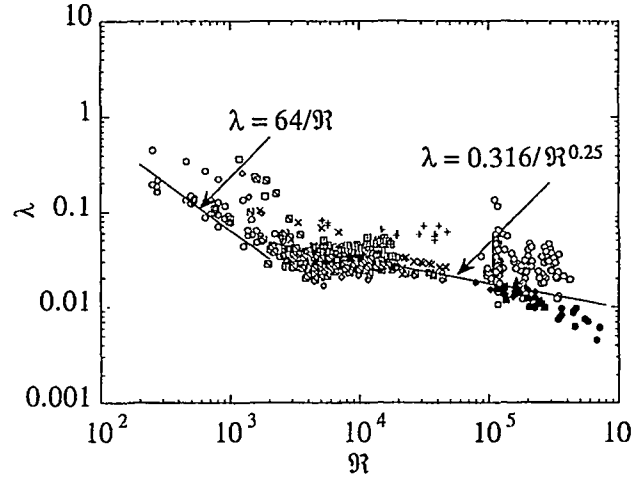

Figure 3. Laminar and Turbulent Core-annular Flow

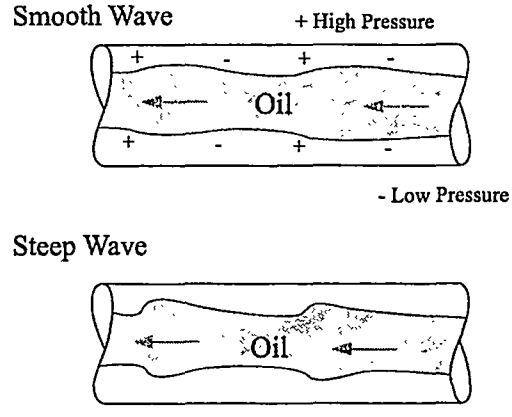

Figure 4. Steep Waves Arise from Smooth Waves

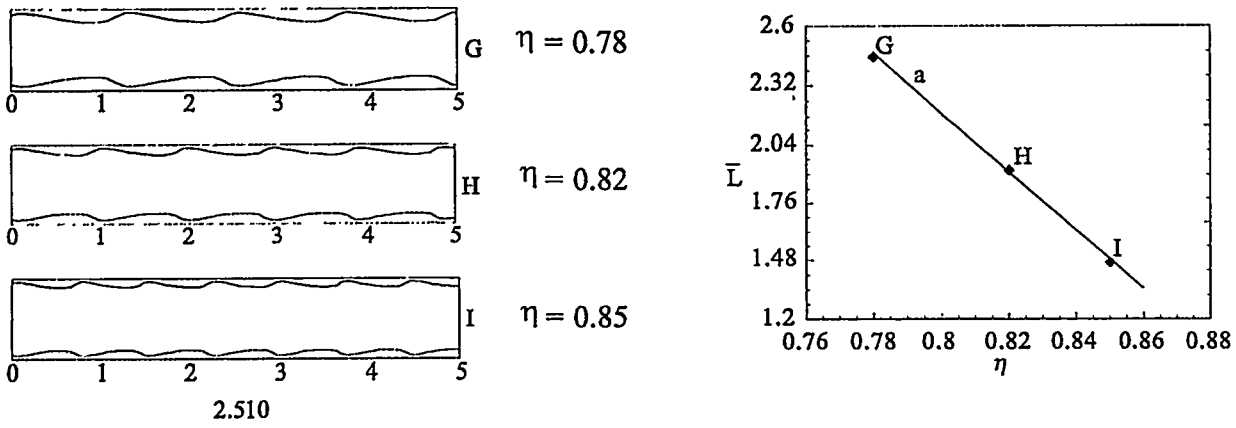

Figure 5. Wave Shortening and Sharkskin

Numerical calculation of BKJ (1996). Wavelength L = 13.5 - 14.1η for (IR, h) = (600, 1.4). The wavelength and amplitude tend together to zero as $\eta \rightarrow 1$ (see JBCR 1997).

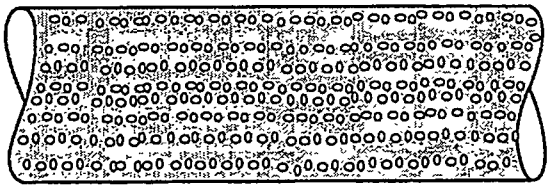

Figure 6. Self-Lubrication

At a critical value of the velocity, the emulsion breaks away from the wall and self-lubricates.

The formation of lubrication layers in o/w emulsions requires that the emulsion breaks and forms a lubricating layer at the wall. This is self-lubrication because water is not added. There are no papers other than the two here on this subject.

Self-Lubrication of Bitumen Froth

The fouled wall is an excellent wall preparation

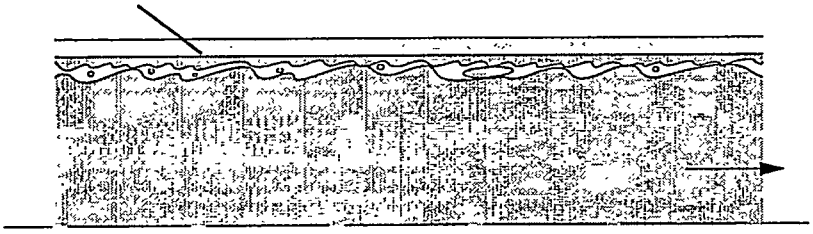

Figure 7. Mechanism of Self-Lubrication "Powdering the Dough"

After the froth breaks it remains lubricated.

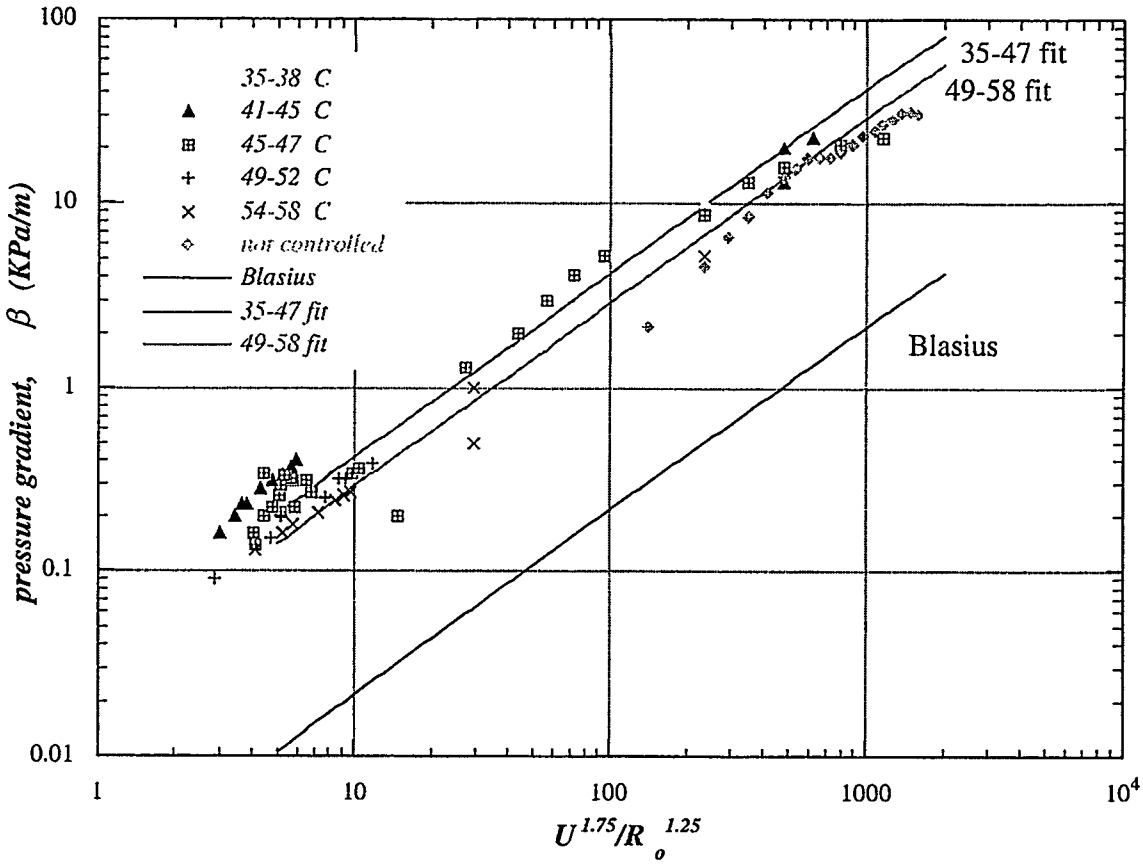

Figure 8. Blasius Correlation for Self-Lubricated Bitumen Froth

SELF-LUBRICATION of an EMULSION of WATER in MIDWAY SUNSET CRUDE

The only other experiment on self-lubrication of oil in water emulsions were done in a 1/2-inch pipeline by Veet Kruka at SHELL HOUSTON. His results are for Midway Sunset Crude oil and are described in his patent. In this case there are no clay particles, nothing special.

Self lubrication of w-o emulsions is of interest to other oil companies but there is *no data* other than Kruka's patent and our bitumen froth. We would like to know effects of *crude oil type*, *water fraction*, *pipe size*, etc.

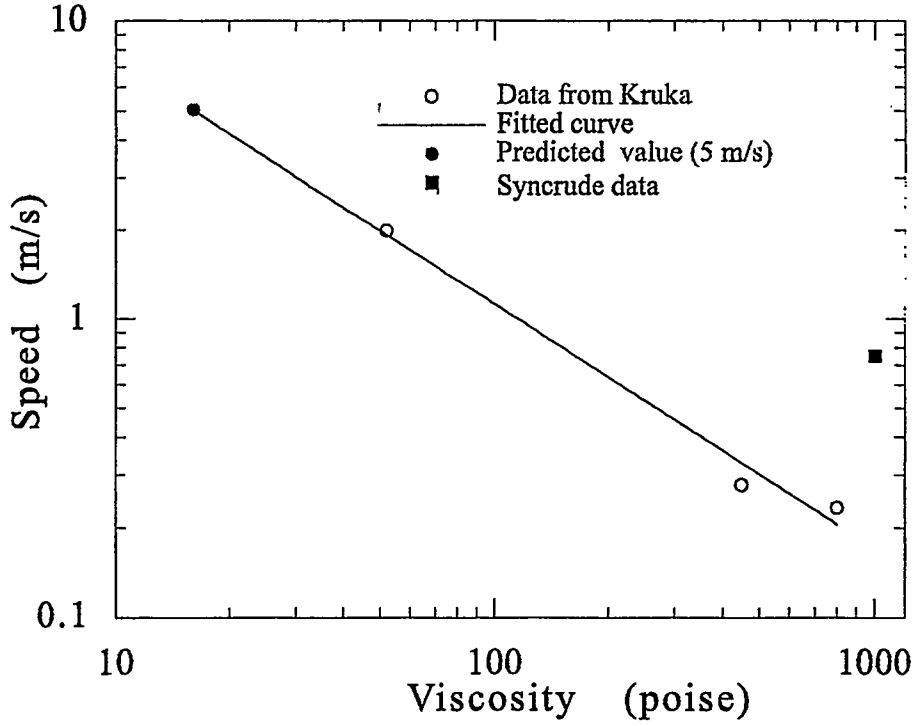

Figure 9. Critical speed vs. viscosity

Kruka's three data points fall on a straight line (see figure 9). The critical speeds for self lubrication are smaller when the viscosity is larger. Our Syncrude data does not lie on the line but it is for a 1-inch rather than 1/2-inch pipeline. We don't know of any other published data.

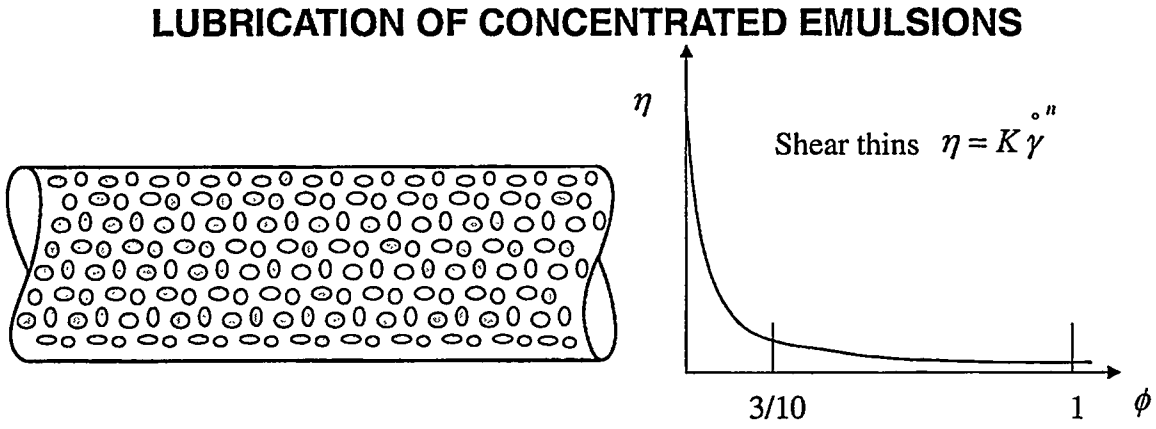

Figure 10. Concentrated Emulsion say 70% oil(left). Viscosity reduction (right).

70% oil is unstable against inversion; stabilized with surfactants The viscosity reduction is a form of lubrication.

Three Commercial Lines Have Used o/w Emulsions to Transport Oil:

1. Indonesia (Shell) carries 40,000 barrels/day of 70% waxy crude in 20"×238 km pipeline.
2. California (Shell) carries 50% heavy crude in 8"×13 mile pipeline.

3. Venezuela (Bitor) carries 70-80% heavy crude in 36"×300 km pipeline.

Other oil companies are working with us to evaluate this transportation option.

The MAIN QUESTION is if, when and how o/w emulsions can be made to enter into core annular flow, giving an additional benefit.

"Self-lubrication" of o/w emulsions involve *migration* of oil away and water toward the pipe wall.

Self-lubrication of w/o emulsions involve *breaking* the emulsion at the wall; it is altogether different.

Because the o/w emulsions shear thin, it is not easy to tell if they self-lubricate.

Comparison of Rheometer and Pipeline Data

We can use the rheometer studies giving K and n to predict pipeline data for a shear thinning fluid.

Dodge-Metzner correlations:

$$R = \frac{D^n V^{(2-n)} \rho}{gK8^{(n-1)}} \qquad f = \frac{D\Delta P g}{2L\rho V^2}$$

Reynolds number *Friction factor*

Laminar flow $$f = \frac{D\Delta P g}{2L\rho V^2}$$

Turbulent flow $$\frac{1}{\sqrt{f}} = \frac{4}{n^{0.75}} \ln\left[Rf^{\left(1-\frac{n}{2}\right)} \right] - \frac{0.4}{n^{1.2}}$$

Rheometer Studies

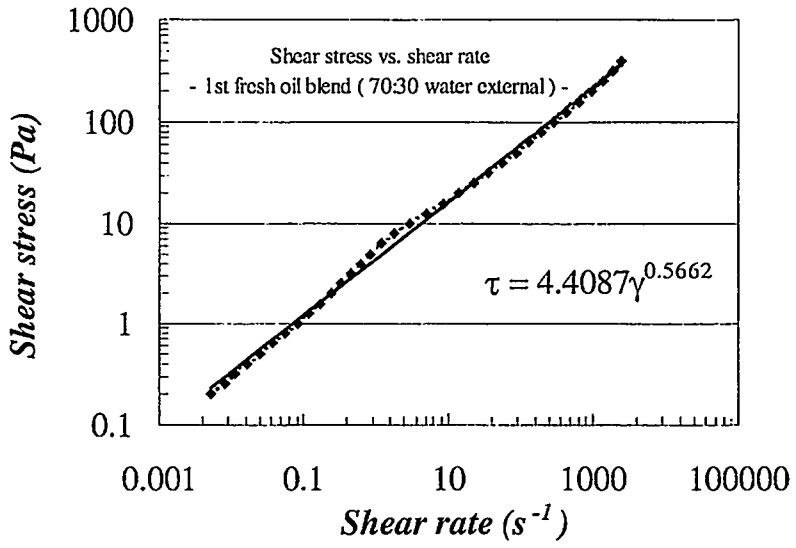

Figure 11. Shear stress vs shear rate.

Rheometer studies are used to get K and n for

$$\tau = K\gamma^{n}$$

The sample may lubricate in the rheometer. If the sample lubricates you will get a different K and n when you change the distance between the plates.

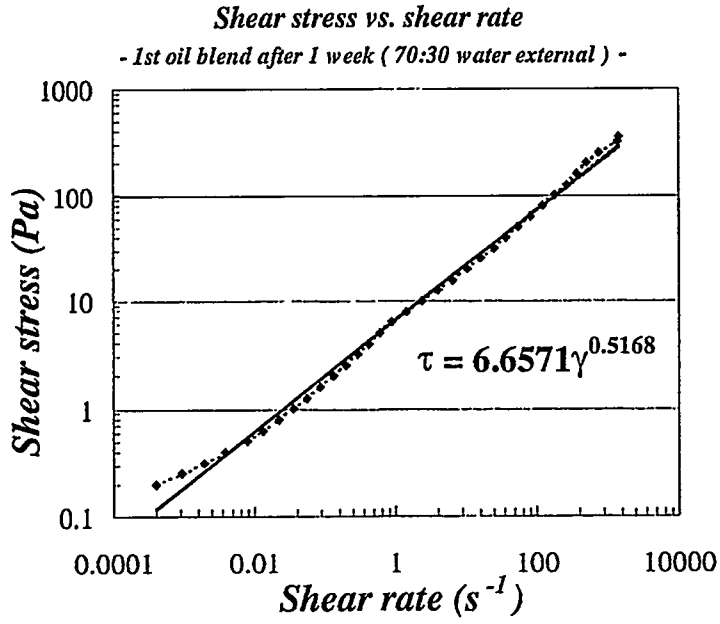

Figure 12. Shear stress vs shear rate.

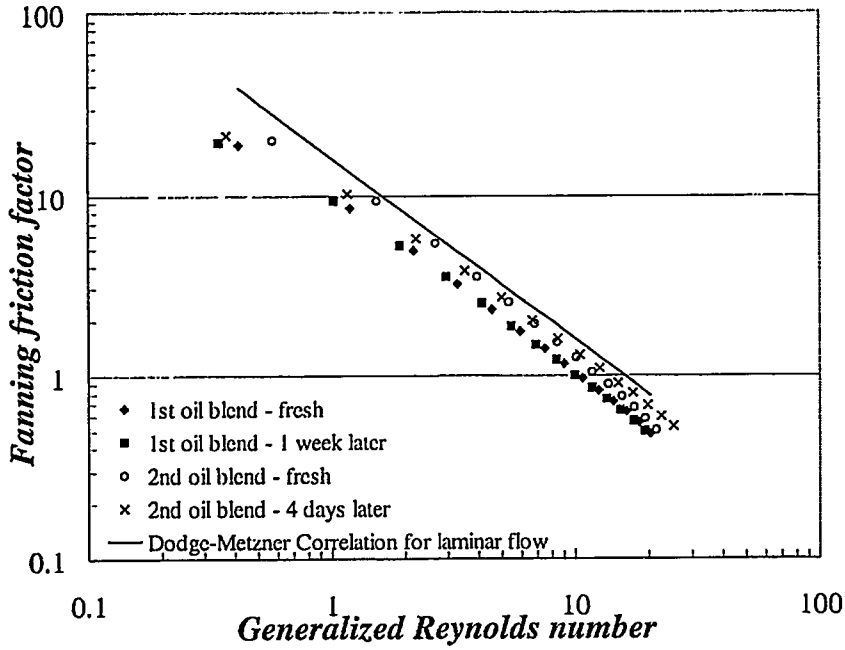

Figure 13. Friction Factor vs. Reynolds Number

The data points are below the theory line suggesting lubrication

Lubrication of Solids in Liquids

Lubrication occurs when particles migrate away from walls. We study this by direct numerical simulation, see

www.aem.umn.edu/Solid_Liquid_Flows

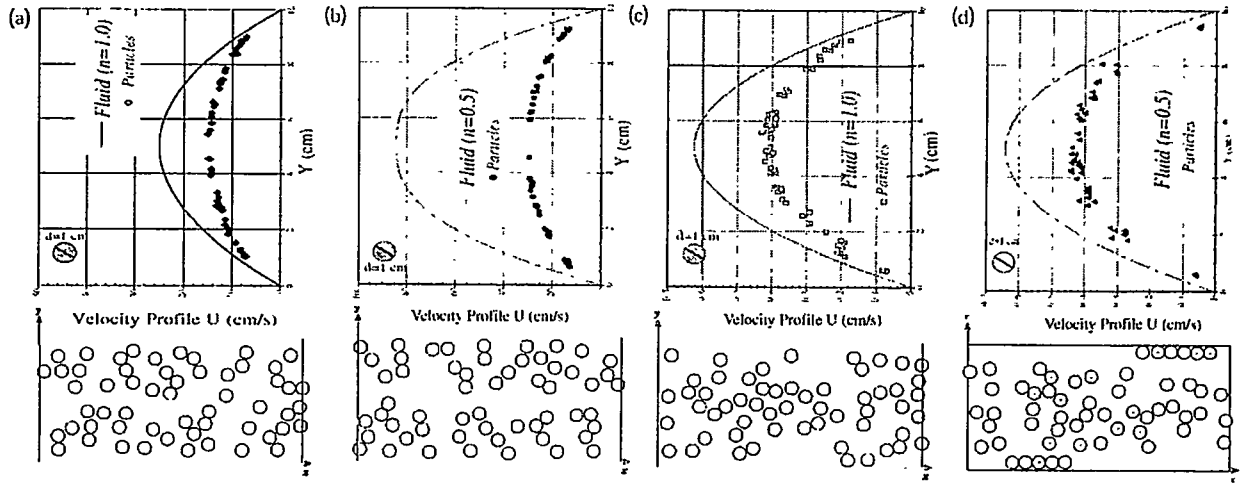

Figure 14. Migration of Neutrally Buoyant Particles in Pressure Driven Flow by DNS (Huang & Joseph JNNFM 1999).

You can isolate and study effects by switching physics on and off in the simulations that you could not do in experiments. (a) Newtonian, (b) Generalized Newtonian with shear thinning index n=0.5, (c) viscoelastic, (d) viscoelastic with shear thinning.

INTERFACIAL WAVE TRANSITIONS IN LIQUID-LIQUID FLOWS AND INSIGHT INTO FLOW REGIME TRANSITION

M. J. McCready, B. D. Woods, M. R. King
Department of Chemical Engineering
University of Notre Dame
Notre Dame, Indiana 46556 USA

ABSTRACT

Measurements of developing interfacial waves on oil-water channel flows show that long wave modes form and grow to large amplitude even though they have much smaller linear grow rates than shorter waves. There is evidence for a "triggering" of these long waves by interaction with much shorter waves, although most of the energy for wave growth comes from the mean flow. Thus linear instability of these long waves is a necessary condition for their formation and consequently, for flow regime transitions from a stratified state. However, experiments in a rotating Couette flow show regimes of no wave growth, even when long waves are unstable. The apparent reason for this is given by numerical integration of the equations that describe weakly-nonlinear wave modes at the interface. The simulations show a cascade of energy from long to short waves and no preferred wavenumber in the spectrum.

INTRODUCTION

Multifluid flows exist in oil wells, oil production and transportation pipelines, heat exchangers, gas-liquid reactors with solid catalyst and various other process piping and vessels. An important emerging issue for multifluid flow research will be how to best solve the contacting/mass and heat transfer problems that will greatly increase, as a new generation of "molecularly-engineered" catalysts developed with much higher dispersion of active metal and more elaborate possibilities of interconnection of pores on different scales. However, given the current uncertainty that exists in the simplest case, gas-liquid flow in pipe, these new problems may be difficult to solve.

Even in light of the need to understand multifluid flow on small scales, their defining characteristic, in channels, pipes and even packed beds is the strength of the *largest* scale disturbances present. For gas-liquid pipe flows, where 6 different flow regimes are possible, slug flow [1] is the regime with large coherent disturbances cause large pressure fluctuations [2] and variations in the gas and liquid flow rates that can affect process equipment. For gas - liquid packed bed flows, the corresponding region is the pulsing flow regime[3], for which the large disturbances have been shown to have the beneficial effect of increased mass transfer rates that can favorably affect the reaction outcome[4].

The existing problem in the prediction of large disturbances leading to slug formation is there are multiple mechanisms that are at work [5] and slugs can form directly from growth of waves on flat layers or by coalescence of several large rollwaves. The standard techniques for the prediction of slugs are various linear stability theories, based on different assumptions and some work that addresses the stability of a slug once it forms. Figure 1 shows several such models. It is readily seen that significant disagreement exists between the different procedures for slug prediction -- even those that are based on the same premise of unstable long waves. If these models are plotted for a model oil-gas flow at 100 ATM in a larger pipe, even bigger disagreement exists. From these results it can be concluded that considerable uncertainty exists in the prediction of slug flow for engineering purposes.

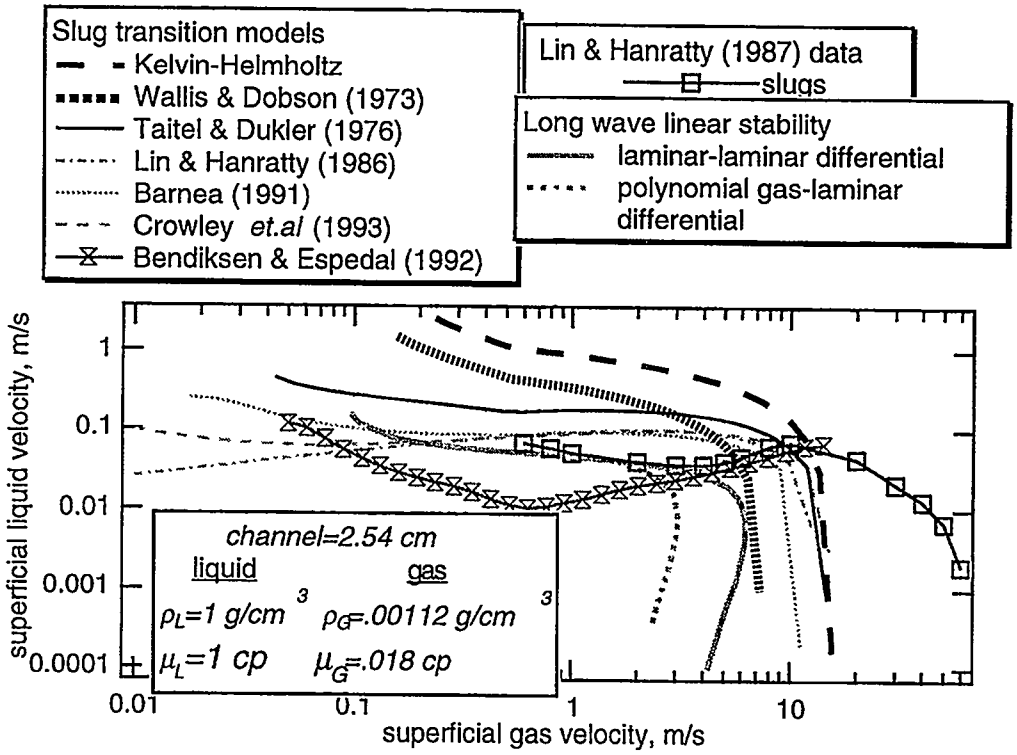

Figure 1. Different slug transition models for air-water in a 2.54 cm, horizontal pipe.

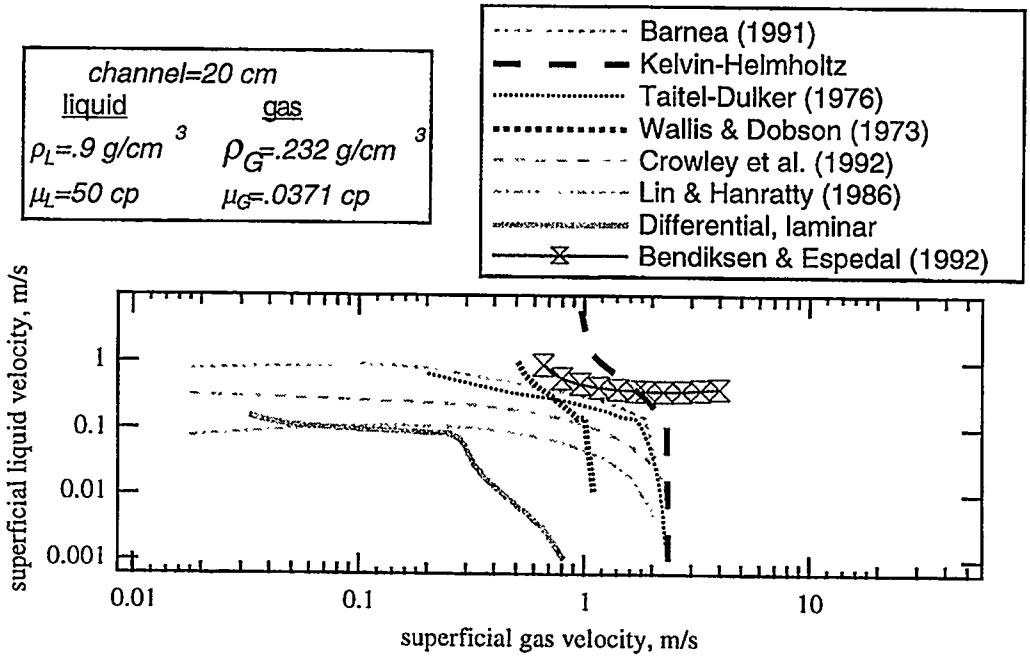

Figure 2. Different slug transition models for oil-gas in a 20 cm horizontal pipe at 20 ATM.

The specific problems addressed in this paper are the mechanisms of wave development on a two-layer stratified flow. Experiments are presented which show that the presence of a long wave instability does not mean that large disturbances will form -- casting doubt on the use of linear stability as a predictive tool for the transition from stratified to slug flow. The development of waves in an oil water channel flow is examined, as a function of distance, showing the development of a long wave peak after the more unstable short waves appear. Numerical simulation of the weakly-nonlinear mode equations is used to provide an explanation of the observed experimental behavior.

EXPERIMENTAL SYSTEMS

Figure 3 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. 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[6].

Figure 4 shows the optical system that is used to obtain data in the oil-water channel flow. The behavior of the interfacial waves is obtained by measuring the time varying wave slope with an optical refraction technique. A laser beam is split into 2 vertical beams a distance db apart and focused onto the interface. The beams are refracted at the interface according to Snell's law due to the instantaneous wave slope. The refracted beams are focused onto position sensing detectors, which provide the displacement of the refracted beam in x,y coordinates from its initial vertical position. From the geometry of the optics and the location of the refracted beam, a time series of the interfacial wave slope is created for each of the two beams. The signals from the two beams can be used to measure the wave velocities.

Experiments were also done using two matched-density liquids in a rotating Couette device. Details have been published previously[7,8].

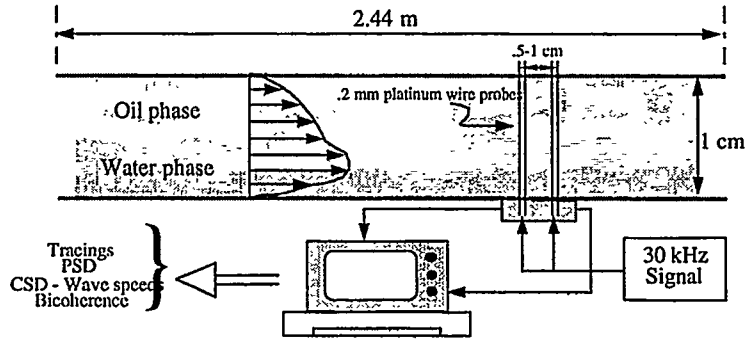

FIGURE 3. OIL-WATER CHANNEL FOR STUDYING INTERFACIAL WAVES

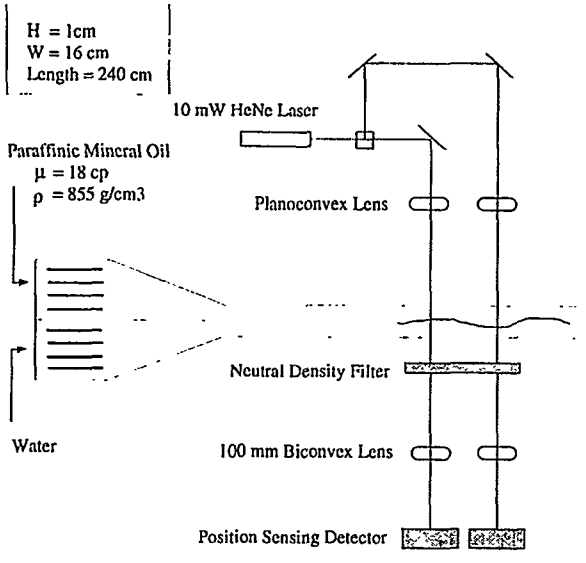

FIGURE 4. SCHEMATIC OF LASER-SLOPE DATA ACQUISITION SYSTEM FOR CHANNEL FLOW

THEORY

Theoretical analysis for this system is based on the complete two-layer Navier-Stokes equations and boundary conditions. The linear stability problem has been solved by Yih[9]and Blennerhassett [10] among others. The weakly nonlinear problem has been formulated with a multiple scales technique by Blennerhassett [10] and an eigenfunction, center manifold approach by [11] and [12]. Instead of confining our analysis to the Stuart Landau equation [12]

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

where A is the complex wave amplitude, $L(\lambda)$ is the linear eigenvalue, β is the Landau coefficient,
we do not use a center manifold approach to simply the equations. The result is then a system of many mode equations of the form

$$\dot{A}_{nl} = L_{nl}A_{nl} + \sum_{p,q,r,s} \vartheta_{nl,pr,qs} A_{pr} A_{qs} + \sum_{p,q,m,r,s,z} \xi_{nl,pr,qs,mz} A_{pr} A_{qs} A_{mz} \qquad [2]$$

where the nonlinear interaction coefficients, for the quadratic terms, ν and the cubic terms, ξ, are functions of liquid depth and wavelength and weaker functions of the degree of shear and the shape of the velocity profile. These equations are solved by numerical integration. Further details on the derivation of these equations and their solution are given in a paper from our group[13].

RESULTS

Figure 5 shows our previous data [7] that indicate regions of the rotation rate- depth ratio space where long waves are linearly unstable and no waves appear. Thus according to the premise of most slug formation theories, large disturbances are expected in this region. Even though this experiment is not channel flow, it calls into question the idea of using instability of long waves as a general criteria for flow regime transition.

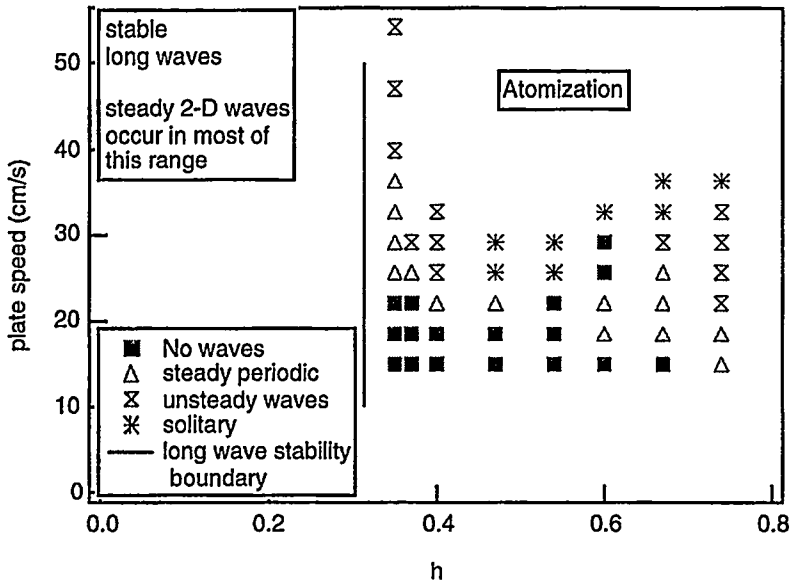

Figure 5. Wave regime map of a rotating Couette flow (left), Simulated spectrum for conditions of h= 0.4, U = 25 cm/s.

A movie of the simulations at h = 0.4 and U = 25 cm/s is available at http://www.nd.edu/~mjm/ specsim.mov. It shows that as the waves grow, there is a cascade of energy from long to short waves, that acts to stabilize the formation of long waves. Further the apparent absence of any waves can be possibly be attributed to the lack of preferred wavelength. Figure 6 shows these spectra at different times during the simulation. At the shortest time, the spectrum matches the linear growth curve. At all longer times there is a continual broadening of the spectrum. However, there is never a clearly-defined wavelength that could be visible in experiments. All of the apparent peaks oscillate in magnitude.

Figure 7 shows wave spectra for a developing oil-water channel flow. It is seen that at the first two positions, the spectra match the predictions of linear growth. However at 60 cm, the spectrum shows a number of distinct peaks that are involved in nonlinear interactions. This is confirmed by the bicoherence spectrum of figure 8. Bicoherence spectra [14] show the strength of quadratic

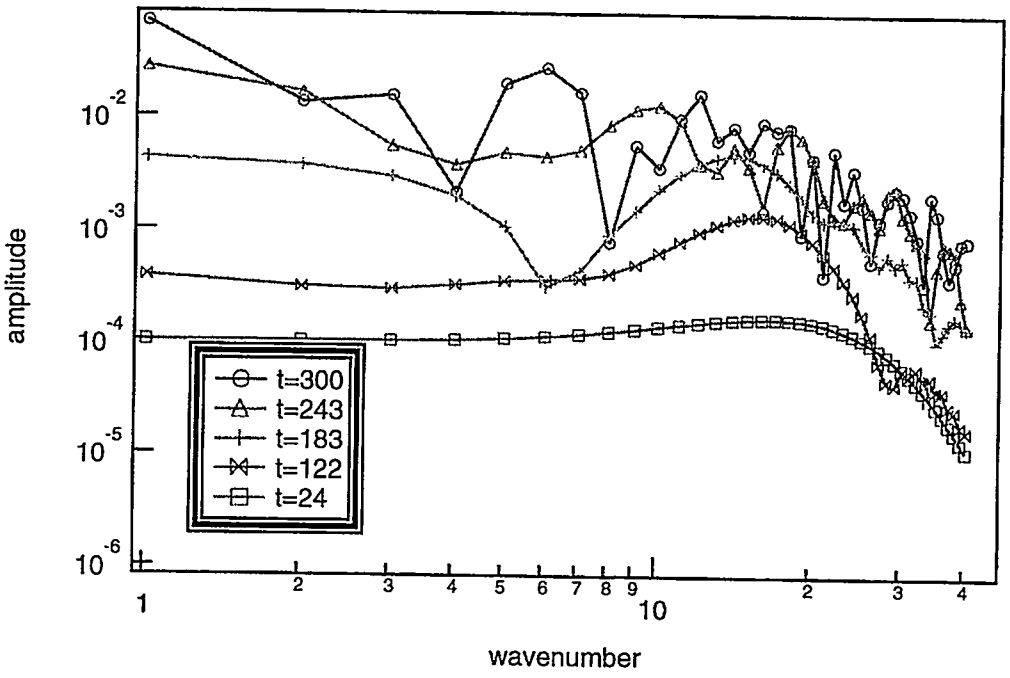

Figure 6. Amplitude spectra for Couette flow. At longer times there is no preferred wavenumber.

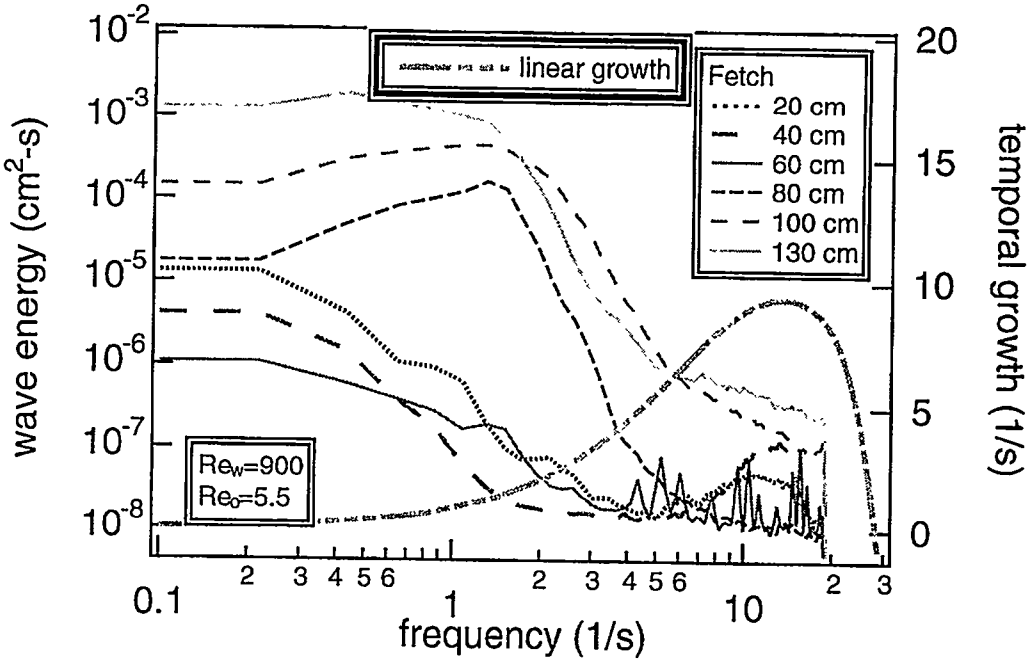

Figure 7 shows the spectra of a developing oil-water channel flow for Rew=900 and Reoil=5.5. The linear growth curve is also shown.

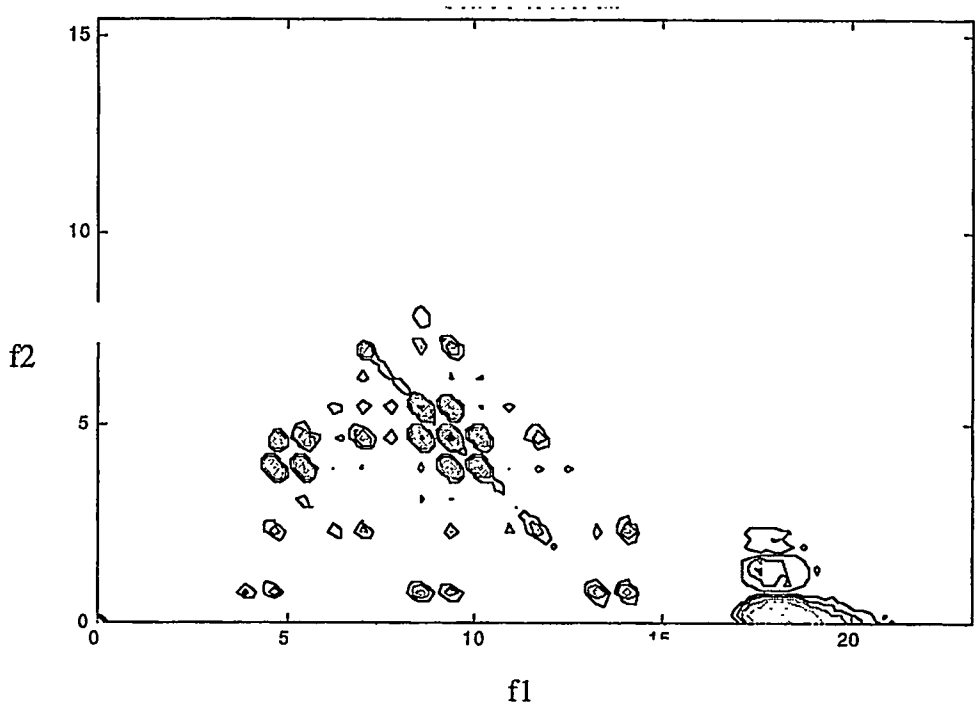

Figure 8. Bicoherence spectrum for oil-water flow of figure 7 at 60 cm. Significant coherence exists between short-short (5,10 Hz) and long short (18,1 Hz).

interactions of the form f1 + f2 = f3. Perfectly coherent modes will have a value of unity. Figure 8 shows that there are interactions between some distinct peaks around 10 Hz and their overtones (f1 = f2) (that are not shown on the power spectra of figure 7) and also with their corresponding subharmonics around 5 Hz. For this subharmonic, f1= f2= 5 Hz. A very strong interaction is seen between a wave mode of about 18 Hz and 1.5 Hz. This is probably a *difference* interaction, f1-f2 = f3, so that the other mode involved is probably about 16.5 Hz. This long-short interaction may be responsible for triggering the formation of a low frequency mode, 1.5 Hz, that is seen to growth substantially between 60 and 80 cm and which completely changes the character of the spectrum.

DISCUSSION

The experiments and simulation presented above provide insight into several different issues. First, figures 1 and 2 show that current predictive methods for the transition to slug flow differ greatly and probably none can be trusted to give reliable results. Second, the presence of no waves in regions where long waves are unstable, calls into question any methods for which long wave stability is the sole criterion for transition. It does appear, however, that long wave stability is a necessary condition for instability. The spectral simulations for the Couette flow suggest that a reason for the absence of visible waves is the lack of a persistent dominant wavenumber. This statement has not been confirmed and it is still possible that imperfections on the experiment may be the reason for the lack of observed waves.

The spectra of developing oil-water flows show that linear growth is followed by nonlinear interactions that can cause subharmonics or trigger low wave modes. These low modes can be precursors of roll waves and slugs. Subharmonics have also been implicated in the transition to slug flow[15].

These results suggest a need for improved procedures for prediction of regime transition that account for the nonlinear processes. They also suggests methods for controlling the transition and perhaps picking the frequency of the large disturbances -- by controlling the frequency of shorter waves that trigger these long wave modes.

Finally because the large disturbances that occur in packed-bed flows[3], pulses, are very similar to slugs in pipe flows -- but not as easy to study -- it could be profitable to look for analogies between the two different systems to better understand the reaction processes that are performed in gas-liquid catalytic reactors.

ACKNOWLEDGMENT

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

REFERENCES

1. P. Y. Lin and T. J. Hanratty, "Prediction of the initiation of slugs with linear stability theory", *Int. J. Mult. Flow.* **12** 79-98 (1986).
2. P. Y. Lin and T. J. Hanratty, "Detection of slug flow from pressure measurements", *Int. J. Mult. Flow.* **13** 13-21 (1987).
3. D. A. Krieg, J. A. Helwick, P. O. Dillon and M. J. McCready, "Origin of disturbances in cocurrent gas-liquid packed bed flows", *AIChE J.*, **41**, 1653-1666, (1995).
4. R. Wu, M. J. McCready and A. Varma, " Effect of Pulsing on Reaction Outcome in a Gas-Liquid Catalytic Packed-Bed Reactor ", *Catalysis Today*, **48,** PP 195-198, (1999).
5. Z. Ruder, P. J. Hanratty and T. J. Hanratty, "Necessary conditions for the existence of stable slugs", *Int. J. Mult. Flow.* **15** 209-226 (1989).
6. William McKee, "An experimental study of interfacial waves in cocurrent oil-water flows" -- M.S. Thesis, University of Notre Dame, (1995).
7. 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).
8. 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).
9. C. S. Yih, "Instability due to viscosity stratification", *J. Fluid Mech.* **27**, 337-352 (1967).
10. P. J. Blennerhassett, "On the Generation of waves by wind," *Proc. R. Soc. Lond.* A **298**, 451-494 (1980).
11. 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).
12. 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).
13. M. R. King and M. J. McCready, "Weakly nonlinear simulation of planar stratified flows", *Physics of Fluids*, **12**, PP. 92-102, (2000).
14. 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).
15. Z. Fan, F. Lusseyran, F. and T. J. Hanratty, "Initiation of slugs in horizontal gas-liquid flows, "*AIChE J.* **39** 1741-53 (1993).