DAMAGE MECHANICS – FAILURE MODES

Dusan Krajcinovic and Milena Vujosevic

Arizona State University, Tempe, AZ 85287-6106, USA

Abstract

The present study summarizes the results of the DOE sponsored research program focused on the brittle failure of solids with disordered microstructure. The failure is related to the stochastic processes on the microstructural scale; namely, the nucleation and growth of microcracks. The intrinsic failure modes, such as the percolation, localization and creep rupture, are studied by emphasizing the effect of the micro-structural disorder. A rich spectrum of physical phenomena and new concepts that emerges from this research demonstrates the reasons behind the limitations of traditional, deterministic, and local continuum models.

INTRODUCTION

All failure modes during the inelastic deformation of materials are classified into two classes: intrinsic or material failures and extrinsic or structural failures. A failure will be referred to as being intrinsic if its threshold can be defined in terms of the effective material properties. The thresholds of extrinsic failure modes depend also on the specimen size and shape. The threshold and type of failure depends on many causes. Temperature, strain rate, embrittleing effect of chemical reactions are some of the most prominently studied agents which affect the brittle to ductile transition. This study is concerned only with the brittle and quasi-brittle failures which occur in solids with a relatively modest cohesive strength as a result of damage evolution.

Damage evolution is caused by the nucleation of new microcracks and growth of the already existing microcracks. The pattern and type of damage evolution is controlled by the stress concentrations at microstructural heterogeneities (hot spots) and/or by existence of internal surfaces of inferior cohesive strength (weak links). In micro-heterogeneous (damage tolerant) solids made of materials with strongly dissimilar fracture strength (fibrous and particulate composites) the microcrack growth can be impeded by the strong phase. The microcrack growth is a basic damage evolution mode in damage sensitive materials, which are characterized by statistically homogeneous microstructures.

A typical engineering material has a disordered microstructure. The spatial distribution of hot spots and weak links is not deterministic. Consequently, the distribution of defects is not deterministic either. As long as the concentration of microcracks is dilute the microstructure is often statistically homogeneous on a rather small scale and the specimen response does not depend on the exact locations of defects. At substantial microcrack concentrations the local fluctuations of stress field depend strongly on the microcrack interaction and, therefore, on the distribution of the distances separating closely spaced microcracks. Additionally, the failure can often depend on the

largest defect or on the smallest distance between interacting defects. It seems, therefore, logical that a purely deterministic (continuum) model of failure may in many cases lead to a poor estimate of the failure threshold. The disorder is annealed in the case when the external stimuli (such as temperature, earthquake and wind loads, etc.) fluctuate randomly. In summary, the disorder of the microstructure of brittle materials is ubiquitous and its effect on the failure mode and the threshold is as a rule significant.

CHARACTERISTIC LENGTHS

To quantify the difference between particular failure modes it is necessary to introduce several characteristic lengths. Some of these lengths are easily singled out. For example, the specimen size L is a characteristic length which plays an important role in brittle fractures. More than one characteristic specimen length is needed when the failure threshold depends on the specimen shape.

The linear size L_{rve} of the smallest volume, referred to as the representative volume element [1], within which the material is statistically homogeneous, is also a characteristic length. Another characteristic length is the size ξ of the largest cluster of interacting microcracks. The cluster is defined [2]: (a) either as a string of concatenated microcracks or (b) an ensemble of microcracks which are close enough to affect their growth pattern and rate by direct interaction. A macrocrack and a shear band are two most commonly encountered examples of two respective types of clusters. The acoustic emission test [3] is a useful diagnostic procedure for the determination of the failure mode. The distance L_{ae} between the two consecutive acoustic signals can be used to determine the damage evolution mode and the failure type.

For computational efficiency a system is often subdivided into many sub-systems (finite elements) within which the considered fields are determined using an ad hoc selected interpolation rule. However, the discretization is not only the matter of efficiency since the defects in a sub-volume see the defects in the adjacent sub-volumes only through the effective moduli. Hence, the defects belonging to different sub-volumes are not correlated. Local fluctuations of the stress, strain and damage fields in a sub-volume will affect the response in the rest of the specimen only if the material within the considered sub-volume is statistically homogeneous, i.e. if the conditions

$$\xi \leq L_{rve} \leq L \quad \text{and} \quad L_{rve} \geq L_{tr} \tag{1}$$

are satisfied. In (1) L_{tr} is the wave length of the tractions applied to the external surfaces of the sub-volume. First of two inequalities (1) renders the material statistically homogeneous while the second condition secures homogeneous response. If the inequalities (1) are satisfied the effective continuum, obtained by mapping representative volume elements on material points, is local.

The distribution of distances between the consecutive acoustic signals is uniform, $p(L_{ae}) = const.$, in damage evolution processes controlled by microcrack nucleation, i.e. in materials with a random distribution of weak spots. A propagation of a single crack is characterized by the fact that the majority of signals are clustered at the crack tip. The distances between pairs of consecutive acoustic signals are equal to the resolution length, $L_{ae} \approx l$. The deformation process is dominated by the interaction induced microcrack growth (cooperative effect) if the distribution of lengths L_{ae}: (a) favors small distances in the direction of the normal to a rather thin elongated region, and (b) is perfectly random in the exterior of this region. The material within the localized region is non-local and non-homogeneous while the material in the exterior volume remains local and homogeneous. The damage evolution and the deformation process are dominated by the kinetics of the largest cluster of interacting microcracks (shear band), i.e. by the rate at which the correlation length ξ increases.

On the microscopic scale the damage evolution depends on the outcome of the competition between the microcrack nucleation and microcrack growth. The intrinsic brittle to quasi-brittle transition, defined as a cross-over from the processes controlled by a single crack to processes dominated by many microcracks, depends on: (a) sign of the long range stresses and (b) the bandwidth of the distribution of fracture strengths of the constituent microstructural phases [4],[2]. The traditional, deterministic definition of the brittle to quasi-brittle transition [5], based on the macroscopic phenomenology, is strongly dependent on the specimen size and shape and is, therefore, not a material parameter.

EXTRINSIC FAILURE MODES

Failure of a micro-homogeneous (rock, concrete, glass or ceramic) specimen subjected to a uniaxial tensile stress is a quintessential example of a perfectly brittle failure. The failure occurs when the largest and preferentially oriented (perpendicular to the tensile axis) existing crack starts propagating. The energy barriers in these materials are insufficient to trap the crack rendering the propagation of the growing crack unstable. If the specimen is subjected to a homogeneous stress field it will fail in an almost instantaneous cleavage mode. The force-displacement relation remains linear until the failure is reached indicating that the accumulation of distributed damage is minimal. The failure threshold is a variate dependent on the specimen size since the probability of finding a large pre-existing crack is proportional to the volume. The statistics of extremes represents an appropriate framework for the determination of the failure threshold as a function of stress and specimen size [6],[7].

The failure of an unconfined specimen subjected to uniaxial compression in a longitudinal splitting mode also belongs to the class of extrinsic failures. The growth of cracks, which are roughly parallel to the compression axis, is initially marginally stable [8]. As soon as the largest of these cracks reaches its critical length the specimen splits into two or more parts. A modest concentration of accumulated damage does not affect the specimen failure. The failure threshold can be predicted from Griffith's. The specimen failure depends on the specimen shape (ratio between its length measured along the compression axis and its width in the lateral dimension). Specimen will split only when this aspect ratio is larger than one (slender cylinders). At small aspect ratios the part near the lateral surface will slab off while the rest of the specimen will resist further load increments.

INTRINSIC FAILURE MODES

The threshold of an intrinsic failure mode depends only on the average stresses, average strains and effective material properties. For a material to fail in an intrinsic mode it must be statistically homogeneous (invariant with respect to translations) on a scale of the sub-system, i.e. the conditions (1) must be satisfied. Conditions (1) are satisfied in the mean-field regime, characterized by a dilute concentration of microcracks which renders the effect of the direct interaction of microcracks on the macroscopic response insignificant.

Tensile Failure of a Damage Tolerant Solid

The damage accumulated in damage tolerant materials subjected to a long range tensile field can be substantial. In a stress controlled test a specimen will fail when the (1111) component of the effective tangential stiffness (or elastic) tensor vanishes. In a elongation controlled test the specimen response may crossover to the softening regime signaling the loss of homogeneity reflected as the loss of ellipticity.

Numerical simulations on two-dimensional, damage tolerant, frozen lattices [9] demonstrate the existence of the size effect. The response in the hardening regime can be very accurately estimated by the mean field model [10] adjusted by the size effect term derived in [9]. The response in the softening regime is controlled by large defect clusters and the stress distribution is multifractal [2]. Large scatter of results for different physical realizations of same statistics is indicative of the dominance of extreme statistics. Consequently, the mean field theories are useless for the modeling of the softening regime.

Elastic Percolation

In the absence of a characteristic length the failure threshold also depends on the effective properties since the material is self-similar (scale invariant). The self-similarity takes place when the correlation length tends to infinity during a random crack nucleation process [2]. The failure occurs when the largest cluster, formed by correlated defects, splits the specimen into two parts. This class of problems is studied using the methods of the percolation theory [11].

A typical elastic percolation test [12] consists of a repetitive sequence of two steps: (1) a defect (perforation or a slit) is randomly placed into an unloaded specimen and (b) the specimen is subjected to a very small traction to determine the effective stiffness of the plate without causing plastic deformation. These two steps are repeated until the specimen stiffness is reduced to zero, i.e. when the cluster of correlated defects splits the specimen into two large fragments. The measured stiffness is equal to the effective secant stiffness in the corresponding, deformation controlled deformation process during which the defects nucleate as a result of the monotonically increasing displacements imparted to the specimen [2].

At the percolation threshold ($\xi \to L^-$) the specimen size is the only remaining characteristic length. Due to the randomness of the defect pattern the material on the scale smaller than the correlation length is self-similar. As a consequence of self-similarity the geometry of the largest cluster is fractal and the scaling law for the correlation length is defined in the form $\xi \propto (f - f_c)^\nu$ as $f \to f_c$, where f and f_c are the density and critical density of randomly inserted defects while the exponent ν is the fractal dimension. The rate at which the order parameter or a macro parameter (the vanishing component of the effective stiffness tensor) approaches zero is controlled only by the correlation length. Hence, the scaling law for the vanishing component of the effective stiffness tensor is also fractal $C \propto (f - f_c)^g$ as $f \to f_c^-$. The exponents ν and g, are universal in being independent of the details of local fluctuations of damage, stress and strain fields.

The elastic percolation threshold f_c and universal exponents, which determine the behavior of the system near the percolation threshold, have to be determined for all elastic materials belonging to the same universal class only once. Some of these parameters are already available. The percolation thresholds and scaling laws for the effective stiffness in the case of random distribution of circular [13], [14] and elliptic [15] holes were determined for all homogeneous, isotropic two-dimensional elastic solids. The percolation thresholds for two-dimensional elastic materials weakened by rectilinear slits were determined by simulations [16] and [17]. The corresponding scaling law for the effective stiffness was derived by Krajcinovic [18]. Analytical expressions for the percolation thresholds in the case of holes [19] and slits [20] fit the simulation data extremely well.

The elastic percolation theory can be, firstly, used to compare different mean field models. The percolation model provides the limit to which the mean field estimates of the effective stiffness must tend as $f \to f_c$. Secondly, the random percolation is a correct model for: (a) the response of a rock or concrete specimen subjected to the hydrostatic compression and (b) the cracks induced during the curing of resins and solidification of metals [21].

Localization

From a continuum viewpoint, localization (or faulting) can be, defined as a bifurcation from a state of homogeneous strain to a state of inhomogeneous strain caused by the emergence of a specimen spanning shear band. The strains within the band are much larger than the strains in the exterior of the band which may even relax after the formation of the shear band. From a micromechanical viewpoint, the deformation regime preceding the localization is controlled by the microcrack nucleation while the microcrack growth is prevented by the lateral confinement and the attendant absence of the singular stress fields at the tips of nucleated microcracks. As the nucleation proceeds the probability that the direct interaction of closely spaced (correlated) cracks may lead to the appearance of singular stress fields sufficient to cause microcrack growth increases. The self-organization of microcracks into clusters is a synergistic process since the largest cluster grow faster than smaller ones.

As the clusters of interacting microcracks grow the degree to which the largest cluster controls the deformation increases. Hence, the macroscopic response is in the neighborhood of the localization a function of the correlation length ξ. The amplification of stress intensity factors also depends on the angle subtended by the cluster and compression axes [22]. The angle which optimizes the cooperative effect, varying between 20 and 30 degrees, is very sensitive to the microstructural imperfections.

The general continuum theory of localization, based on the Hadamard's concepts of stability, was formulated by Rudnicki and Rice [23] and Rice [24]. The elegance of these two papers spawned a host of analytical and computational studies. In the process of the development of several sophisticated models the sage advice that the localization conditions "depend critically on subtle features of (constitutive) descriptions" and that the "role of deformation field non-uniformities or imperfections... (are)... of great importance for the initiation and spreading of localized deformation zones" [24] was summarily dismissed. The onset of the localization was defined in terms of the acoustic tensor, i.e. by $det(\mathbf{n} \cdot \mathbf{C}^i \cdot \mathbf{n}) = 0$ where $\mathbf{C}^i$ is the effective tangential tensor at the point of bifurcation and $\mathbf{n}$ the normal to the localized band. The bifurcation is not possible in the von Mises material and the plasticity models must be embellished either by a non-associative flow rule or by allowing the formation of vertices in the yield surface.

The added problem related to the application of local continuum models is related to the absence of a characteristic length. As a result the localized zone has a zero thickness and the density of the energy dissipation is infinite. This fact triggered development of several non–local and gradient localization models.

None of these models treats the material microstructure as being random and the process as being dynamic. In consideration of correlation-induced (Mott) and disorder-induced (Anderson) localizations it was concluded that the delocalization-localization is impossible in the absence of some disorder. It is, therefore, not surprising that the deterministic models were unable to capture the essential non-deterministic nature of the phenomenon which looms behind the tests (Lockner, et al [3]). The localization belongs to the class of the short to long-range transitions. As the damage concentration increases the correlation range ξ grows from zero to the specimen size L. The correlation length ξ controls the specimen macroscopic response only in the phase preceding the onset of localization and the post-localization regime. However, well before that point the influence of ξ on the macroscopic response cannot be neglected. The onset of the localization and exact thickness of the localized band are largely in the eyes of the beholder. In a material such as granite and limestone the band "thickness" varies along its length. The material in the exterior of the band seems to be statistically homogeneous while the characterization of the interior of the non-local material within band is not as simple.

The localization process can be modeled using the molecular dynamics [25]. The disorder is quenched into the lattice. To replicate the Lockner, et al. [3] tests the selected two-dimensional, irregular, elastic triangular lattice is subjected first to biaxial compression $\sigma_{lat} = \sigma_{axial} > 0$. Subsequently, the lattice length in the direction of the longer axis is slowly decreased in order to be able to consider both hardening and softening regimes.

In the early stages of the simulation the damage evolution is controlled by weak spots (rupture of links of inferior strength). As the defect concentration increases the self-organization of defects into the largest cluster becomes evident. The increase of the correlation length ξ is initially slow and crosses over to a power law as $\xi \to L^{-}$. The true nature of the phenomenon is reflected in the transfer of the energy through the disordered lattice. The dynamic and stochastic nature of the phenomenon can be illustrated using the "ant in the labyrinth" model [11]. An ant (random walker), located within the localized band, is allowed to move randomly from a site to any of the other nearest neighbor sites providing that the link between two sites exists. The ant must stay in its site each time it makes an attempt to move along a ruptured link. In an undamaged lattice the mean square distance traversed by the ant is linearly proportional to the walking time. In a localized shear band the time the ant needs to reach the less damaged material increases with the density of ruptured links. However, as the time of walking approaches infinity all ants, starting their walk in different sites within the band, will eventually find their way out from the labyrinth (shear band) assuming that the lattice is connected. Thus, the localized band "width" depends on the walking time allotted to the ant.

As a mechanical analogy of this ant a link within the band can be subjected to a periodic change of its length. The fraction of the imparted energy recorded at a point in the band exterior depends on the damage within the band and the frequency of the imparted vibrations. At very small frequencies (statical loads) large part of the energy percolates through the band. As the frequency of the imparted vibrations increases the fraction of energy recorded in the exterior of the band decreases.

In summary, the question of what is the "width" of the localized band is not well posed. In fact, the width is a function of the correlation length and the frequency. In a purely static approximation the band width is equal to zero.

Annealed Disorder

The creep rupture of resins is an example of failure modes attributed to the annealed disorder. The time to creep rupture depends on the magnitude and sign of the applied tractions, temperature and material microstructure. Most phenomenological and micromechanical [26] models are deterministic in form. Random fluctuations of temperature and stress are neglected. A different viewpoint was taken by Regel', et al. [27] and Vujosevic and Krajcinovic [28] who treat the temperature as the stimulus which enhances the mobility of atoms and increases the probability of the dislocation motion and rupture of atomic bonds.

The microstructure of epoxy resins [21] emphasizes dense nodules interconnected by lower density material (a relatively vacuous network of cross–linked molecular chains). A triangular central force lattice with nodules as sites and molecular chains as links is a realistic model of the described microstructure. The quenched disorder is neglected and the lattice is in the pristine state both geometrically and topologically regular, and all links have identical stiffness and strength.

The damage evolution, i.e. bond rupture sequence, is treated as a random process which is activated by spatially and temporally random thermal fluctuations. The probability that a ith link will rupture during the time interval Δt , was based on the absolute reaction–rate theory [29]. The sequence of the link ruptures is determined by the Monte Carlo lottery. The simulations demonstrate that the temperature driven damage evolution leads to the percolation transition. The

specimen stiffness decreases linearly. The loss of stiffness becomes nonlinear and very rapid when the specimen is subjected to tensile loads. In all cases the strain vs time curves exhibit all trends characteristic of creep. The exponential dependence of the time to failure on the load parameter derived from simulations is in agreement with the experimentally observed trends [27].

The simulations data in [28] were used to construct approximate analytical model which fits the simulation data surprisingly well. A mean field model was derived to estimate the effect that stress concentrations have on the time to creep rupture. Final, and perhaps the most important, result was that the time to rupture does not depend on the lattice size.

SUMMARY AND CONCLUSIONS

The principal message of this research is that the disorder in macrostructure and the random fluctuations of externally imparted stimuli cannot be easily dismissed when estimating the type of failure and its threshold. Purely phenomenological and deterministic continuum models based on fitting experimental data may rather often lead to wrong conclusions. Direct and careless extensions of mean field models beyond the limits of the dilute concentration of defects is not always a prudent strategy. Abandoning familiar and elegant mean field methods may, indeed, be traumatic. At the same time a search of a more rational model, which acknowledges the loss of homogeneity, may in many cases be the only avenue to a reliable analytical prediction of the failure threshold. The development of non-local theories was in the majority of cases channeled along the familiar deterministic routes. Preservation of mathematical elegance may, indeed, be a virtuous goal. At the same time it cannot be the only guideline leading to the improvement of modeling techniques. It seems to the authors of this study that physics should have a precedent over mathematics.

ACKNOWLEDGEMENT

The authors gratefully acknowledge the financial support in form of a research grant from the U.S. Department of Energy, Division of Engineering and Geosciences to the Arizona State University which made this work possible.

REFERENCES

1. S. NEMAT–NASSER and M. HORI, *Micromechanics: Overall Properties of Heterogeneous Materials*, North–Holland, Amsterdam, The Netherlands (1993).
2. D. KRAJCINOVIC, *Damage Mechanics*, North–Holland, Amsterdam, The Netherlands (1996).
3. D.A. LOCKNER, D.E. MOORE and Z. RECHES, 'Microcrack Interaction Leading to Shear Fracture', in: *Rock Mechanics*, J.R. Tillerson and W.R. Wawersik, Eds., Balkema Publ., Rotterdam, The Netherlands, (1992).
4. S. ROUX and A. HANSEN, 'Early Stages of Rupture of Disordered Materials', *Europhys. Lett.* 11, 37 (1990).
5. M.S. PATERSON, *Experimental Rock Deformation–the Brittle Field*, Springer–Verlag, Berlin, Germany (1978).
6. W.A. CURTIN and H. SCHER, 'Algebraic Scaling of Material Strength', *Phys. Rev. B*, 45, 2620 (1992).
7. P.M. DUXBURY, S.G. KIM and P.L. LEATH, 'Size Effect and Statistics of Fracture in Random Materials', *Mat.Science and Engineering*, 176, 25 (1994).

8. H. HORII and S. NEMAT–NASSER, 'Brittle Failure in Compression: Splitting, Faulting and Brittle–Ductile Transition', *Phil.Trans. Royal Soc. London*, 319, 337 (1986).
9. A. HANSEN, S. ROUX and H.J. HERRMANN, 'Rupture of Central–Force Lattices', *J. of Physics France*,50, 733 (1989).
10. D. KRAJCINOVIC and M. BASISTA, 'Rupture of Central–Force Lattices Revisited', *J. Phys. I*, 1, 241 (1991).
11. D. STAUFFER and A. AHARONY, *Introduction to Percolation Theory*, 2nd edition, Taylor & Francis, Washington, DC (1992).
12. L. BENGUIGUI, 'Experimental Study of the Elastic Properties of a Percolating System', *Phys. Rev. Lett.* 53, 2028 (1984).
13. B.I. HALPERIN, S. FENG, P.N. SEN, 'Difference Between Lattice and Continuum Percolation Transport Exponents', *Phys. Rev. Lett.* 54, 2391 (1985).
14. L. BENGUIGUI, 'Elasticity and Fracture near a Percolation Threshold in two Dimensions', in: *Fragmentation, Form and Flow in Fractured Media*, R. Englman and Z. Jaeger, Eds.,*Annals of Israel Phys. Soc.* 8, 288 (1986).
15. K. MALLICK, D. KRAJCINOVIC, D. SUMARAC and M. VUJOSEVIC, 'Critical State of a Two–Dimensional Elastic Continuum Containing Elliptical Voids', *Engng. Fracture Mech.* 46, 553 (1993).
16. E.G. PIKE and C.H. SEAGER, 'Percolation and Connectivity: A Computer Study I, *Phys. Rev. B*, 10, 1421 (1974).
17. P.C. ROBINSON, 'Connectivity of Fracture Systems–A Percolation Theory Approach, *J. Phys. A: Math. Gen.* 16, 605 (1983).
18. D. KRAJCINOVIC, 'Scaling Law for Blue Cheese Model of Damage', *Phys. Lett. A*, 173, 433 (1993).
19. D. KRAJCINOVIC, K. MALLICK, M. BASISTA and D. SUMARAC, 'Elastic Moduli of Perforated Plates in the Neighborhood of Critical States', *Int. J. Solids Structures*, 29, 1837 (1992).
20. D. KRAJCINOVIC, D. SUMARAC and K. MALLICK, 'Elastic Parameters of Brittle, Elastic Solids Containing Slits, *Int. J. Damage Mech.* 1, 386 (1992a).
21. D. KRAJCINOVIC and K. MALLICK, 'Micromechanics of Process Induced Damage in Thermosets, *J. Mech. Phys. Solids*, 43, 1059 (1995).
22. H. HORII and S. NEMAT–NASSER, 'Compression–induced Microcrack Growth in Brittle Solids: Axial Splitting and Shear Failure', *J. Geophys. Res.* 90, 3105 (1985).
23. J.W. RUDNICKI and J.R. RICE, 'Conditions for Localization of Deformation in Pressure–sensitive Dilatant Materials', *J. Mech. Phys. Solids*, 23, 371 (1975).
24. J.R. RICE, 'The Localization of Plastic Deformation', in: *Proc. of 14th International Congres of Theoretical and Applied Mechanics*, W.T. Koiter, Ed., North Holland, The Netherlands (1976).
25. D. KRAJCINOVIC and M. VUJOSEVIC, 'Localization: dynamic transition of the correlation range', (in preparation)
26. A.C.F. COCKS and F.A. LECKIE, 'Creep Continuum Equations for Damaged Materials', in: *Advances in Applied Mechanics*, J.W. Hutchinson and T.J. Wu, Eds., 25, 239, Academic Press, New York, NY (1987)
27. V.R. REGEL', A.I. SLUTKER and E.E. TOMASHEVSKII, *Kinetic Theory of Strength of Solids*, Nauka, Moskow, USSR (in Russian) (1974).
28. M. VUJOSEVIC and D. KRAJCINOVIC, 'Creep Rupture of Polymers: a Statistical Model', *Int. J. Solids Structures* (in print)
29. H. EYRING, 'Viscosity, Plasticity and Diffusion as Examples of Absolute Reaction Rates', *J. Chem. Phys.* 4, 283 (1936).

T^*_ε INTEGRAL ANALYSIS OF FRACTURE SPECIMENS

Y. Omori [1], K. Perry, Jr. [2], J.S. Epstein[2], L. Ma[1], H Okada[3], S. N. Atluri*[3] and A. S. Kobayashi[1]

[1]University of Washington Department of Mechanical Engineering, Seattle, WA 98195-2600
[2]Idaho National Engineering Laboratory, Material Science Division, Idaho Falls, ID 83415
[3]Georgia Institute of Technology, Computational Mechanics Center, Atlanta, GA 30332-0356

ABSTRACT

T^*_ε integral values associated with stable crack growth in thin 2024-T3 aluminum compact (CT) specimens and A606 HSLA steel single edge notched (SEN) specimens were determined directly from the crack tip displacement field obtained by moiré interferometry. Stable crack growth in the SEN specimen was also simulated by an elastic-plastic finite element (FE) model which was driven by the experimentally determined boundary conditions. T^*_ε obtained experimentally and by FE were in reasonable agreements with each other. Unlike the vanishing J integrals with crack extension, T^*_ε reached steady state values with stable crack growth. Thus, for a given integration contour, Γ_ε, near the crack tip, T^*_ε can be used as a stable crack growth as well as a ductile fracture criteria.

INTRODUCTION

In a series of papers, May and Kobayashi [1] showed that the ASTM J resistance curve does not represent the crack tip state of stress since J is not a path independent integral under stable crack growth. Moreover, the near field J integral vanished after reaching a maximum value at the initial phase of stable crack growth. Brust et al [2,3], on the otherhand, showed through finite element (FE) simulation, that the near field T^* integral [4] remained a steady state value under loading, unloading, reloading and under stable crack growth.

Unlike the J integral, the T^* integral is based on the incremental theory of plasticity, T^*_ε, and is defined by Stonesifer and Atluri [4] as

$$\Delta T_{\varepsilon}^{*} = \int_{\Gamma_{\varepsilon}} \left[\Delta W \cdot n_1 - (t_i + \Delta t_i) \cdot \Delta u_{i,1} - \Delta t_i \cdot u_{i,1} \right] ds$$

where W is the work density and t_i and u_i are the traction and displacement, respectively. Here ΔT^*_{ε} denotes the ΔT^* value associated with a contour, Γ_{ε}, in the very vicinity of the crack. The proximity of Γ_{ε} is restricted to the region where plane stress condition prevails and is outside of the region of three dimensional state of stress at the crack tip. The latter is generated by the finite thickness of the plate specimen. For a flat crack without a shear lip, Narasimhan and Rosakis [5] have shown that the state of plane stress prevails outside of one half of the plate thickness.

In theory, the total T^*_{ε} is obtained by summing the incremental ΔT^*_{ε} through the plastic deformation process. Fortunately Pyo et al [6] have found that the total T^*_{ε} computed directly by using the stresses and strains based on the incremental theory of plasticity was for all practical purpose equal to the summed ΔT^*_{ε}. Thus T^*_{ε} can be computed directly without the cumbersome incremental summation procedure provided the stresses and strains, on which T^*_{ε} is based, are obtained through the use of the incremental theory of plasticity. Like the J integral, T^*_{ε} is path dependent in the presence of large scale plasticity and plastic unloading. Thus, if T^*_{ε} is to be used as a fracture criterion, it is imperative that T^*_{ε} be evaluated along a contour, Γ_{ε}, very close to the crack tip, such that it can be considered a crack tip parameter.

In this paper, we present the T^*_{ε} integral values which were determined **directly** from the displacement field surrounding a stably growing crack in thin 2024-T3 aluminum compact (CT) specimens and A606 HSLA steel single edge notched (SEN) specimens.

METHOD OF APPROACH

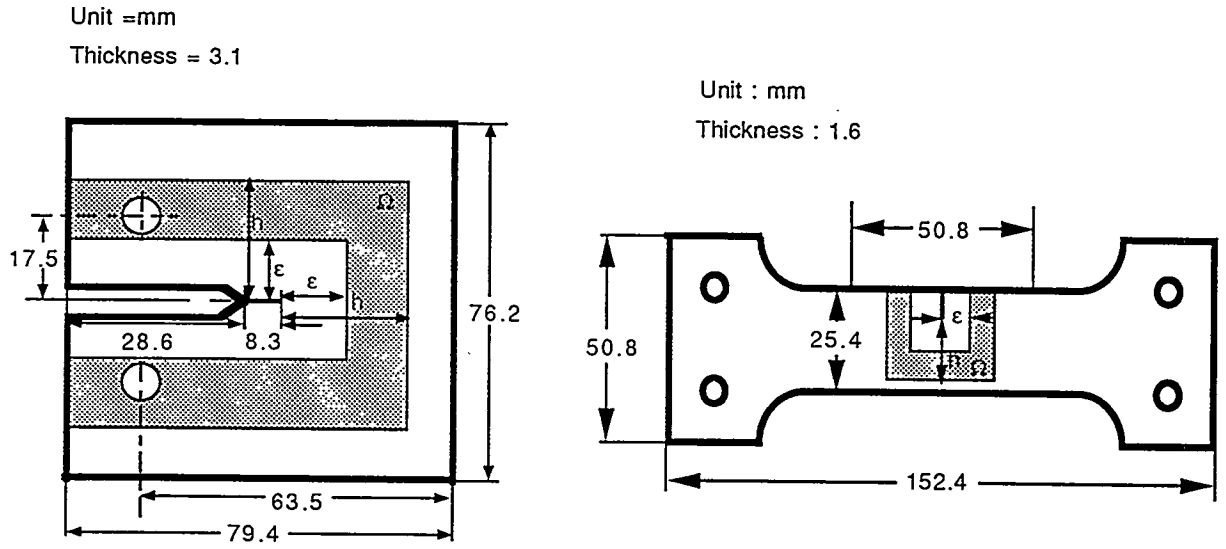
Experimental Procedure

The experimental procedure consisted of measuring the two orthogonal displacement fields surrounding a stably growing crack in the CT and SEN specimens using Moiré interferometry. Figure 1 shows the 2024-T3 aluminum CT specimen and the A606 HSLA steel SEN specimen, respectively. A coarse cross diffraction gratings of 40 lines/mm was used to determine the large plastic strains surrounding the extending crack. A special Moiré interferometry procedure [7], which combines the advantages of geometric Moiré and the traditional Moiré interferometry and uses a low frequency Moiré grating for measuring large strains in the vicinity of the crack tip, was developed for this ductile fracture study. A special four-beam Moiré interferometry bench was also constructed for use with this low frequency Moiré diffraction grating.

The strains along the line integration contour for the T^*_{ε} integral computation were determined directly from the orthogonal displacement field obtained by Moiré interferometry. The stresses corresponding to the total strains were then computed using the equivalent stress-strain and the measured uniaxial stress-strain data of the 2024-T3 aluminum and the A606 HSLA steel sheets. This use of the deformation theory of plasticity to compute stresses does not account for the unloading process which occurs in the trailing wake of the extending crack. However, by restricting the integration contour very close to and along the extending crack, Okada and Atluri

[8] has shown that the contour integration trailing the crack tip can be neglected by virtue of the closeness of the integration path, Γ_ε, to the traction free crack surface. This approximation not only simplified the integration process but also eliminated the undesirable effect of the deformation theory of plasticity which is used to compute the stresses from the measured strains.

The crack tip opening angle (CTOA) was also computed by the angle subtended by the measured crack opening displacement (COD) at a distance 1 mm from the crack tip.



(a) 2024-T3 CT specimen.

(b) A606 HSLA SEN Specimen.

Figure 1. Specimens.

Numerical Procedure

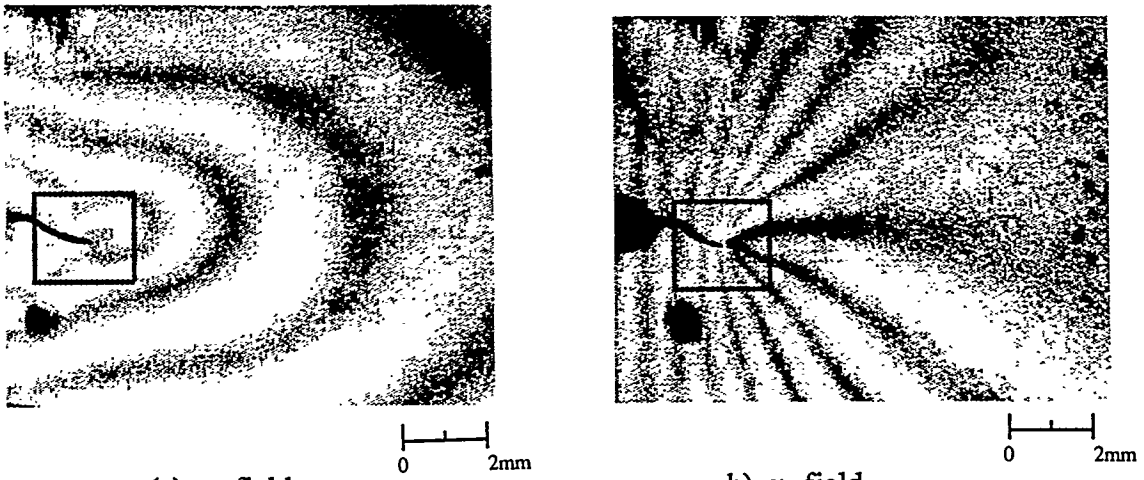
A plane stress finite element (FE) model of portion of the A606 HSLA steel SEN specimen was used for numerical analsysis. The FE model of the SEN specimen was truncated at a 20.8 mm distance from the crack in order to conserve compute time. The measured displacements along the truncated width of the SEN specimen together with the measured instantaneous crack length and the measured equivalent stress-strain relation of A606 HSLA steel were used to drive the FE model in its generation mode based on the incremental theory of plasticity.

The T^*_ε integrals along an elongated contour surrounding the stably growing crack of the SEN specimen was then computed. Unlike the stresses used in the experimental procedure for T^*_ε evaluation, the FE analysis provided stresses which accounted for the unloading effect in the trailing wake of the extending crack tip. Therefore, the entire contour was used for T^*_ε evaluation. Numerical errors in the FE data in the vicinity of the crack tip were masked by replacing the contour integral by the equivalent domain integral of Nikishkov and Atluri [7]. To recapitulate, T^*_ε evaluation procedures for the Moiré and FE studies differ in that the former

involved only the frontal segment of a near-field contour, Γ_ε, while the latter involved an equivalent domain integral over the entire crack length.

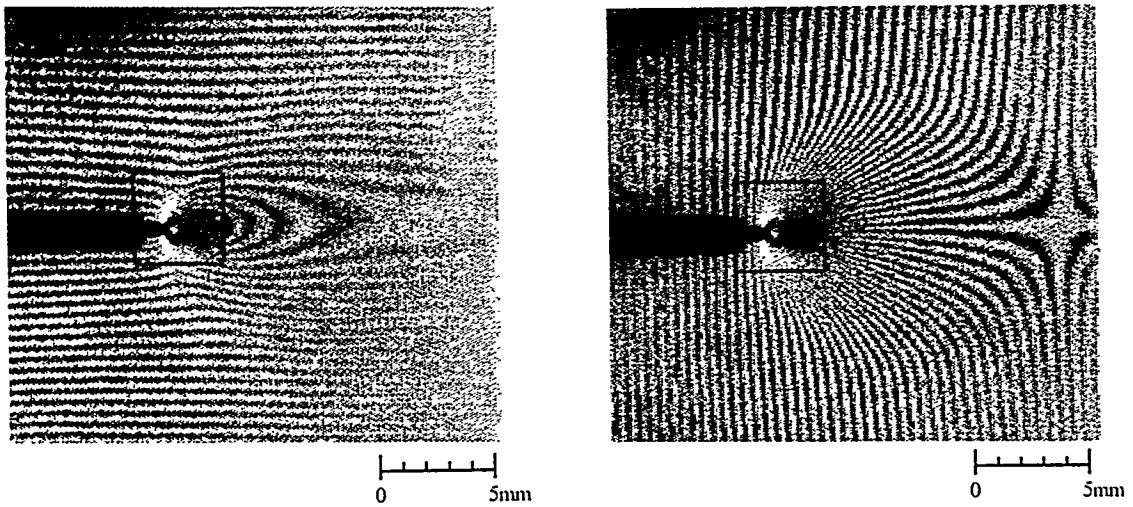
RESULTS

A total of four 2024-T3 aluminum CT specimens and four A606 HSLA SEN specimens were analyzed. Figures 2 and 3 show typical Moiré interferometry patterns of the u- and v-



(a) u- field. b) v- field.

Figure 2. Moiré fringe patterns of 2024-T-3 CT specimen.



(a) u-field. (b) v-field.

Figure 3 Moiré fringe patterns of A606 HSLA SEN specimen.

displacement fields of an 2024-T3 aluminum CT specimen and an A606 HSLA steel SEN specimen, respectively. Also shown are the elongated Γ_ε contour of $\varepsilon = 2.0$ mm along the extended crack in the 2024-T3 aluminum CT specimen and $\varepsilon = 2.0$ mm for the A606 HSLA steel SEN specimen. These ε's are over a one half of a plate thickness away from the crack tip where the state of plane stress is thought to prevail even in the presence of 100 percent shear lips in the crack.

Figure 4 shows the experimentally and FE determined T^*_ε for the 2024-T3 aluminum CT and the A606 HSLA SEN specimens. The maximum crack extension, Δa, for the 2024-T3 aluminum and A606 HSLA steel specimens were 5.5 and 8.0 mm, respectively. Also shown for comparison is the experimentally and numerically determined T^*_ε for the same Γ_ε in a thinner 2024-T3 aluminum SEN specimen [10].

Figure 5 shows the experimentally determined crack tip opening displacements (CTOA). The CTOA for the aluminum SEN specimen is in excellent agreements with that of the 2024-T3 aluminum SEN specimen [10] as well as that of Dawicke et al [11].

Unlike the increasing J resistance curve determined by the ASTM Standard testing procedure, or the vanishing J for the near field solution, the T^*_ε for a given Γ_ε shown in Figure 4 and the CTOA shown in Figure 5 both reached physically more realistic steady state values with stable crack growth.

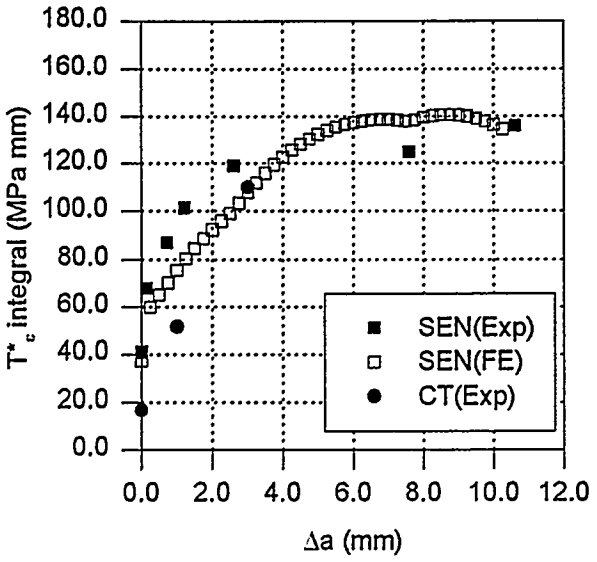
CONCLUSIONS

1. T^*_ε computed from the displacement fields, which were obtained experimentally and from finite element analysis, of the A606 HSLA steel SEN specimens were in good agreement with each other.

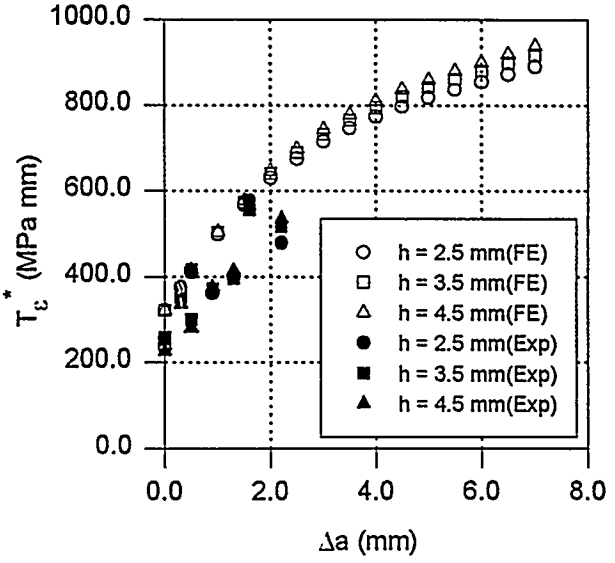
2. Computed T^*_ε values are domain-size independent and tend to converge to a stationary value for a smaller inner contour, Γ_ε of $\varepsilon = 2.0$ mm for the 2024-T3 aluminum CT specimen and $\varepsilon = 2.0$ mm for th A606 HSLA specimen.

3. T^*_ε did not reach a steady state value for $\Delta a = 8.0$ mm during stable crack growth in a thin A606 HSLA SEN specimen.

4. Computed and measured CTOA reached a steady state value of 5^o and 15^o, respectively, after a stable growth of about $\Delta a = 2.0$ mm in the 2024-T3 aluminum CT specimen and A606 HSLA steel SEN specimen.

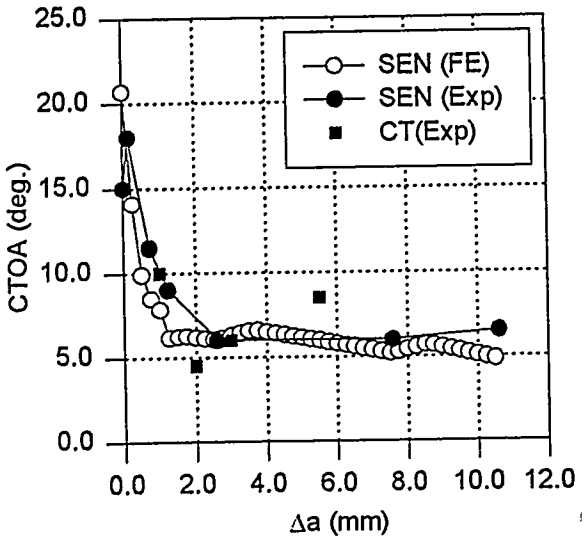


(a) 2024-T3 CT specimen

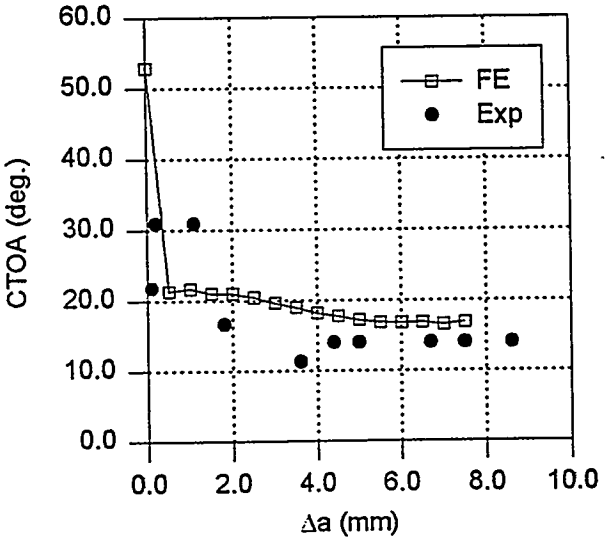


(b) A606 HSLA SEN specimen

Figure 4. Experimentally determined and FE T^*_{ε} integrals.



(a) 2024-T3 SEN and CT specimens



(b) A606 HSLA SEN specimen

Figure 5. CTOA.

ACKNOWLEDGMENT

The results reported here were obtained through DOE Grants DE-FG06-94ER14490 and DE-FG0694ER 14491.

REFERENCES

1. G.B. May. and A.S. Kobayashi, "Plane stress stable crack growth and *J*- Integral/HRR field," *Int. J. Solid and Structures,* 37/617, 857-881 (1995).

2. F.W. Brust, T. Nishioka, S.N. Atluri, and M. Nakagaki, "Further studies on elastic-plastic stable fracture utilizing the *T** integral," *Engineering Fracture Mechanics*, 22, 1079-1103 (1985).

3. F.W. Brust, J.J. McGowan, and S.N. Atluri,, "A combined numerical/experimental study of ductile crack growth after a large unloading, using *T**, *J* and CTOA criteria," *Engineering Fracture Mechanics*, 23, 537-550 (1986).

4. R.C. Stonesifer and S.N. Atluri, "On a study of the (ΔT) and C* integrals for fracture analysis under non-steady creep," *Engineering Fracture Mechanics,* 16, 769-782 (1982).

5. R. Narasimhan and A.J. Rosakis, "Three-Dimensional Effects Near a Crack Tip in a Ductile Three-Point Bend Specimen: Part I--A Numerical Investigation," *ASME Journal of Applied Mechanics,* 57, 607-617, 1990.

6. C.-R. Pyo, H. Okada, and S.N. Atluri, "An Elastic-plastic finite element alternating method for analyzing wide spread fatigue damage in aircraft structures," *Computational Mechanics*, 16, 62-68 (1995).

7. F.X. Wang, G.B. May and A.S. Kobayashi, "Low-spatial-frequency steep grating for use in Moiré interferometry," *Optical Engineering*, 33, 1125-1131 (1994).

8 H. Okada, and S.N. Atluri, "T^*_ε integral evaluation from experimental displacement field for a plate with stably propagating crack: development of calculation procedure and implication of T^*_ε", (1996), to be submitted.

9. G.P. Nikishkov and S.N. Atluri, "An equivalent domain integral method for computing crack-tip integral parameters in non-elastic, thermo-mechanical fracture," *Engineering Fracture Mechanics,* 26, 851-867 (1987).

10. Y. Omori, H. Okada, L. Ma, S.N. Atluri, A.S. Kobayashi, and P. Tan, "Further Studies on T^*_ε Integral Under Plane Stress Crack Growth," to be published in *Localized Damage, '96,* Wessex Institute of Technology.

11. D.S. Dawicke, M. Sutton, J.C. Newman, and C.A. Bigelow, "Measurement and analysis of critical CTOA for thin-sheet aluminum alloy materials," *Fracture Mechanics 25th Symposium,* ASTM STP 1220, eds. F. Erdogan and R. J. Hartranft, 358-379 (1995).

MICROSCOPIC INTERFACIAL PHENOMENA DURING FLOW IN POROUS MEDIA

Michael J. Miksis and Michael P. Ida

Department of Engineering Sciences
and Applied Mathematics
Northwestern University
Evanston, Illinois 60208

Abstract

A fundamental process during any multiphase flow in porous media is the breaking apart of one of the phases into smaller components. Here we investigate this breaking process as applied to a thin liquid film. We study the breaking of both a two dimensional planar film and a cylindrical thread of liquid using both analytical and numerical methods.

INTRODUCTION

Multiphase flows in porous media occur in many situations of practical interest. An example with application to enhanced oil recovery is foam flow in porous media. Here a gas and surfactant solution is injected into the ground. This then generates a foam which is used to drive the oil out. The foam is composed of alternating regions of liquid and gas with the liquid primarily in the form of a thin liquid film coating the solid portions of the pores or in the form of thin liquid lamellae which separate the gas regions and move with the gas down the pore channel. As the foam moves within the porous material, these liquid films can break or rupture forming larger gas bubbles within the material. In addition, the formation of the foam itself is associated with a film rupture process called snap-off where a large gas bubble is driven through a constriction in the pore channel, becomes unstable and splits into two parts (Tsai and Miksis [1].) Our aim here is to study this rupture process. We will consider a model for the dynamics of a thin viscous film which is valid for long wave (relative to the film thickness) disturbances. The model accounts for the effects of surface tension, inertia and van der Waals forces. We begin by studying a two dimensional planar film. Numerical and analytical results close to rupture will be presented. We will also study a thin three dimensional axisymmetric liquid thread.

There have been a number of works recently concerned with the breaking of either thin 2-D liquid films or axisymmetric threads of liquid. For example, Dupont et al. [2] analyzed the interfaces between two liquids under the influence of surface tension (but without van der Waals forces). Rupture in a finite time is observed numerically and the region near rupture is scrutinized for a similarity solution. Considering the finite-time rupture of thin fluid layers in

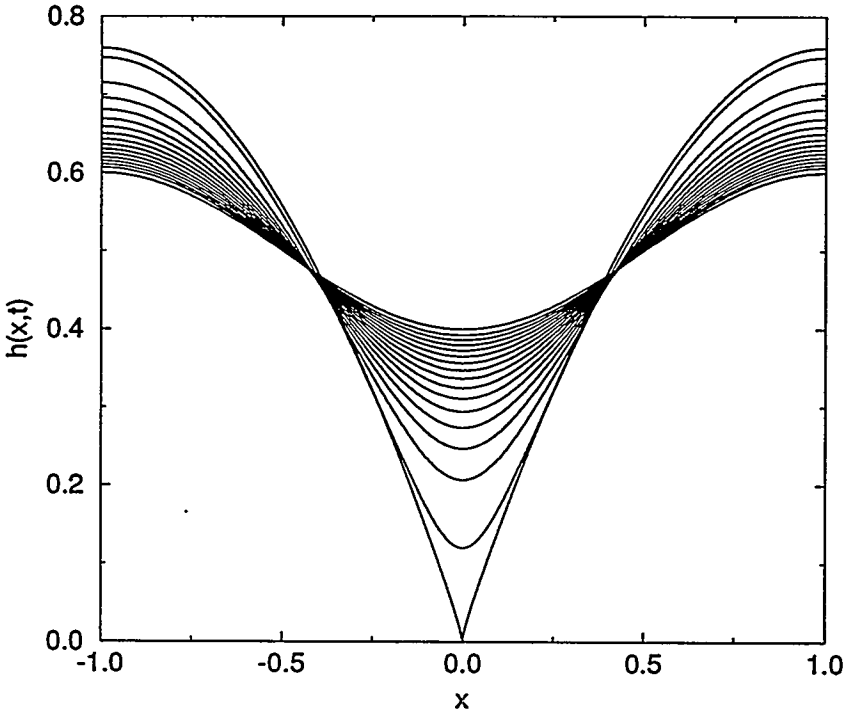


Figure 1: Time series evolution of an unstable thin film for $A = 1$ and $S = A/\pi^2 = S_c/2$. $h^{(0)}(x,t)$ is shown for $t = 0, 0.1, 0.2, \ldots, 1.4, 1.5, 1.52129982$.

Hele-Shaw flows, Goldstein et al. [3] studied a lubrication model for the dynamics of the fluid interfaces. Based upon their numerical studies, they concluded that the Bond number (van der Waals forces were not included in their model) plays a decisive role in the character of the solutions near rupture. Additionally, they investigated the region near rupture numerically and conjectured a power-law form for the interface ($h \propto (t - t_r)^\beta$ where β was numerically found to be equal to 1.45 ± 0.05). Claiming the existence of "universal exponents," Eggers [4] considered the rupture regions of thin axisymmetric threads of liquid. In the same context, i.e., the thinning of threads of liquid, Eggers and Dupont [5] analyzed the "pinching" singularity observed experimentally. Their model equations included the effects of surface tension with the full curvature term, but neglected the role of van der Waals forces.

TWO DIMENSIONAL THIN FILMS

We will use the two-dimensional free film equations of Erneux and Davis [6]. These equations were derived asymptotically in the long wave limit. For convenience, we summarize those equations here. Consider an incompressible film of density ρ and dynamic viscosity μ. Introduce a two-dimensional (x, z) cartesian coordinate system. Let b be the unit of length in the x direction and let h_0, the undisturbed film thickness, be the unit of length in the z direction. In addition let $\mu/b\rho$ be the unit of velocity and $\rho b^2/\mu$ the unit of time. Specify the symmetric film

interfaces by $z = \pm h(x,t)$, here t is time, and the transverse component of the fluid velocity in the x-direction by $u(x,t)$. In the limit $h_0/b \ll 1$, the leading-order evolution equations for these two quantities are then,

$$\frac{\partial h}{\partial t} = -u\frac{\partial h}{\partial x} - h\frac{\partial u}{\partial x}, \tag{1}$$

$$\frac{\partial u}{\partial t} = \frac{4}{h}\frac{\partial h}{\partial x}\frac{\partial u}{\partial x} + 4\frac{\partial^2 u}{\partial x^2} - u\frac{\partial u}{\partial x} + 3S\frac{\partial^3 h}{\partial x^3} + \frac{3A}{8h^4}\frac{\partial h}{\partial x}, \tag{2}$$

where $S = \gamma_0 \rho h_0/3\mu^2$ is a nondimensional constant which encompasses the dimensional surface tension γ_0, and $A = \rho b^2 \tilde{A}/6\pi h_0^3 \mu^2$ is a dimensionless constant which encompasses the effects of van der Waals forces through the Hamaker constant $\tilde{A}$. We will assume periodicity on the interval $x \in [-1,1]$ so that about the origin $x = 0$, h is even ($h(-x,t) = h(x,t)$) and u is odd ($u(-x,t) = -u(x,t)$). Fig. 1 demonstrates a typical unstable numerical solution to these equations. The parameter values are $A = 1$ and $S = A/\pi^2$. The initial conditions used are $u(x,0) = 0$ and $h(x,0) = 1/2 - (0.1)\cos\pi x$, and solutions are shown for $t = 0, 0.1, 0.2, \ldots, 1.4, 1.5, 1.52129982$. Note that near rupture ($h \to 0$ at $x = 0$), the numerical solution appears to have taken on a similarity form. Our aim is to identify this dynamics. Erneux and Davis [6] have show that a uniform thickness film is linearly stable when $S > 2A/\pi^2$. Hence we will select our parameters to fail this stability criterion in order to observe rupture.

In order to resolve the form of the solution near the point of breakage, we define the following similarity variables,

$$\tau = t_r - t, \tag{3a}$$

$$\xi = x\tau^{-\beta}, \tag{3b}$$

where t_r is the time of rupture. We then assume the following forms for h and u,

$$h = \tau^{\alpha}\mathcal{H}(\xi), \tag{4a}$$

$$u = \tau^{\Gamma}\mathcal{U}(\xi). \tag{4b}$$

The forms of these solutions are suggested by the rapid variations in both space and time near the point of rupture. Now substitute the similarity forms into the evolution equations (1) and (2). Suppose we balance all the terms in (1). We find that $\Gamma = \beta - 1$. Now consider the other equation. Here there are several possible terms which could be balanced, each giving a different answer. We claim that the viscous and van der Waals terms must balance (this will be verified numerically later). Hence this assumption implies that we need to balance the terms of $O(\tau^{\Gamma-2\beta})$ with those of $O(\tau^{-3\alpha-\beta})$. This implies that $\alpha = 1/3$. Hence we have determined that $\Gamma = \beta - 1$ and $\alpha = 1/3$ but we have not determined the specific values of β and Γ from this dominate balance argument.

Now we will examine the evolution of the interface numerically. To obtain extreme spatial accuracy in the region near rupture while maintaining a reasonable execution time, an adaptive regridding finite-difference algorithm based upon that employed by Bernoff and Bertozzi [7], Bertozzi [8] and Dupont et al. [2] has been used to solve equations (1) and (2) (see Ida [9] for details). The results of a systematic examination of the observed exponents are summarized in table 1. In each case, the same initial conditions $((h(x,0), u(x,0)) =$

exponents	$A=1$	$A=5$	$A=10$
α	$0.3438428 \pm 0.002986701$	0.3381307 ± 0.00352386	$0.3365617 \pm 0.003716294$
$\alpha - 2\beta$	$-0.4422666 \pm 0.006927182$	$-0.5384250 \pm 0.012368135$	$-0.58326818 \pm 0.01416819$
$\Gamma - \beta$	$-1.0124770 \pm 0.004877930$	$-1.0181044 \pm 0.004328344$	$-1.0192778 \pm 0.004913314$

Table 1: Results of a systematic study of the exponents in the similarity solution for the free film near breaking. For each value of A, values are averaged over $S = \{0, A/4\pi^2, A/2\pi^2, 3A/4\pi^2, A/\pi^2\}$.

$(1/2 - (0.1)\cos\pi x, (0.1)\sin\pi x))$ are used; and for each value of $A = \{1, 5, 10\}$, values of $S = \{0, A/4\pi^2, A/2\pi^2, 3A/4\pi^2, A/\pi^2\}$ are employed. For each value of A used, the values of the exponents are then averaged, with the results and their standard deviations presented in table 1. From the values of the standard deviations, which never exceed approximately 2%, we see that the values of all exponents appear to be independent of S. We find, however, that while α and $\Gamma - \beta$ appear to be independent of A (each is within an interval of approximately two standard deviations in width, and within approximately four standard deviations of their expected values of 1/3 and -1), $\alpha - 2\beta$, the exponent for $\frac{\partial^2 h}{\partial x^2}(0,t)$ which would result from (4), is not (the interval needed to contain it is on the order of 20 standard deviations). Thus, our assumption concerning the balancing of the terms seems to be confirmed.

We would now seem to have satisfactory proof as to the correct form of the similarity solution. Thus, we find that $\alpha = 1/3$, and that $\Gamma - \beta = -1$ but Γ and β are undetermined. We speculate that the constants β and Γ are to be determined through a matching procedure to a more complex solution of the full evolution equations in the outer region away from the point of rupture.

In conclusion we note that we have identified a similarity form for the film thickness and transverse velocity of a two-dimensional thin free film near rupture. We have found that van der Waals and viscous forces dominate the evolution, still leaving a single degree of freedom (the value of either β or Γ) in the precise specification of the form of the solution. This undetermined constant could in principle be determined through a matching procedure to the outer flow field away from the immediate vicinity of rupture. Similarity solutions of this type have been discussed by Barenblatt and Zel'dovich [10].

THREE DIMENSIONAL LIQUID THREADS

In this section, we investigate a set of evolution equations for the dynamics of a thin, axisymmetric thread of liquid [9]. The equations can be derived by a long wave analysis [9]. The effects of a van der Waals like force will be included in our model, in addition to the effects of surface tension, inertia and viscosity.

We begin by considering an infinite, periodic, axisymmetric thread of liquid of density ρ and kinematic viscosity ν oriented along the $\tilde{z}$ axis of a cylindrical coordinate system. Denote the fluid velocity by $\tilde{\mathbf{v}}(\tilde{r}, \tilde{z}) = \tilde{u}_r(\tilde{r}, \tilde{z})\,\mathbf{r} + \tilde{u}_z(\tilde{r}, \tilde{z})\,\mathbf{z}$ where $\mathbf{r}$ and $\mathbf{z}$ are the unit vectors in the $\tilde{r}$ and $\tilde{z}$ directions respectively. Let $\tilde{r} = \tilde{h}(\tilde{z}, \tilde{t})$ represent the fluid interface, $\mathbf{n}$ and γ the surface tension on it. In order to extract a tractable problem we perform a long wavelength analysis of the the equations of motion. Let $\tilde{b}$ be a typical wavelength of the system and $\tilde{h}_0$ be a typical

thread radius such that $\epsilon = \tilde{h}_0/\tilde{b} \ll 1$. The leading order in ϵ equations of motion are [9]

$$\frac{\partial h}{\partial t} = -u_z \frac{\partial h}{\partial z} - \frac{h}{2}\frac{\partial u_z}{\partial z}. \tag{5}$$

and

$$\begin{aligned}\frac{\partial u_z}{\partial t} = &\frac{6}{h}\frac{\partial h}{\partial z}\frac{\partial u_z}{\partial z} + 3\frac{\partial^2 u_z}{\partial z^2} \\ &-u_z\frac{\partial u_z}{\partial z} + \frac{1}{h^2}\frac{\partial h}{\partial z}\left(S + \frac{3A}{h^2}\right).\end{aligned} \tag{6}$$

Here we have introduced the dimensionless variables $\tilde{u}_r = (\epsilon\nu/\tilde{b})u_r$, $\tilde{u}_z = (\nu/\tilde{b})u_z$, $\tilde{z} = \tilde{b}z$, $\tilde{r} = \tilde{h}_0 r$ and $\tilde{h} = \tilde{h}_0 h$. In addition we have scaled time by $\tilde{b}^2/\nu$ and the pressure difference by $(\rho\nu^2/\tilde{b}^2)$. Similar to the previous section we introduce the dimensionless constants $S = \gamma\tilde{b}/\epsilon\rho\nu^2$ and $\tilde{A} = (48\rho\nu^2\tilde{h}_0^3/\tilde{b}^2)A$, where we have assumed that there is an attractive van der Waals like potential of the form $\tilde{A}/\tilde{h}^3$. Together, equations (5) and (6) constitute a set of evolution equations for the two quantities h and u_z. These equations are identical to the evolution equations obtained by Eggers and Dupont [5] with gravity replaced by van der Waals attractions.

Note that the pressure scale specified here is slightly different from that used in the analysis of thin films. In particular, if p_{film} is the pressure scale used in the thin film scalings of the previous section and p_{thread} is the pressure scale used here, then we have that $p_{film}/p_{thread} = \epsilon$. This larger pressure scale is necessary to balance the radial curvature which is now large because of the thinness of the thread and the hoop stress it induces in the normal stress boundary condition.

Suppose that we use the full curvature in the equations, not the asymptotic correction as done above. Then we find that equation (6) becomes,

$$\begin{aligned}\frac{\partial u_z}{\partial t} = &\frac{6}{h}\frac{\partial h}{\partial z}\frac{\partial u_z}{\partial z} + 3\frac{\partial^2 u_z}{\partial z^2} - u_z\frac{\partial u_z}{\partial z} + \frac{3A}{h^4}\frac{\partial h}{\partial z} \\ &+S\left\{\frac{\dfrac{\partial h}{\partial z}}{\sqrt{1+\left(\dfrac{\partial h}{\partial z}\right)^2}}\left[\frac{1}{h^2} + \frac{\dfrac{\partial^2 h}{\partial z^2}}{h\left[1+\left(\dfrac{\partial h}{\partial z}\right)^2\right]} - \frac{3\left(\dfrac{\partial^2 h}{\partial z^2}\right)^2}{\left[1+\left(\dfrac{\partial h}{\partial z}\right)^2\right]^2}\right]\right. \\ &\left. + \frac{\dfrac{\partial^3 h}{\partial z^3}}{\left[1+\left(\dfrac{\partial h}{\partial z}\right)^2\right]^{3/2}}\right\}.\end{aligned} \tag{7}$$

The form of equation (7) is now quite similar to the transverse velocity evolution equations previously derived for thin films (equations (1) and (2)). In particular, the highest-order derivative multiplying the surface tension parameter, S in (7) is now a third-order spatial derivative as it is in the thin film evolution equations. This is in marked contrast to the evolution equation (6) derived using the correct asymptotic expression for the curvature where the the highest-order derivative multiplying S is only of first-order.

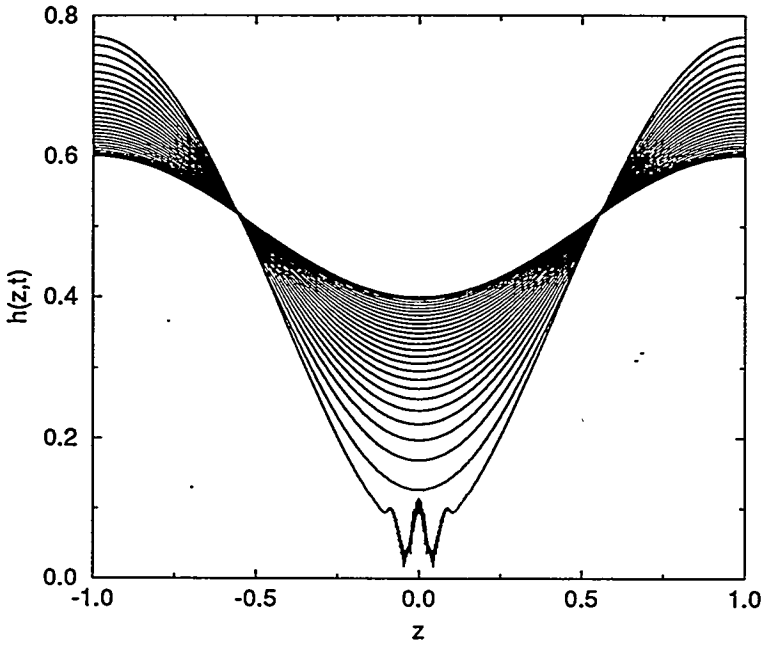


Figure 2: Evolution of an unstable perturbation to the thin thread equations solved using the finite-difference scheme. The parameter values used are $S = A = 1$; and solutions are shown for $t = 0, 0.01, \ldots, 0.25, 0.2572390625$.

To consider the linear stability of the evolution equations (5) and (6) we perturb in normal modes about the steady-state solutions $h = \hat{h} = const.$ and $u_z = 0$,

$$\begin{pmatrix} h \\ u_z \end{pmatrix} = \begin{pmatrix} \hat{h} \\ 0 \end{pmatrix} + \begin{pmatrix} \check{h} \\ \check{u}_z \end{pmatrix} e^{\omega t + ikz}. \tag{8}$$

Linearizing and applying a solvability condition, we obtain the dispersion relation for the growth rate, ω,

$$\omega = \frac{3k^2}{2}\left[-1 \pm \sqrt{1 + \frac{2}{9k^2\hat{h}}\left(S + \frac{3A}{\hat{h}^2}\right)}\right]. \tag{9}$$

Since the larger value of $\omega > 0$ for all values of the parameters involved, we find that the steady-state solution is unconditionally unstable.

If we now use the full expression for the curvature and we analyze the linear stability of equations (5) and (7), we find that the dispersion relation is instead given by,

$$\omega = \frac{3k^2}{2}\left\{-1 \pm \sqrt{1 + \frac{2}{9k^2\hat{h}}\left[S\left(1 - k^2\hat{h}^2\right) + \frac{3A}{\hat{h}^2}\right]}\right\}. \tag{10}$$

Since there are now values of the parameters for which both values of ω may be negative and the steady-state stable, we find that the condition for stability is $k > k_c$ where,

$$k_c = \frac{1}{\hat{h}}\sqrt{1 + \frac{3A}{S\hat{h}^2}}. \tag{11}$$

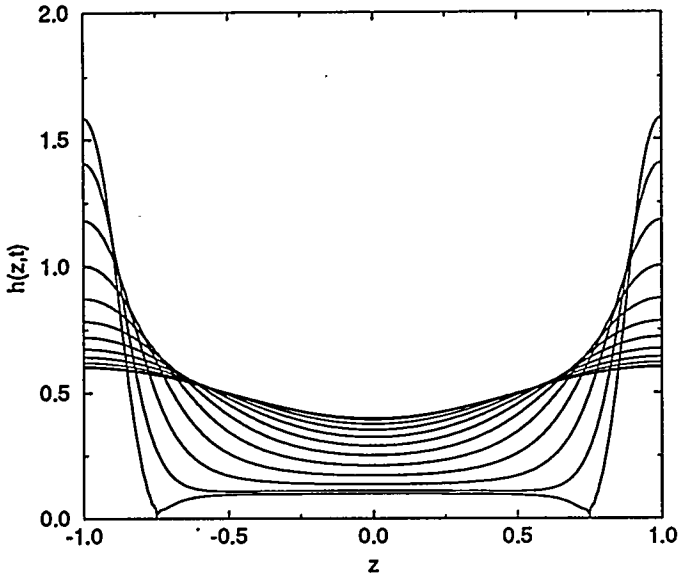


Figure 3: Evolution of an unstable perturbation to the thin thread equations solved using the finite-difference scheme. The parameter values used are $S = 100$ and $A = 1$; and solutions are shown for $t = 0, 0.01, \ldots, 0.09, 0.10, 0.10756$.

In this instance, the largest eigenvalue is positive only over a finite range of k.

Now consider the numerical solution of the evolution equations discussed above assuming periodicity on the interval $z \in [-1, 1]$. The finite-difference approximation used earlier for the two-dimensional thin film (see Ida [9] for details) is used here.

Figs. 2 and 3 are representative of the types of solutions of the coupled nonlinear system (5) and (6). Each figure is started with the same initial conditions, $(h(z,0), u_z(z,0)) = (1/2 - (0.1)\cos\pi z, 0)$, and with the same value of $A = 1$. In fig. 2 we set $S = 1$ and show solutions for $t = 0, 0.01, \ldots, 0.25, 0.2572390625$ and in fig. 3 for which $S = 100$, solutions are shown for $t = 0, 0.01, \ldots, 0.10, 0.10756$. What we find, by resolving the region near rupture more closely, is that a pair of "dimples" forms off of the axis of symmetry ($z = 0$) and leaves an isolated thin packet of fluid remaining. We see that as S increases, the size of this packet of fluid increases with the rupture points moving further away from the axis of symmetry. If we were to focus our attention on the region near rupture, we see the formation of the "dimpled" structure before the onset of numerical instabilities. (We found these instabilities extremely difficult to eliminate, because the growth-rate in equation (9) is monotonically increasing $\forall k > 0$, and the film is unconditionally unstable. Thus, the onset of instabilities exhibited in shorter wavelength modes is not limited by a cutoff wavelength.)

If we solve the equations utilizing the full curvature term, (5) and (7), we obtain solutions where the rupture point is more localized and the dimple seems to disappear. Only order one values of S have been considered and larger values need to be studied to confirm this conclusion.

In conclusion we note that in this section we have studied a set of leading-order evolution equations for the interface shape and transverse velocity of an axisymmetric thin liquid thread

subject to a van der Waals like force. Based upon a linear stability analysis, we have found these equations to be unconditionally unstable. These equations have been solved numerically. We find that as the effects of surface tension are increased, rupture occurs over a wider spatial domain, and a "dimpled" structure is exhibited. This is in marked contrast to the characteristic rupture behavior of thin films which is confined to a small spatial region.

ACKNOWLEDGEMENT

This research was supported in part by Department of Energy grant DE-FG02-88ER13927.

References

[1] T.-M. Tsai and M.J. Miksis, Dynamics of a drop in a constricted capillary tube. *J. Fluid Mech.*, 274, 197-217 (1994).

[2] T. F. Dupont, R. E. Goldstein, L. P. Kadanoff, and S.-M. Zhou, Finite-time singularity formation in Hele-Shaw systems. *Phys. Rev. E*, 47, 4182–96 (1993).

[3] R. E. Goldstein, A. I. Pesci, and M. J. Shelley, Topology transitions and singularities in viscous flows. *Phys. Rev. Lett.*, 70, 2043–2046 (1993).

[4] J. Eggers, Universal pinching of 3d axisymmetric free surface flow. *Phys. Rev. Lett.*, 71, 3458–3460 (1993).

[5] J. Eggers and T. F. Dupont, Drop formation in a one-dimensional approximation of the Navier-Stokes equation. *J. Fluid Mech.*, 262, 205–21 (1994).

[6] T. Erneux and S. H. Davis, Nonlinear rupture of free films. *Phys. Fluids A*, 5, 1117–22 (1993).

[7] A. J. Bernoff and A. L. Bertozzi, Singularities in a modified Kuramoto-Sivashinski equation describing interface motion for phase transition. *Physica D*, 85, 375-404 (1995).

[8] A. L. Bertozzi, Symmetric singularity formation in lubrication-type equations for interface motion. preprint.

[9] M. P. Ida, The Dynamics of Thin Liquid Films, Ph.D. thesis, Northwestern University, Dec. 1995.

[10] G. I. Barenblatt and Ya. B. Zel'dovich, Self-similar solutions as intermediate asymptotics. *Ann. Rev. Fluid Mech.*, 4, 285-312 (1972).

DETERMINATION OF NEAR-SURFACE MATERIAL PROPERTIES BY LINE-FOCUS ACOUSTIC MICROSCOPY

Jan D. Achenbach and Wei Li

Center for Quality Engineering and Failure Prevention
Northwestern University
Evanston, IL 60208, U.S.A.

ABSTRACT

A line-focus acoustic microscope is used in conjunction with a multiple wave-mode method to determine elastic constants from a single $V(z)$ measurement. $V(z)$ curves which include contributions from different wave modes, measured using the line-focus acoustic microscope at 225 MHz, have been compared with theoretical results predicted by a $V(z)$ measurement model. The determination of elastic constants has been achieved numerically by seeking a set of elastic constants that leads to the best fit, in the least square sense, of the theoretical results to the experimental ones. The method has been applied to isotropic materials in bulk, and plate and thin-film configurations. Elastic constants for each of these cases have been determined. The consistency, convergence, sensitivity and accuracy of the procedure have been investigated.

INTRODUCTION

Line-focus acoustic microscopy (LFAM) provides a method to determine the elastic constants of homogeneous specimens and thin-film/substrate configurations. The elastic constants are determined from the velocities of leaky acoustic waves that can be obtained from $V(z)$ measurements. The $V(z)$ curve, which is a record of the transducer output voltage V as a function of the distance z between the lens focus and the specimen surface, is unique to a material and referred to as the material signature. Generally speaking, more than one elastic constant has to be determined, and hence more than one data point is required. Hence, for isotropic materials sufficient data cannot be procured with a single mode. For anisotropic solids the velocity can be measured as a function of the angle defining the propagation direction on the surface to yield a sufficiently large set of data. The technique has been discussed in great detail in a recent review article, which also lists numerous references [1]. For thin-film/substrate configurations measurements at various frequencies or for different film thicknesses may be carried out to obtain sufficient data using standard measurement procedures [1]. There are, however, obvious advantages to work with a single specimen and at a single frequency. As discussed in this paper, this can be done by considering the contributions of more than one leaky wave mode to the $V(z)$ curve.

This paper presents a multiple mode method to determine, with a single-frequency $V(z)$ measurement on a single specimen, the elastic constants of bulk isotropic solids, thin isotropic plates and thin-films on substrates. The $V(z)$ curve which includes contributions from multiple leaky acoustic waves has been measured experimentally using a line-focus acoustic microscope at 225 MHz. $V(z)$ curves have also been simulated numerically using a measurement model with selected elastic constants. Both the experimental and the numer-

ical $V(z)$ curves go through the same $V(z)$ analysis to yield their respective predictions for the leaky acoustic wave velocities. The determination of elastic constants is then achieved through minimization of the differences between the theoretical predictions and the experimental results by a numerical iterative searching procedure known as the simplex method. The accuracy, consistency, convergence and sensitivity of the numerical inversion have also been studied in this paper.

THIN FILMS

Thin film materials are already widely used, and they promise to have significant additional applications in future technology. There are many present and potential applications of configurations consisting of thin films deposited on a substrate.

A first important application is concerned with thin coatings to protect surfaces of components from wear, impact, corrosion, and thermal disturbances. Hard and wear-resistant coatings are an important segment in the US and world economy. The need to extend the wear-life using coatings is not only to save cost, but also reduction of downtime.

Diamond film has a number of remarkable properties. It is the hardest substance known, and it has a higher modulus of elasticity than any other material. When free of impurities, it has one of the highest resistivities. It also combines a very high thermal conductivity with a low thermal expansion coefficient to yield high resistance to thermal shock. Lastly, diamond is very resistant to chemical attack. However, diamond coatings are not good with ferrous alloys because diamond reacts with steel at high temperatures.

Transition-metal nitride films are commonly used as hard, protective coatings for softer surfaces. Superlattice films, including TiN/NbN, TiN/VN, and TiN/VNbN, have, however, been shown to exhibit much higher hardness than homogeneous nitride films. The elastic constants of these films on a substrate are difficult to measure.

LINE-FOCUS ACOUSTIC MICROSCOPY

An acoustic microscope consists of four main components: the acoustic probe, the pulse-mode measurement system for transmitting and receiving electrical signals, the mechanical systems for alignments and movements of the sample and a computer for controlling the system and processing the recorded wave forms.

The propagation of surface acoustic waves along the interface of a solid material and air or water provides a useful means to determine material constants of the solid by measuring propagation velocities. For small specimens, or when local values of the material constants must be determined, the measurements require small wavelengths corresponding to high frequencies. For such cases a single water coupled focused transducer can be used advantageously, and the ultrasonic technique is referred to as quantitative acoustic microscopy.

Particularly useful is a water-coupled line-focus acoustic lens since such a lens allows the measurement of the SAW velocity in specified directions. Hence line-focus acoustic microscopy (LFAM) has been used to determine the elastic constants of anisotropic materials.

The best known technique measures the $V(z)$ curve, which is a record of the magnitude of the transducer's voltage output V as a function of the distance z between the focal line of the lens and the specimen surface. It can be shown that the spacing of peaks or valleys in the oscillatory $V(z)$ curve is directly related to the propagation velocity of the surface wave, see [1]-[2].

$V(z)$ *Curve Measurement*

The experimental results have been obtained with a Honda AMS-5000 ultrasonic measurement system and a line-focus acoustic lens operating at $225MHz$. The detailed description of this system can be found in Reference [1]. Figure 1 shows schematically the configuration of the acoustic probe and the specimen. A ZnO-film transducer generates and detects longitudinal waves at the flat surface of a Z-cut sapphire rod. The acoustic beam is focused by an acoustic lens with a cylindrical concave surface at the other end of the rod. The cylindrical concave surface has a radius of $1.0mm$ and an aperture half-angle of 60°. The operating frequency is around $225MHz$ and the focal length of the lens is $1.15mm$. For efficient transmission of acoustic waves through the lens couplant interface, a chalcogenide glass film with a quarter wavelength thickness is deposited on the cylindrical concave surface. A specimen is placed on a mechanical stage, translated in the vertical direction and rotated around the axis of the rod.

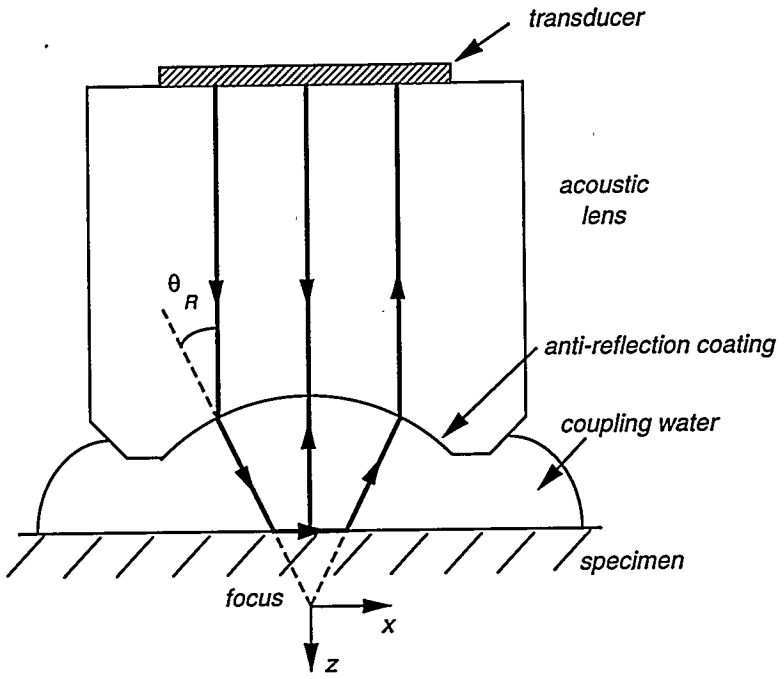


Figure 1: Cross-sectional configuration of the line-focus acoustic probe.

$V(z)$ *Measurement Model*

Basic to the interpretation of $V(z)$ curves is a reliable measurement model. A $V(z)$ measurement model simulates the measurement procedure, including any systemic errors that may occur in the determination of the velocity from experimental $V(z)$ curves. For example, effects due to multiple modes, will be replicated in the numerical model. The material constants obtained from comparisons of results from the measurement model and experiments will, therefore, be free of these systemic errors.

A measurement model for the $V(z)$ curve has been described in detail in Ref. [1]. According to the model, the output voltage of the transducer can be expressed as

$$V(z) = \int_{-\infty}^{+\infty} \exp(2ik_z z) L_1(k_x) L_2(k_x) R(k_x) dk_x \tag{1}$$

where $k_z = (k_w^2 - k_x^2)^{1/2}$ and k_w is the wave number in the coupling fluid. $L_1(k_x)$ is the angular spectrum of the wave field at the focal plane generated by a plane wave in the buffer rod, and $L_2(k_x)$ is the voltage response of the transducer when a plane wave of unit amplitude and wave vector (k_x, k_z) is insonifying the lens. A detailed description and discussion of the angular spectrum function $L_1(k_x)$ and the response function $L_2(k_x)$ can be found in Ref. [1], where integral expressions as well as numerical calculations have been presented. $R(k_x)$, which is the reflectance function of the fluid-loaded specimen, can generally be expressed as:

$$R(k_x) = \frac{A - \beta B}{A + \beta B} \tag{2}$$

where $\beta = i\rho_w \omega^2 / k_z$, and ρ_w is the coupling fluid density. A and B have different definitions for different sample geometries. For a fluid-loaded bare-substrate-specimen or a fluid-loaded layered-specimen, A and B can be found in Ref. [1]. For a fluid-loaded plate-specimen, a numerical approach proposed by Chimenti and Nayfeh [3] has been adopted here. A summary of the calculation of $R(k_x)$ for an isotropic plate is given in the Appendix of Ref. [4].

To determine the leaky surface wave velocities a procedure completely analogous to the one for the experimental $V(z)$ curve is applied to the theoretical result given by Eq. (1).

Determination of Elastic Constants

The elastic constants are obtained by seeking a set of elastic constants that yields the best fit in the least square sense of the theoretical predictions to the experimental results. A numerical iterative searching procedure known as the simplex method has been used to find the set of elastic constants that minimizes the deviation between theoretical and experimental results.

A dimensionless deviation function, D, is defined in terms of the measured and the calculated quantities as:

$$D = \sum_{n=1}^{N} \left\{ W_n^{(v)} \left| \frac{v_n^c - v_n^m}{v_n^m} \right| + W_n^{(A)} \left| \frac{A_n^c - A_n^m}{A_n^m} \right| \right\} \tag{3}$$

where v_n^m and A_n^m are the velocity and the peak-amplitude in the Fourier domain of the n-th leaky surface mode obtained from the measured $V(z)$ curve; v_n^c and A_n^c are the same quantities obtained from the calculated $V(z)$ curve. It is noted that different weights, $W_n^{(v)}$ and $W_n^{(A)}$, have been imposed on the differences in velocities and spectrum amplitudes in order to get the best results. Both experimental and numerical errors have been taken into consideration in choosing these weights. For example, more weight has been put on the velocity part than on the amplitude part, because the accuracy of the velocity measurements is usually about an-order-of-magnitude better.

The selection of the starting values for the application of the simplex method is in general a matter of trial and error. Different selections of the starting values should yield the same answer, but due to the existence of local minimums, a good initial guess of the starting values is very important for quick convergence to the true values.

SOME RESULTS

Specimens with three different sample configurations have been tested. Bulk specimens include a modified borosilicate glass AF45, a glass slide and an aluminum sample; the plate specimen is a modified borosilicate thin glass plate D263 of 50 μm thickness; and thin-film

specimens include a 2.2 μm titanium-film deposited on an aluminum substrate and a 0.1 μm gold-film deposited on a glass-slide substrate.

For the bulk specimens, the leaky Rayleigh wave and the LSSCW have been used for the determination of the elastic constants. For the D263 glass plate the zeroth-order Lamb-mode and a higher-order Lamb-mode have been used. For the thin-film/substrate specimens, multiple LSAW modes have been used to determine the thin-film elastic constants. The dispersive leaky Rayleigh wave and the LSSCW have been used for the titanium-film on an aluminum substrate configuration, and two generalized leaky Lamb modes, namely the dispersive leaky Rayleigh wave and the leaky Sezawa wave mode, have been used for the gold-film on a glass-slide substrate configuration.

The determined elastic constants are listed in Table 1. Known values of the densities and thicknesses have been used. For the thin-film/substrate cases, the elastic constants of the substrates were known. It should be noted that accurate information of the substrate properties is very important for the determination of thin-film constants. Elastic constants determined in this paper for the bulk aluminum and the glass-slide have been used for the substrates.

Table 1: Elastic constants determined for materials in various configurations.

Specimen	Thickness (μm)	Density (kg/m^3)	Waves or Modes used	Determined Elastic Constants (GPa)	
				E	G
AF45 Glass (bulk)	N/A	2720	leaky Rayleigh & LSSCW	63.2	25.6
Glass Slide (bulk)	N/A	2459	leaky Rayleigh & LSSCW	71.0	28.5
Aluminum (bulk)	N/A	2700	leaky Rayleigh & LSSCW	70.4	26.3
D263 Glass (plate)	50	2510	0-th Lamb & higher Lamb	73.1	30.6
Titanium film (on Aluminum)	2.2	4508	dispersive leaky Rayleigh & LSSCW	87.7	33.2
Gold film (on Glass slide)	0.1	19281	dispersive leaky Rayleigh & leaky Sezawa	79.5	28.0

DISCUSSION

Comparisons between Theoretical Predictions & Experimental Results

The sets of elastic constants that have been determined can be verified by comparing the corresponding theoretical velocities with the experimental results. Figure 2(a)(b) shows $V(z)$ curve and $V(k)$ curve comparisons, where the dashed lines are the experimental results

and the solid lines are the theoretical predictions using the determined elastic constants.

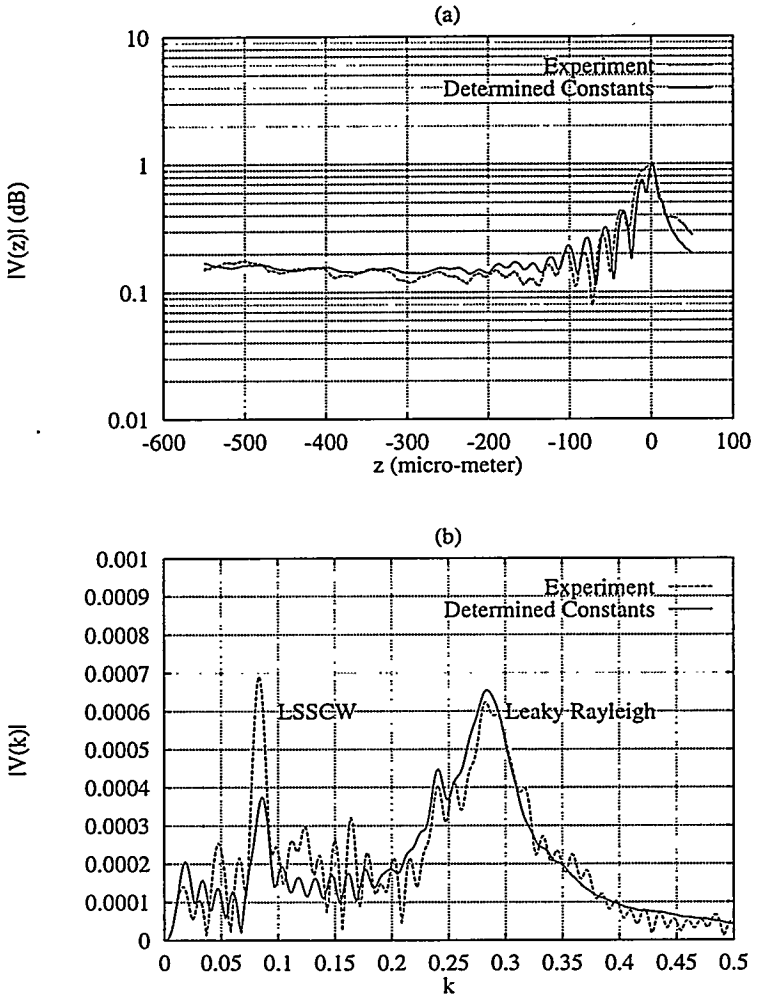


Figure 2: (a) $V(z)$ curve comparisons for AF45 bulk glass. (b) $V(k)$ curve comparisons for AF45 bulk glass.

Consistency of Determined Elastic Constants

The elastic constants that have been determined from one set of experimental data can be checked by using them to calculate $V(z)$ curves for another independent experimental setup. For example, the elastic constants of D263 glass plate that have been determined from the water/plate/air configuration can be verified using the water/plate/water configuration.

Convergence due to Use of Multiple Leaky Acoustic Waves

As discussed in section 2.4, the numerical procedure includes information from different leaky acoustic waves. The use of multiple leaky wave information reduces the region in the plane of the material constants E and G for which the deviation D, defined by Eqn. (3), has

a prescribed small value. This improves the convergence of the procedure. For the AF45 bulk glass sample, Figure 3 shows the reduction of the region for $D < 0.024$ due to the use of multiple leaky acoustic waves. Figure 3 displays the changes of the deviation field D as the Young's modulus and the shear modulus vary around the determined values for the five cases: (a) if only leaky Rayleigh wave velocity information is used for the minimization of D; (b) if only LSSCW velocity information is used; (c) if only leaky Rayleigh wave amplitude information is used; (d) if only LSSCW amplitude information is used; (e) if all the above (a)-(d) have been used and combined in Eqn. (3).

It is interesting to note, from Figures 3(a)-(d), that despite the obvious convergence problems implied by the extent of the dark areas, the areas have different slopes. It is the different directions of the slopes in Figure 3(a) and Figure 3(b)(c)(d) that lead to the convergence of our numerical inversion when Eqn. (3) is used, as indicated by the finite extent of the dark domain in Figure 3(e).

CONCLUSIONS

$V(z)$ measurements of multiple leaky acoustic waves by line focus acoustic microscopy have been used to determine elastic constants. The method presented in this paper has the advantage that it requires only a single V(z)-curve measurement by the LFAM. It serves as a very effective approach to determine elastic constants of a single isotropic specimen since only one single-frequency $V(z)$ measurement is required. Elastic constants for isotropic materials in bulk, and for plate and thin-film/substrate configurations have been determined to satisfactory accuracy. It has been shown that the use of multiple leaky wave information reduces the region in the plane of the material constants E and G for which the deviation D has a prescribed small value. This improves the convergence of the method. Consistency of the results has been shown by verifying the elastic constants determined from one set of experimental data with results from another experimental configuration. The sensitivity of the deviation D to variations of the elastic constants has also been investigated. Generally D is more sensitive to changes of Young's modulus.

ACKNOWLEDGMENT

This work was supported by the Department of Energy, Division of Engineering and Geosciences, Office of Basic Energy Sciences, under Grant DE-FG02-86ER13484.

REFERENCES

1. J. D. Achenbach, J. O. Kim and Y. C. Lee, "Measuring Thin-Film Elastic Constants by Line-Focus Acoustic Microscopy," in *Advances in Acoustic Microscopy*, Vol. 1, edited by Andrew Briggs, 153-208, Plenum Press, New York (1995).

2. J. Kushibiki and N. Chubachi, "Material Characterization by Line-Focus Beam Acoustic Microscope," *IEEE Trans. Sonics Ultrason.*, SU-32, 189-212 (1985).

3. D. E. Chimenti and A. H. Nayfeh, "Ultrasonic Reflection and Guided Waves in Fluid-coupled Composite Laminates," *J. Nondestr. Eval.*, 9, 51-69 (1990).

4. W. Li and J. D. Achenbach, "$V(z)$ Measurement of Multiple Leaky Wave Velocities for Elastic Constant Determination", *J. Acoust. Soc. Am.*, in press.

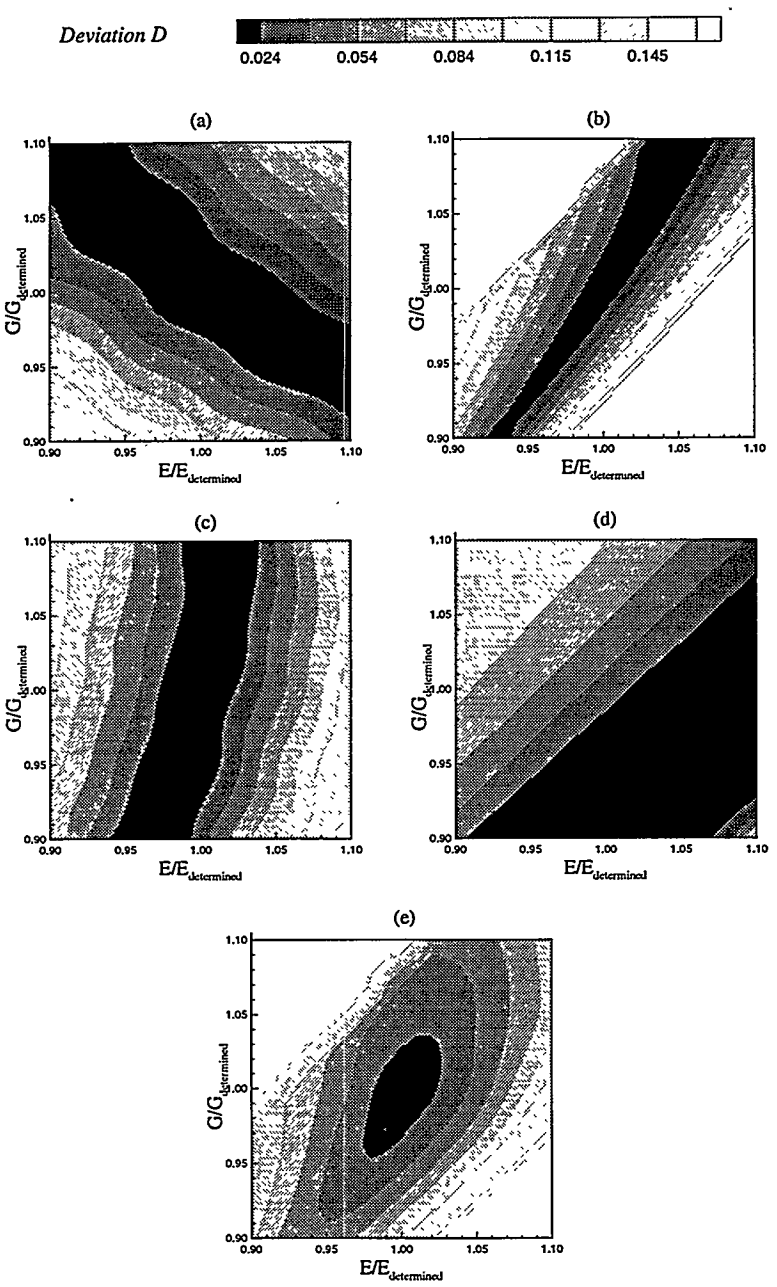


Figure 3: Reduction of region for $D < 0.024$ when multiple leaky acoustic waves/wave parameters are used for AF45 bulk glass. (a) Deviation field when only leaky Rayleigh wave velocity is used; (b) Deviation field when only LSSCW velocity is used; (c) Deviation field when only leaky Rayleigh wave amplitude is used; (d) Deviation field when only LSSCW amplitude is used; (e) all of the above have been used.

CRYOTRIBOLOGY OF DIAMOND AND GRAPHITE

Yukikazu Iwasa, Ahmet F. Ashaboglu, and Ernest R. Rabinowicz

Francis Bitter Magnet Laboratory
and
Department of Mechanical Engineering
Cambridge MA 02139 USA

ABSTRACT

An experimental study was carried out on the tribological behavior of materials of interest in cryogenic applications, focusing on diamond and graphite. Both natural diamond (referred in the text as diamond) and chemical-vapor-deposition (CVD) diamond (CVD-diamond) were used. The experiment was carried out using a pin-on-disk tribometer capable of operating at cryogenic temperatures, from 4.2 to 293 K. Two basic scenarios of testing were used: 1) frictional coefficient (μ) *vs* velocity (v) characteristics at constant temperatures; 2) μ *vs* temperature (T) behavior at fixed sliding speeds. For diamond/CVD-diamond, graphite/CVD-diamond, stainless steel/CVD-diamond pairs, μ's are virtually velocity independent. For each of diamond/graphite, alumina/graphite, and graphite/graphite pairs, the $\partial\mu/\partial v$ characteristic is favorable, *i.e.*, positive. For diamond/CVD-diamond and graphite/CVD-diamond pairs, μ's are nearly temperature independent between in the range 77 – 293 K. Each μ *vs* T plot for pin materials sliding on graphite disks has a peak at a temperature in the range 100 – 200 K.

INTRODUCTION

The principal objectives of this research were to advance the theoretical understanding of low-temperature sliding behavior and to expand the cryogenic tribology data base, particularly for very hard materials. Hard materials are of interest because of their generally high load bearing capacities, low friction coefficients, and wear resistance. The emphasis here is on the frictional behavior of materials of interest to cryogenic applications sliding against CVD-diamond film and graphite.

Early in the next century, high-temperature superconducting (HTS) technology and cryogenics are expected to play key roles in competitive global markets encompassing energy, information, medicine, transportation, space, and basic science [1]. We have been studying

the cryotribology of materials relevant to applications in these fields. A notable recent development in "new materials" is the remarkable progress achieved in materials fabrication techniques, *e.g.*, chemical-vapor-deposition (CVD), for synthesizing diamond films. The CVD-diamond films possess properties very similar to those of bulk diamond, making them promising for new applications such as in HTS magnet technology. The CVD-diamond may be usable in HTS magnets as an interface material as a good electrical insulator as well as a good thermal conductor.

TRIBOLOGY OF DIAMOND AND GRAPHITE

The friction of diamond and graphite has long been known to be low [2]. Gardos and Soriano [3], Samuels and Wilks [4], Kohzaki, *et al.* [5], Jia, *et al.* [6], Bowden and Hanwell [7] have reported on the frictional behavior of diamond, either natural or CVD-processed film, and found both to have low values of friction coefficient, typically 0.05-0.1, with materials exposed in the atmosphere. Mody, *et al.* [8], Petlyuk, *et al.* [9], Khopin [10], Nishiyama, *et al.* [11], Cameron [12] have reported frictional data for graphite and graphite-filled composites. Their results indicate values in the range of 0.05 – 0.15.

EXPERIMENTAL PROCEDURE AND MATERIALS

Apparatus The tests were performed with a rotational pin-on-disk tribometer built to keep specimens at cryogenic temperatures in the range 4.2 – 293 K [13]. A rotating specimen disk slides against three hemispherical specimen pins, symmetrically spaced on a 51-mm bolt circle diameter. The specimens are immersed in a cryogen to achieve and maintain the test temperature: nitrogen (77 K) and helium (4.2 K). The tribometer operates at a constant nominal sliding velocity, ranging from 10^{-7} to 0.1 m/s. Velocity is regulated through a computer-controlled DC servo motor and a set of precision gear reducers. Normal loads (7.5 – 22.4 N) are applied by placing weights on the load ring which transfers the force through a pulley system. A constantan-chromel thermocouple imbedded in the stationary disk sample monitors the specimen temperature.

The friction force is measured by strain arms which prevent the rotation of the specimen pins. The friction force is monitored continuously during test with a computer-based data acquisition system. Although the use of spherical pins results in highly localized contact at the sliding surfaces, the friction coefficients obtained under these conditions are valid because the frictional force is generally independent of the apparent contact area [14].

Materials Table 1 presents a listing of the materials tested, both in disk and pin form. The mechanical properties of the materials are well known at room temperature and to a lesser extent at cryogenic temperatures [15 – 17]. It is not the intention of this project to examine all possible combinations of materials, but rather to test certain combinations that were of interest.

Table 1: Pin and Disk Materials

Pin Materials	*Disk Materials*
Natural diamond	CVD-diamond film
Graphite	Graphite
Copper	Copper
Stainless steel*	Stainless steel†
Alumina	

* AISI440C (pins) and AISI316 (disks).

Both diamond and graphite are crystalline forms of carbon [18]. Diamond is a covalently bonded solid; its bonding arrangement comprises four equivalent covalent bonds directed towards the four corners of a tetrahedron. By comparison, graphite is a layer-lattice material consisting of three equivalent 120° bonds in the basal plane with a fourth, hybrid resonant orbital directed perpendicular to the plane [19]. Graphite's solid lubricant behavior is chiefly due to the easy shear of these interplanar hybrid orbitals. The microwave plasma CVD method was used to synthesize the diamond-coated films for our cryotribological study. In the microwave method, a substrate is placed in a quartz tube and a stream of hydrocarbon gas, *e.g.*, methane diluted by hydrogen, is converted into a film of crystalline diamond by microwave heating. Although silicon is the standard material for substrates, several other substrate materials are possible, including ceramics (silicon carbide, tungsten carbide, silicon nitride, alumina) and metals (tantalum, molybdenum, and tungsten). The diamond is deposited on the substrate at a rate of $0.3 - 0.5\,\mu$m/h. The substrate material used in the experiments was silicon.

Pin and Disk Preparation Graphite pins were obtained as 6.35-mm (1/4-in) diameter rods and turned to their final shape with a radius form tool. Alumina and AISI440C stainless steel pins were obtained in the form of 6.35-mm diameter balls. Three natural diamond pieces, each approximately 1/3 carat and of irregular shape were mounted to allow contact with the specimen disks. The graphite disks were machined from plate stock and faced with a single point cutting tool to remove surface irregularities. The disks were then randomly abraded against 320-grit silicon carbide paper under running water. This cleaned the disks and provided a consistent surface finish between tests. Since water does not wet a surface contaminated with organic contaminants, a disk was deemed clean when a water layer wetted the surface.

Surface Cleaning Before a test, three of the desired pins were mounted in a brass specimen holder, wiped clean with a swab soaked in methanol, allowed to air dry, and placed into the apparatus. Each disk was first rinsed with methanol to displace the water layer and air dried on clean tissue paper. Later, it was placed in an oven at approximately 100°C overnight, to bake away any residual contaminants on the surface.

Testing Two types of measurement were performed: 1) friction (μ) *vs* velocity (v) at constant temperatures (4.2, 77, and 293 K); 2) μ *vs* temperature (T) at a constant sliding velocity of 10^{-5} m/s. For each speed setting, a distance of at least 3 mm was covered to allow the measured friction to achieve a steady value [20]. A velocity of 10^{-5} m/s used in μ *vs* T measurements allowed a sufficient number of data points at a given, though not absolutely constant, temperature and yet limited the total sliding distance of the run to ~2 m. A slowly changing temperature environment could be created in the cooldown process during cryogen transfer into the test cryostat or in the subsequent warm-up process after the cryogen transfer. The friction data could be monitored during either the cooldown or the warm-up process. Both processes showed essentially the same μ *vs* T behaviors. However, because of the lower magnitudes of $\partial T/\partial t$, typically of less than 20 K/h in the warm-up processes compared with 60 – 80 K/h in the cool-down processes, the warm-up processes were used. Because a quiescent helium boil-off rate from the test cryostat was in the range 0.10 – 0.20 l/h, an initial pool of liquid helium provided ~30 hours for each μ *vs* T measurement sequence.

RESULTS AND DISCUSSIONS

Tribological data are often unreproducible and sometimes inexplicable. Friction by its nature is very complex. The factors that contribute to a tribological process may be divided into three classifications: those we know, those we don't, and those we can control. The first is finite, the second may not be, and the third is very finite indeed. The conditions of tribological experiments are particularly crucial. Although utmost care for consistency was taken for each experimental sequence, often ambient conditions, *e.g.*, room temperature, air humidity, cleanliness of the air, have influenced the outcome. When natural (expected) variation of the structures and composition of the materials tested are added to the equation, large variability in results was sometimes encountered. The data gathered will be presented according to two classifications: μ *vs* v; and μ *vs* T. Because of space limitations, only selected sets of data are presented here.

Friction vs Velocity Figures 1 – 6 show μ *vs* v graphs, each with three plots corresponding to temperatures of 4.2, 77, and 293 K. Each μ *vs* v plot is the arithmetic mean of two sets of measurement. Data for the two sets are within 10% at all points for all material pairs. For a given velocity, each point represents the average value of at least 100 data points. The sliding speeds tested ranged from 10^{-6} m/s to 0.1 m/s. For each speed setting, a distance of at least 3 mm was covered to allow the measured friction to achieve a steady value. The applied load was 14.3 N.

Figure 1 – 3 shows, respectively, data diamond/CVD-diamond, stainless steel/CVD-diamond, and graphite/CVD-diamond pairs at three temperatures. Data clearly show that μ at each temperature is virtually velocity independent, confirming a general friction law on the independence of friction on velocity. Data for diamond/graphite, alumina/graphite, and graphite/graphite pairs are shown, respectively, in Figs. 4 – 6. Here for each pair,

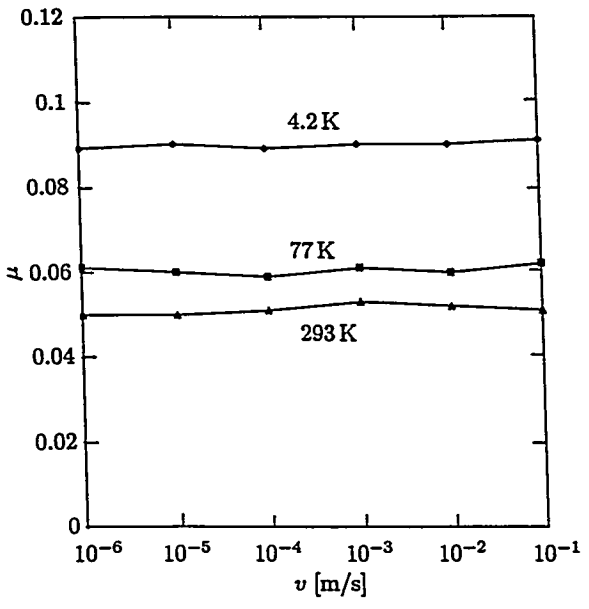


Fig. 1 μ *vs* v for diamond/CVD-diamond pairs.

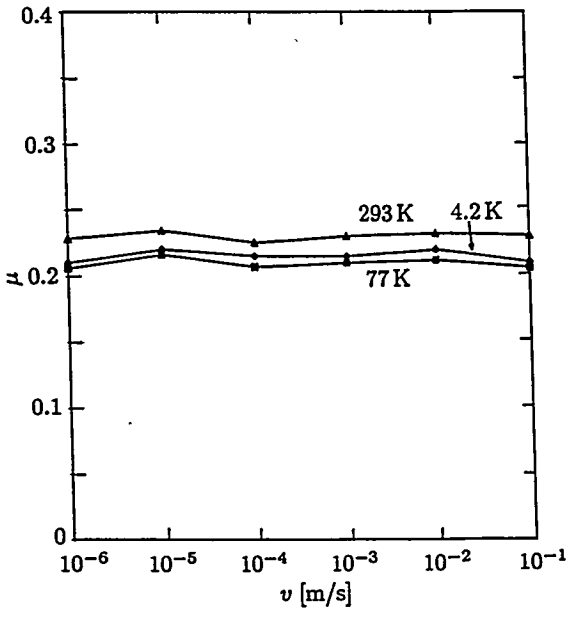


Fig. 2 μ *vs* v for stainless steel/CVD-diamond pairs.

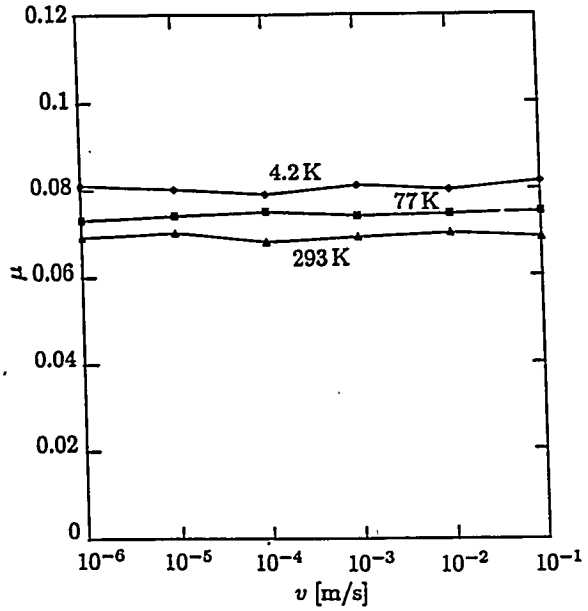


Fig. 3 μ *vs* v for graphite/CVD-diamond pairs.

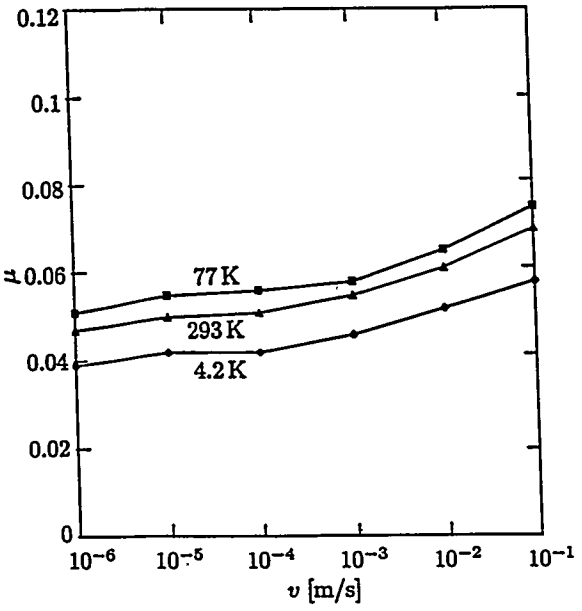


Fig. 4 μ *vs* v for diamond/graphite pairs.

μ increases with v at all three temperatures. The increase can be attributed to the continuous creation of wear particles that are later entrapped between the sliding surfaces. Since graphite is a softer material in comparison with materials like diamond and alumina, the hard pins plow the surface of graphite, wearing graphite and generating wear particles during sliding. When the sliding speed is increased, the rate at which these wear particles are generated is also increased. This results in more wear particles to be entrapped between the sliding surfaces, increasing μ. When graphite pins slide on graphite, again wear particles are generated, this time from both surfaces. Here, because pins and disk are of the same material, adhesion also plays a major role.

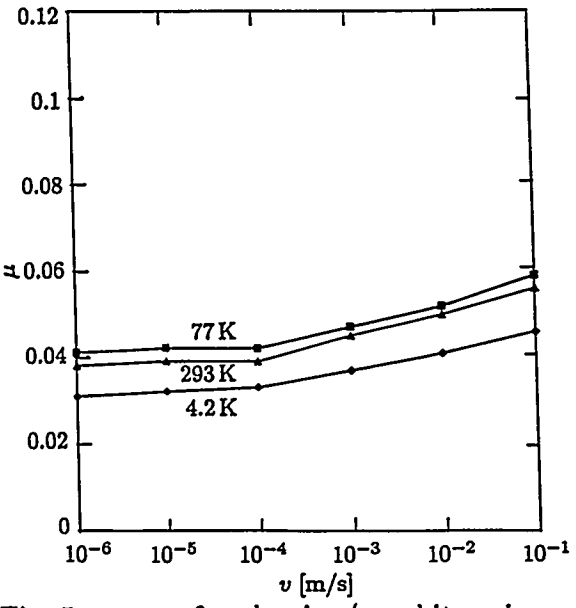


Fig. 5 μ *vs* v for alumina/graphite pairs.

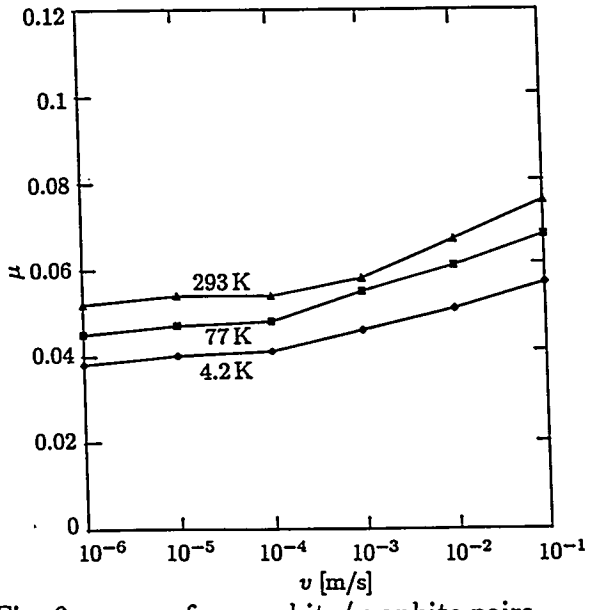


Fig. 6 μ *vs* v for graphite/graphite pairs.

Friction vs Temperature Figures 7 – 10 show μ *vs* T traces, each taken at a constant velocity of 10^{-5} m/s. For each temperature span, the total sliding distance was ~2 m.

Data in Figs. 7 and 8 are, respectively, diamond/CVD-diamond and graphite/CVD-diamond pairs. For the diamond/CVD-diamond pair, μ is 0.09 at 4.2 K and decreases with temperature to 0.05 at 293 K, remaining constant at 0.06 in the middle range 75 – 225 K. According to the adhesion theory, this pair, because both materials are identical, should have high values of μ [21]. Our results appear to be quite the opposite: it has a low value of 0.05 for most of the temperature range tested.

Data for the CVD-diamond/graphite pair (Fig. 8) show a fairly constant μ (0.07 – 0.08) with temperature. Although small, μ varies somewhat in the temperature range 50 – 250 K. This variation may be attributable to the wear particles of graphite, the softer of the material pair. As sliding proceeds wear particles are generated and entrapped between the sliding surfaces, increasing the friction. As the sliding continues, these particles are either plastically deformed or removed from the sliding path, hence reducing the friction. This process of wear particles being generated and later eliminated from the sliding surface could account for the variation in the coefficient of friction. As might be surmised from the three data points at $v = 10^{-5}$ ms/s of Fig. 2, the μ *vs* T plot of stainless steel/CVD-diamond pairs (not shown here) indicates that μ increases gradually above 77 K, reflecting the fact that μ is inversely proportional to stainless steel's hardness.

Data in Figs. 9 and 10 are, respectively, for pins of natural diamond and graphite *vs* graphite disks. For the diamond/graphite pair, μ is roughly temperature independent and equal to ~0.04, from 4.2 K to ~70 K. μ increases with temperature in the range ~70 – ~160 K. This increase is attributed to macroscopic stick-slip, typically observed in this temperature range [22 – 24]. Above ~160 K, μ decreases, reaching at room temperature a value roughly the same as that at 4.2 K. The graphite/graphite data (Fig. 10) are similar to the diamond/graphite data, except here the peak value is ~0.06.

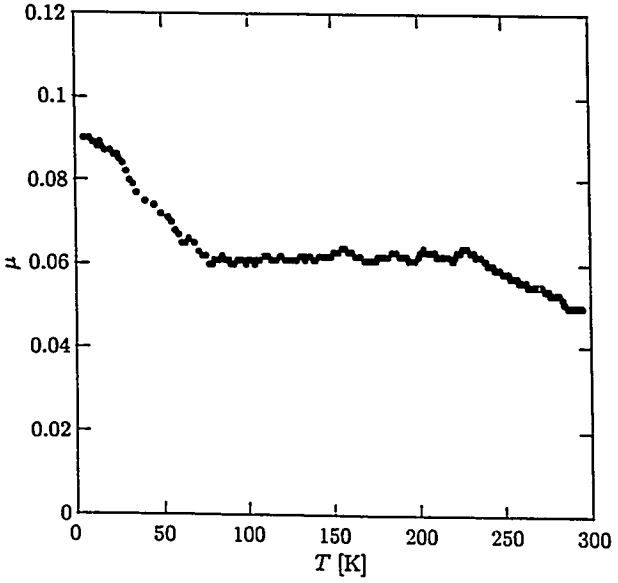


Fig. 7 μ *vs* T at $v = 10^{-5}$ m/s for diamond/CVD-diamond pairs.

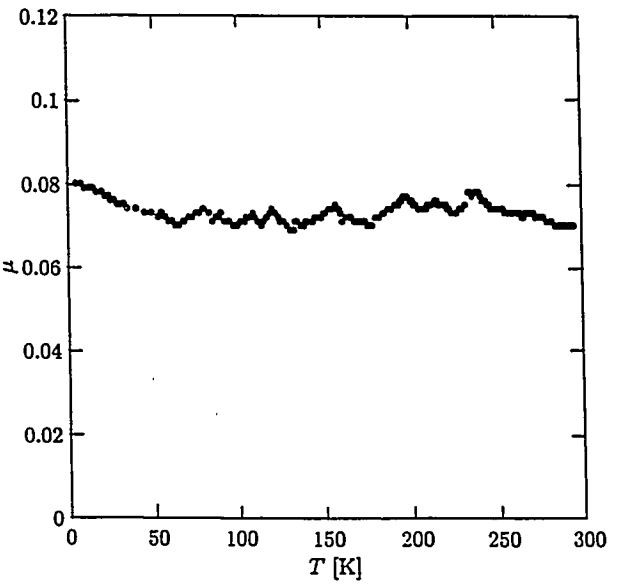


Fig. 8 μ *vs* T at $v = 10^{-5}$ m/s for graphite/CVD-diamond pairs.

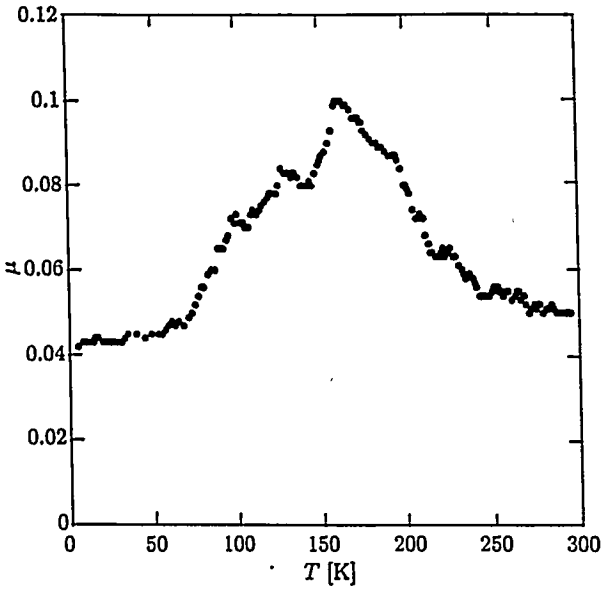


Fig. 9 μ *vs* T at $v = 10^{-5}$ m/s for diamond/graphite pairs.

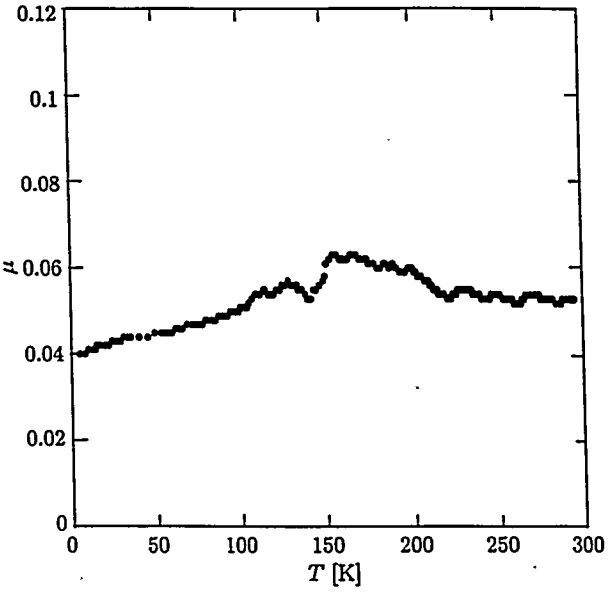


Fig. 10 μ *vs* T at $v = 10^{-5}$ m/s for graphite/graphite pairs.

CONCLUSIONS

We have presented a selection of cryotribology data for diamond (natural and CVD films) and graphite, all in terms of μ *vs* v plots and μ *vs* T plots. Based on our data, both presented here and not presented owing to lack of space, we may make the following conclusions. 1) For natural diamond, graphite, and stainless steel pins *vs* CVD-diamond coated disks, μ's are essentially velocity-independent. For diamond pins sliding on stainless steel disks, μ is roughly velocity-independent at 4.2 and 77 K, but increases with v at 293 K; apparently at temperatures above 77 K and certainly at 293 K, stainless steel is a "soft" material in comparison with diamond. 2.) Although for natural diamond, graphite, and stainless steel pins *vs* CVD-diamond coated disks, μ's are fairly temperature-independent, there are subtle variations. In the diamond/CVD-diamond pairs, μ is at a maximum at 4.2 K and decreases with temperature. In the graphite/CVD-diamond pairs there is a variation in μ with T in the mid temperature range, due perhaps to wear particles generated by the graphite pins. 3.) The materials sliding against graphite disks show a peak frictional coefficient value within the temperature span 100 – 200 K. As is the case with diamond/stainless steel pairs, there is a definite correlation between friction and material hardness: μ is inversely proportional to material hardness. Thus μ is higher at higher temperatures, because hardness generally decreases with temperature.

ACKNOWLEDGEMENT

The work presented here was performed under the auspices of the U.S. Department of Energy, Office of Basic Sciences and is based in part on a thesis submitted by AFA in partial fulfillment of the requirements for the degree of Master of Science in Mechanical Engineering at MIT, February 1996. The authors thank Drs. Koji Kobashi and Takeshi Tachibana of Kobe Steel Ltd. for supply of CVD-diamond coated disks.

REFERENCES

1. Yukikazu Iwasa, *Case Studies in Superconducting Magnets*, (Plenum Press, New York, 1994).
2. F.P. Bowden, and D. Tabor, *Friction and Lubrication*, (Latimer Trend & Co., 1967).
3. M.N. Gardos, and B.L. Soriano, "The effect of environment on the tribological properties of polycrystalline diamond films," *J. Mater. Res.* **5**, No. 11 (1990).
4. B. Samuels, and J. Wilks, "The friction of diamond sliding on diamond," *J. Mater. Science* **23**, 2846 (1988).
5. M. Kohzaki, K. Higuchi, S. Noda, and K. Uchida, "Tribological characteristics of polycrystalline diamond films produced by chemical vapor deposition," *J. Mater. Res.* **7**, (1992).
6. K. Jia, Y.Q. Li, T.E. Fischer, and B. Gallois, "Tribology of diamond-like carbon sliding against itself, silicon nitride, and steel," *J. Mater. Res.* **10**, No. 6 (1995).
7. F.P. Bowden, and A.E. Hanwell, *Proc. Roy. Soc.* **A 295**, 233 (1966).
8. P.B. Mody, T. Chou, and K. Friedrich, "Effect of testing conditions and microstructure on the sliding wear of graphite fiber/PEEK matrix composites," *J. Mater. Science* **23**, 4319 (1988).
9. Y. Yamada, S. Ueda, and K. Tanaka, "Effects of temperature on the friction and wear of poly(amide-imide)-based composites," *J. Japan Soc. Lub. Eng.* **33**, 458-466 (1988).
10. A.M. Petlyuk, L.N. Sentyurikhina, O.V. Lazovskaya, and T.P. Yukhno, "Temperature dependence of wear resistance of solid lubricant coatings containing MoS_2 and graphite," *Sov. J. Fric. and Wear*, **8**, 132 (1987).
11. Katsuhiro Nishiyama, Keisuke Taguchi, Tsuneo Ohdaira, and Abe, "Graphite solid-state lubricant generating low coefficient of friction at high temperatures," *J. Jap. Soc. Powder and Powder Metallurgy* **41**, 613 (1994).
12. A. Cameron, *The Principles of Lubrication*, (Longmans Green & Co., 1966).
13. Y. Iwasa, R. Kensley, and J.E.C. Williams, "Frictional properties of metal-insulator interfaces," *IEEE Trans. Mag.* **MAG-15**, 36 (1979).
14. E. Rabinowicz, *Friction and Wear of Materials*, (John Wiley, New York, 1965).
15. R.B. Scott, *Cryogenic Engineering* (Reprinted by Met-Chem Research, Boulder, CO.)
16. D.A. Wigley, *Materials for Low-temperature Use*, (Oxford University Press, 1978).
17. *Cryogenic Materials Data Handbook*, (Air Force Materials Laboratory, Wright-Patterson Air Force Base, OH, 1970).
18. A.R. Von Hippel, *Molecular Science and Molecular Engineering*, (MIT Press and John Wiley & Sons, New York, 1959).
19. *Physical Properties of Diamond*, Ed. R. Berman, (Clarendon Press, Oxford, 1965).
20. E. Rabinowicz, "Autocorrelation analysis of the sliding process," *J. Appl. Phys.* **27**, 131 (1956).
21. E. Rabinowicz, "Determination of compatibility of metals through static friction test," *ASLE Transactions* **14**, 198 (1971).
22. F.P. Bowden and G.W. Rowe, "The friction and mechanical properties of solid krypton," *Proc. Roy. Soc.* **A.228**, 1 (1955).
23. M.B. Peterson, and S. Ramalingam, "Coatings for tribological applications," *Fundamentals of Friction and Wear of Materials* (ASM, Metals Park, OH) 119 (1981).

THE TEMPORAL NATURE OF FORCES ACTING ON METAL DROPS IN GAS METAL ARC WELDING

Lawrence A. Jones Thomas W. Eagar Jeffrey H. Lang

Massachusetts Institute of Technology
Cambridge, MA, 02139 USA

ABSTRACT

At moderate and high welding currents, the most important forces in gas metal arc welding acting on the molten electrode are magnetic forces arising from the interaction between the welding current and its own magnetic field. These forces drive the dynamic evolution of the drop and also depend on the instantaneous shape of the drop. In this paper, experimentally observed manifestations of magnetic forces are shown, and a technique for approximating the temporal evolution of these forces from experimentally measured drops shapes is reported. The technique provides quantitative data illustrating the large increase in the magnetic forces as a drop detaches from the electrode.

The temporal evolution of the metal drop geometry during drop detachment in gas metal arc welding (GMAW) has an important effect on the process because the forces acting on the drop depend on its shape and they change dramatically over the course of drop detachment. If the geometric evolution of the drop is ignored, it is impossible to quantitatively explain such phenomena as the substantial initial velocity a drop has at the moment it detaches from the electrode. It is also difficult to quantitatively explain the effectiveness of current pulsing. As another example, if the geometric evolution of the drop is ignored, it is difficult to quantitatively explain why some welding conditions result in the stable, axisymmetric detachment of drops and other conditions do not.

Over a wide range of conditions, the effects of magnetic forces may be seen acting on drops detaching from a GMAW electrode. A drop detaching at 260 A and 29 V is shown in Figure 1.* For a 1/16-inch diameter ER70S-3 electrode in Ar–2%O_2, this current is in the upper end of the globular transfer region. The distinctly flattened (oblate) shape of the drop is a result of the magnetic forces acting on the drop. It will be seen below that when a drop of approximately the same size detaches under very low-current conditions, the drop is largely spherical and is not flattened (Figure 5). Also, although it is not apparent

*The images in Figure 1 and subsequent figures were obtained using the optical technique described in [1]. An arc is present in all of the images, but it is rendered virtually invisible by the optical technique.

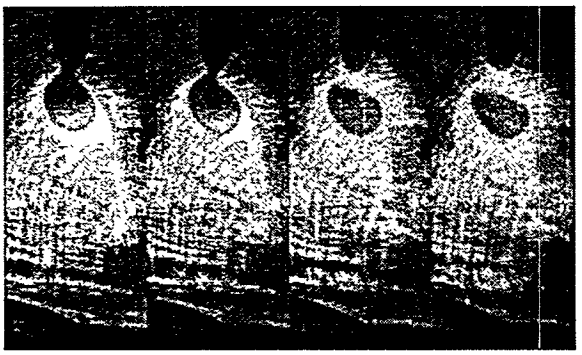


Figure 1: Drop detachment with constant current 260 A and 29 V electrode positive. The electrode is 1/16-inch diameter ER70S-3 in Ar–2%O_2. Note the vertically flattened (oblate) shape of the drop during and after detachment.

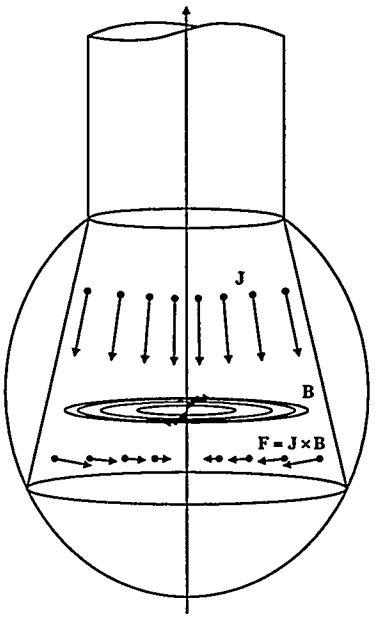


Figure 2: Magnetic forces in a welding drop arise from the interaction between the welding current and its own magnetic field.

from the images in Figure 1, the drops accelerate off the end of the electrode at a rate substantially greater than the acceleration of gravity. This excess acceleration is due to the sharp increase of magnetic forces acting while the drop is detaching and due to the force of plasma flow while the detached drop is in flight.

The magnetic forces arise from the interaction of the welding current with its own magnetic field as illustrated in Figure 2. If the current diverges in the drop (as shown in Figure 2), then downward forces act on the fluid in the drop, and if the current converges in the drop, then upward forces act on the fluid in the drop. Unlike gravity, which acts uniformly in the vertical direction on the fluid (assuming the density of the fluid is spatially uniform) and is an irrotational force, the magnetic forces do not act uniformly and there is a rotational component of force acting on the fluid.

At higher currents (~400–470 A), the heat of the arc causes molten metal to stream off of the electrode forming a column of liquid which then breaks up into drops. The magnetic forces due to the current flowing through this liquid aid the breakup of this column into drops. Even without magnetic forces acting, the breakup of a liquid column will occur due to mechanical disturbances in the fluid flow which cause disturbances in the curvature of the surface of the column. If current is flowing in the column, the magnetic forces aid the breakup of the column. At very high currents, the magnetic forces are apparent in the appearance of rotating streams of metal as shown in Figure 3 for 480 A and 35 V. At such high currents the magnetic forces are significant compared to the inertial forces in the column and slight asymmetries in the column cause asymmetric radial magnetic forces which move the column away from its straight line of flow. It can be shown that such a bend in the current path also results in azimuthal components of the magnetic forces. The combination of asymmetric radial forces and azimuthal forces results in the spiraling mo-

tion of the column as seen in Figure 3. This motion is known as a kink instability [2].

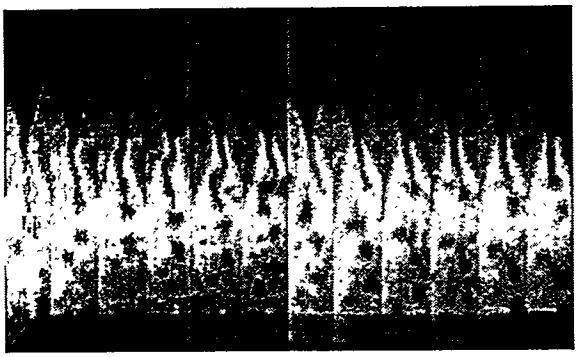


Figure 3: Drop detachment with constant current 480 A, 35 V electrode positive. The electrode is 1/16-inch diameter ER70S-3 in Ar–2%O_2. Note the spiraling detachment of drops from a column of liquid metal.

By pulsing the welding current, magnetic forces may be used to detach drops from the electrode. The shape of the pulse, its magnitude, duration, repetition rate (frequency), and the current level between pulses are all parameters in this process. An incorrect choice of parameters may result in an unstable arc, a reduced metal deposition rate, incorrect base-plate heating, and/or spatter. Figure 4 shows the results of applying a pulse having too great an amplitude. The magnetic forces overwhelm the surface tension forces and the detachment process proceeds asymmetrically. A thin, filamentous neck is violently snapped to the side resulting in fine spatter. Due to the violence of the detachment, the main drop sometimes breaks apart resulting in coarse spatter. If the current is not reduced immediately after the detachment of a drop, the heat of the arc will generate a subsequent small drop (a droplet) and the magnetic forces will detach it asymmetrically which also generates coarse spatter. Such a droplet is seen forming in the last image in Figure 4. In this particular case,

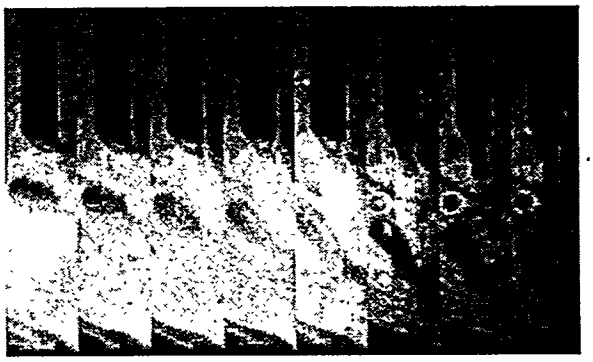


Figure 4: Drop detachment with pulsed current 550/40 A, 6 Hz, 2% duty cycle, and 18 V electrode positive. The electrode is 1/16-inch diameter ER70S-3 in Ar–2%O_2. Note the whipping of the drop neck, the severe distortion of the drop, and the presence of a small droplet on the electrode at the end of the sequence which often detaches and flies to the side as spatter.

the current is reduced before the droplet detaches from the electrode and the droplet is pulled back onto the electrode by surface tension.

Under certain pulsing conditions, drops will detach from the electrode with minimal distortion and no droplets. Such a case is shown in Figure 5. In this case, the current pulse imparts enough momentum to the drop to cause detachment, but ends well before the drop detaches, thereby avoiding the formation of droplets. At the time of detachment, the welding current is 40 A and the drop is almost spherical. The lack of distortion of the drop is in marked contrast to the distortion of the drop seen in Figure 1 at 260 A constant current. When the current is very low at drop detachment, the magnetic forces do not act violently on the neck of the drop as seen in Figure 4, and since the drop is largely undistorted as it detaches, the probability is low that it will break apart.

If a current pulse lower than 290 A

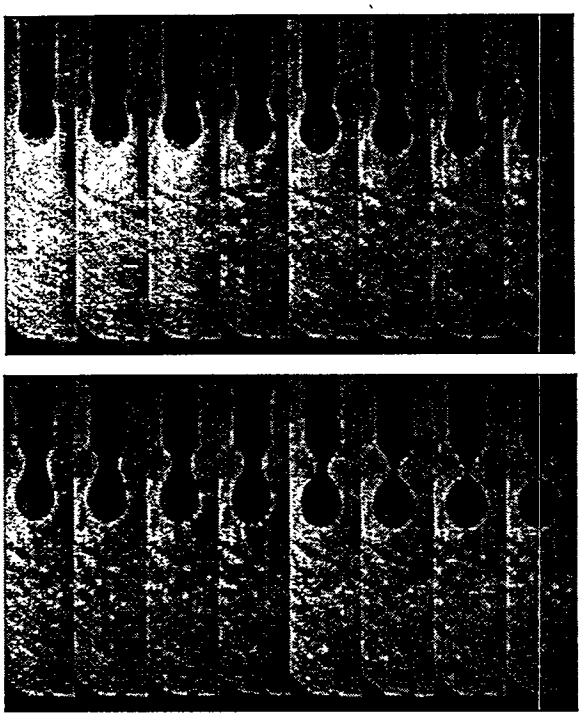


Figure 5: Drop detachment with pulsed current 290/40 A, 5 Hz, 2% duty cycle, and 18 V electrode positive. The electrode is 1/16-inch diameter ER70S-3 in Ar–2%O_2. Note the symmetric detachment of a drop and the lack of drop distortion.

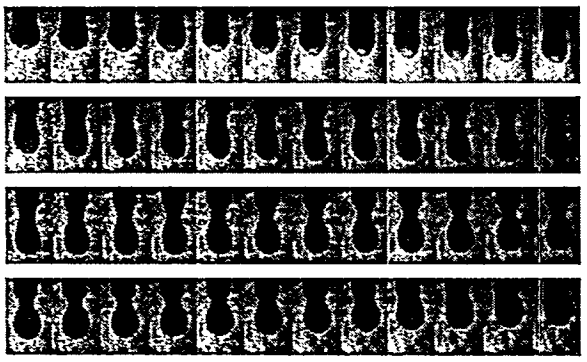


Figure 6: Drop response with pulsed current 280/40 A, 5 Hz, 2% duty cycle, and 18 V electrode positive. The electrode is 1/16-inch diameter ER70S-3 in Ar–2%O_2. The current pulse is insufficient to detach the drop. Subsequent current pulses will also fail to detach the growing drop.

is used, the momentum impulse supplied by the magnetic forces will not be great enough to overcome surface tension and the drop will not detach, as shown in Figure 6. Without the assistance of magnetic forces, the surface tension forces, which act to restore equilibrium, dominate and overcome the inertial force of the drop. During the next several current pulses, the mass of the drop will be greater and the magnetic forces will be unable to impart enough momentum to the drop to make it detach. The drop will continue to grow until gravity, aided by the disturbances caused by current pulses, overcomes the surface tension holding the drop on the electrode.

The magnetic diffusion time and the magnetic Reynolds number in a drop of molten steel on the end of a GMAW electrode are both very small. The very short magnetic diffusion time indicates that on the time-scale of drop motion, the diffusion of the magnetic field throughout the drop is essentially instantaneous. The very small magnetic Reynolds number indicates that the magnetic diffusion process is much faster than fluid convection in the drop. Therefore, in a gas metal arc welding electrode the magnetic field is unaffected by the fluid velocity in the drop. Rather, the distribution of the magnetic field is determined by the instantaneous geometry of the current path in the drop and, assuming uniform fluid conductivity, the current path is determined by factors other than the fluid velocity.

In [3], the magnetic stress tensor is used to calculate the total vertically-directed magnetic force acting on a generalized, axisymmetric drop shape. The motion of the fluid interacts with the magnetic field only in that it affects the resulting shape of the drop and hence the shape of the current path. Therefore, if the shape of the current

path is known at each instant, the magnetic stress tensor may be used to calculate the total vertically-directed magnetic force on the drop (the sum of the irrotational and rotational parts).

As a drop attempts to detach from the solid electrode, a neck forms. The current density in the drop neck increases and the divergence of the current increases. Both phenomena cause the magnetic forces acting on the drop to increase and result in a measurable acceleration of the drop upon detachment. The time during which a neck initiates and collapses is short compared to the total drop growth time but it is during this brief time that the magnetic forces are most important.

During drop necking, the lower part of the drop is well modeled by a truncated ellipsoid and the upper part is well modeled by a volume formed by rotating a third-order polynomial about the vertical axis, as illustrated in Figure 7. Third-order polynomial volumes were chosen to model the neck because these shapes are completely defined by the geometrical boundary conditions at the top of the truncated ellipsoid and the bottom of the electrode.

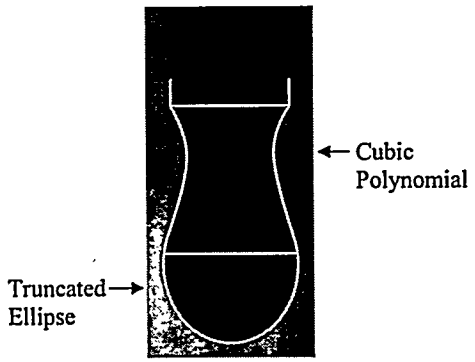


Figure 7: Experimentally-observed necking drop shape modeled with a truncated ellipsoid and a polynomial volume.

A truncated ellipsoid is shown in Figure 8 along with the portion of a polynomial volume below its waist (its narrowest point). In [3], the total vertically-directed magnetic force on a generalized shape is calculated to be

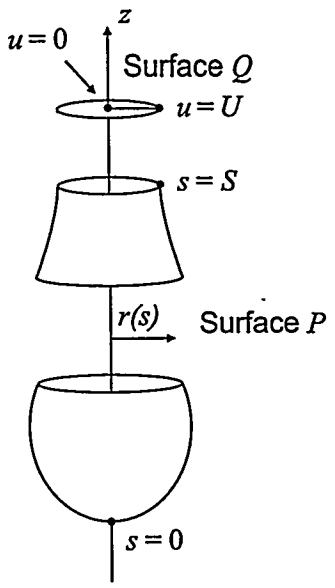


Figure 8: Stress tensor surfaces.

$$f_z = -\frac{\mu_o I^2}{16\pi} + \mu_o \pi \int_0^S \left[\int_0^s j(s')r(s')ds'\right]^2 \frac{dr}{ds'}\frac{ds'}{r(s')}, \quad (1)$$

where I is the portion of the welding current emerging from the drop and $j(s)$ is the density of this current along the surface s of the drop. If there is no drop neck, if the truncated ellipsoid is a truncated sphere, and if the current density $j(s)$ is taken to be constant,* then both integrals in Eq. 1

*No experimental measurements are available in the literature about the distribution of the current density on a GMAW electrode due to the difficulty of making such measurements in the harsh environment of the arc next to the free surface of a drop. However, observations of drops detaching from the electrode empirically suggest that a constant $j(s)$ may be a reasonable approximation [4].

may be solved in closed form. The resulting expression is commonly used to compute the total vertically-directed magnetic force acting on a welding drop:

$$f_z = \frac{\mu_o I^2}{4\pi}\left[\frac{1}{4} - \ln\left(\frac{a\sin\Phi}{r_e}\right) + \frac{1}{1-\cos\Phi} + \frac{2}{(1-\cos\Phi)^2}\ln\left(\frac{2}{1+\cos\Phi}\right)\right], \quad (2)$$

where a is the radius of the sphere, r_e is the radius of the electrode, and Φ is the angle describing the portion of the drop over which current is emerging ($\Phi = 0 \Rightarrow$ no current emerges from the drop, and $\Phi = \pi/2 \Rightarrow$ current emerges from the lower hemisphere of the drop).

While a drop is necking—a time when the magnetic forces increase dramatically—truncated-ellipsoid and polynomial-volume combinations are required to model the shape of the drop, as shown in Figure 7. The integrals in Eq. 1 cannot be solved analytically even for these simple approximations of the fluid surface of a necking drop. Over the surface of the truncated ellipsoid, the inner, squared integral may be solved analytically, but not the outer integral. Over the surface of the polynomial volume, the inner, squared integral cannot be solved analytically (and therefore the outer integral also cannot be solved analytically).

Truncated ellipsoids and polynomial volumes were fit to drop profiles observed during current pulsing similar to those shown in Figure 5. The instantaneous vertically-directed magnetic forces acting on these fitted shapes were then computed numerically using Eq. 1. Constant current density $j(s)$ was assumed to cover the entire surface of the shapes up to the waist of the polynomial volume. Experimentally-measured current pulses and the corresponding computed vertically-directed magnetic forces are shown in Figure 9.

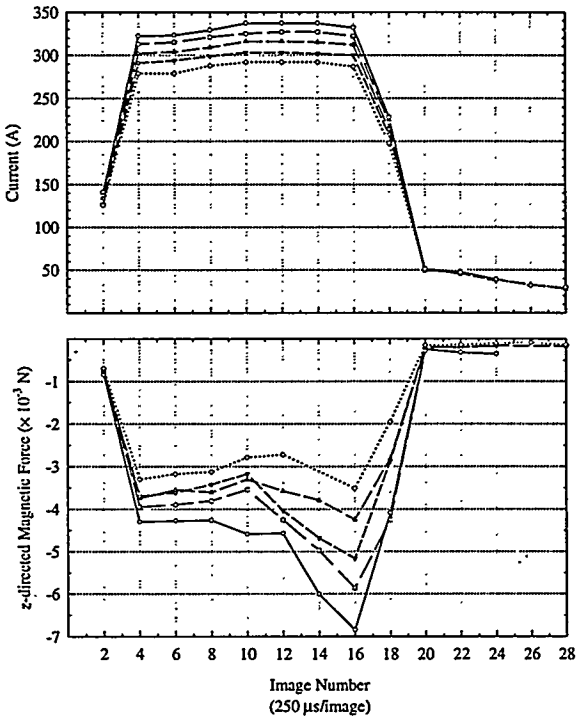


Figure 9: Total vertically-directed magnetic force during current pulsing with the Ar–2%O_2 current emission model.

The forces in Figure 9b show that initially as a drop elongates in response to the downward (negative) magnetic force, the magnitude of the force decreases slightly. This decrease is because in an elongated (prolate) drop the current diverges less. Once a neck begins to form, however, the magnitude of the magnetic force increases rapidly. This increase is because the narrowing of the neck results in greater subsequent current divergence. This observation is not surprising or new since it has long been known that magnetic forces act to "pinch" drops off the end of an electrode. The results shown in Figure 9, however, represent the first time the temporal evolution of the magnetic force has been computed using close approximations of experimentally measured drop shapes.

By using shape approximations as in Figure 7, a simple, dynamic model of drop detachment may be developed, as illustrated in Figure 10. Such a model is developed in [4] and is compared with extensive measurements of drop detachment obtained from high-speed images of gas metal arc welding.

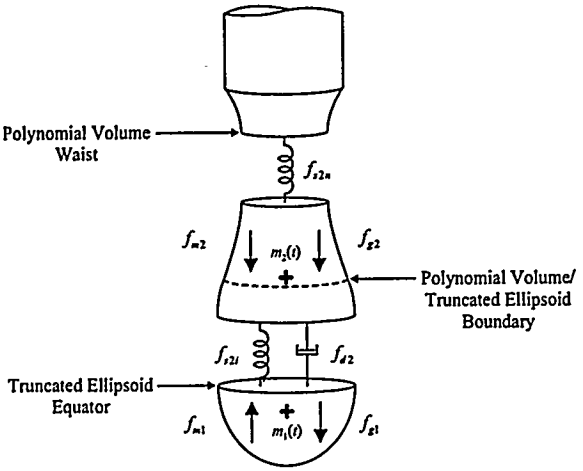


Figure 10: Illustration of the dynamic model developed in [4] using truncated-ellipsoid/polynomial-volume shapes.

Acknowledgement

This work was supported by the United States Department of Energy, Office of Basic Energy Sciences.

References

[1] C. D. ALLEMAND et al., "A Method of Filming Metal Transfer in Welding Arcs," *Welding J.*, **64**(1), 45–47 (1985).

[2] A. DATTNER, "Current-Induced Instabilities of a Mercury Jet," *Arkiv för Fysik*, **21**(7), 71–80 (1962).

[3] J. C. AMSON, "Lorentz Force in the Molten Tip of an Arc Electrode," *British J. of Appl. Phys.*, **16**, 1169–1179 (1965).

[4] L. A. JONES, *Dynamic Electrode Forces in Gas Metal Arc Welding*, PhD thesis, Massachusetts Institute of Technology, 1996.

PARTIAL CONTROL OF COMPLEX SYSTEMS WITH APPLICATION TO THE FLUIDIZED CATALYTIC CRACKER

Irven H. Rinard and Reuel Shinnar

Department of Chemical Engineering
City College of CUNY

ABSTRACT

The research deals with the control of complex nonlinear system with a limited number of manipulated variables. In many chemical processes the number of variables that make up the specifications and constraints exceeds the number of manipulated variables available. Furthermore, model information is limited. The goal of our work is to study the design of the control system and the conditions required to achieve adequate control for such cases. A Fluid Catalytic Cracker was chosen to illustrate and test our approach. This paper presents a short overview and summary of our approach and results.

INTRODUCTION

The fluidized catalytic cracker (FCC) is one of the most important processing systems in oil refineries. It is a complex system to operate and control, one which is not fully understood, and one on which the petroleum industry has spent considerable time and research effort over the past sixty years. The heart of the problem is that there are many more process variables that one needs to control than there are manipulatable variables with which to do so. This is characteristic of most complex systems such as oil refineries and chemical plants as well as social systems, even the economy. The other salient feature of such systems is that their behavior is not well understood. This is true of the FCC; it is certainly true of the economy.

The goal of the present work is to develop a practical theory of partial control that can be applied to complex systems. However, to insure that it is a relevant theory, its development has been closely tied to a significant real-world example, namely, the FCC. Further, this work has been done in close collaboration with Mobil, one of the major petroleum companies. Our research in this area has reached the state where its results can be implemented and utilized by industries in many areas. It

already has led to improvements in the control of FCC's and these have been adopted by industry. It is the goal of this paper to give an overall picture of the theory and its application to control of FCC's. For details the reader is referred to the papers that have already been published under this grant [1-4].

CONCEPTS OF PARTIAL CONTROL

It is not obvious that one can control a complex nonlinear system by only a limited number of manipulated variables, particularly when the behavior of that system is not completely understood. Based on our theoretical results and the study of actual complex processes, we have developed an approach and a theoretical framework as how to choose the manipulated and measured variables for a control system in order to get the best results. The control ideas underlying this research were developed and published by Professor Shinnar in two papers [5,6] and were based on his work in process control in both academia and industry.

In general, in the design of a control system, one first identifies the set of variables that need to be controlled, namely, those that are related to product specifications and process operating constraints. Denote these by Y_{pj} and their number by N_c. Next, one identifies a set of suitable manipulated variables with which to effect this control. Denote these by U_m and their number by N_m. If $N_m = N_c$, then all of the controlled variables can be controlled, at least in principle, to exact set points. We call this exact control. It is the type of control to which the vast majority of the control literature and textbooks are devoted. A different situation arises quite often in the control of large, complex systems, one where $N_m < N_c$. Now, not all the Y_{pj} can be controlled at exact set points. Instead, we must settle for something less strict, namely, that these variables are kept within an acceptable operating space, i.e., $Y_{pjmin} \le Y_{pj} \le Y_{pjmax}$. We call this partial control.

The basic idea of partial control is that we choose a subset of the controlled variables, Y_{cd}, and the manipulated variables, U_{cd}, which results in a square control matrix. This primary control matrix will be used for exact control of the chosen subset of controlled variables. Which and how many variables are chosen depends upon a number of factors that are discussed below. For many systems an appropriate choice of the primary matrix will allow us to maintain all the Y_{pj} within the acceptable operating space by adjusting the set points of the variables in the control matrix (Fig. 1). Choice of this primary control matrix is critical to the success of the overall control. If one has a very good model, one would feed the measurements of Y_{pj} directly into a model based on direct computer control to adjust all U_j. This does not allow integral control and requires a much better model than we normally have.

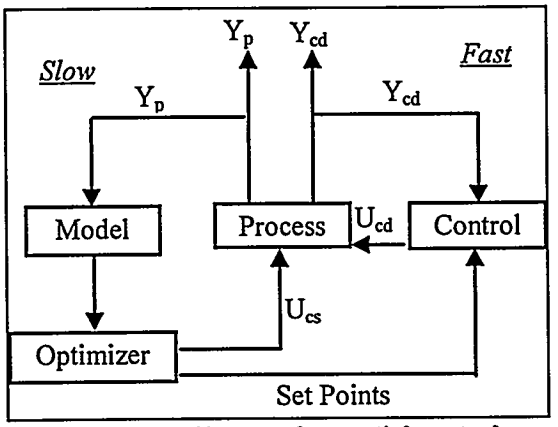


Figure 1: Block diagram for partial control

There are several considerations in the choice of the primary control structure. We would like the primary matrix to be as small as possible. To achieve this, our choice Y_{cd} and U_{cd} must satisfy the following criteria:

A. Dominance

In partial control one chooses a set of measured variables Y_{cd} that are kept by exact control at chosen set points. If one or more of the output variables of concern Y_{pj} moves out of the desired range,

one adjusts the set points of Y_{cd} to keep the operation of the process acceptable, both economically and in terms of product specifications.

It is very important how one chooses this set of measured variables Y_{cd}, some of which may be of no direct interest by themselves. The most important concern is dominance which means that changing the set point of a measured variable has a strong impact on many or all the process variables of interest. To do so requires that we have available manipulated variables which have a strong impact on Y_{pj} and Y_{cd}. A good example is chemical reactor temperature which affects all the kinetic processes just as climatic temperature affects all biological processes. The difference is that in a reactor we can control it. Change the temperature and we change almost everything else: the production rate, the yield or product distribution, even the stability of the operating point. Similarly, interest rates are a dominant variable in the economy. Raise interest rates and many, many other things change. There are many other examples.

B. Sufficiency

It is overly optimistic to believe that a single dominant control loop will lead to acceptable control of the many important variables that arise in complex systems. This leads to the concept of sufficiency. How many and what dominant variables must be controlled so that the system is stable and that all the other variables are maintained within specified limits? What constitutes a sufficient set of such variables?

It is important to ensure that a process has enough manipulated variables with a wide enough range of variability to insure that the process can deliver product that meets its quality specs. The control must simultaneously ensure that process conditions do not violate operating constraints, such as metallurgical limits, flooding or entrainment flow velocities, etc., which we include in Y_p. The control should also achieve this goal in economically acceptable ways. Since the availability of the manipulated variables is determined during the design phase, one needs to understand how the design impacts control. Another important consideration is that of the product specifications. These are determined externally by market requirements but set many of the constraints on Y_{pj}. This will have a major impact on the design of both the unit itself as well as the control system. One has to understand how the design interacts with the specifications in order to come up with an economically acceptable compromise. Better partial controllability is achieved at the expense of adding more manipulated variables but these must be chosen to provide the widest possible range of operating conditions at acceptable cost.

Our research deals with this important problem and we have developed these criteria for the FCC. These results are important in developing strategies for better design procedures, and are a good example demonstrating how basic research in control can lead to better design of chemical processes.

C. Modellability

The choice of dominant variables is itself constrained by the number of manipulated variables available for control. These must satisfy several criteria, the primary ones being that they have a significant gain with respect to the dominant variables to be controlled and that they affect these variables on a suitable time scale.

This leads to the very important question of modellability. To indirectly control a complex system where anywhere from a dozen to a hundred or more variables are of interest and to do this by direct control of a small set, say two to six, dominant variables requires a model. This model must satisfy two criteria:

- How well can the process variables Y_{pj} be related to the dominant variables, particularly in steady-state? This is the major determinant of how well we can do with respect to constraint management.
- How well can the dominant variables be related to the manipulated variables? This is a major

determinant of how well we can stabilize and control the primary control structure. This has to be addressed from system to system; examples will be discussed later. In general full scale reactors have more complex and less uniform flows and hopefully the variables sensitive to these variations are not dominant for Y_p. This is essential not only for control but also for safe scale up.

D. Time Scale of Response

The primary control structure must be chosen so that it responds on a time scale commensurate with the time scale on which the process variables Y_{pj} respond to disturbances to the system. This time scales are determined by the design. The designer faces some compromises for complex processes. In adiabatic processes such as the FCC one introduces intentionally sufficient thermal inertia to allow the operator to intervene in order to prevent catastrophic instabilities and correct mistakes. In the FCC the holdup in the regenerator provides this thermal inertia. The reactor has a much faster response which is beneficial. Fast response in the process industries is often less important than in aerospace.

E. Nonlinearities and Stability

In complex, nonlinear systems assuring stability of the control system is a major concern. Such systems often exhibit multiple steady-states, some of which are open-loop stable, others which are not. Input multiplicities and regions of zero open-loop gain further complicate the task of choosing a control structure. The pertinent nonlinear features of the FCC were analyzed in [2] as part of this program and have critical impact on the control strategies.

APPLICATION TO THE FCC

The fluidized catalytic cracker (FCC) was chosen as an example of the kinds of complex systems encountered in the petroleum and chemical industries which are operated subject to partial control. As shown in Fig. 2, the FCC consists of two principal units. In the first, the reactor, hot catalyst is contacted with oil in a fluidized bed. It supplies heat whereby the oil is first vaporized, then cracked to more desirable products such as gasoline and heating oil. The catalyst is cooled down and simultaneously deactivated by the formation of coke which is a byproduct of the cracking reactions. The catalyst is reactivated and reheated by circulating it to the second unit, the regenerator, where it is contacted with air to combust the coke. The energy released in the regenerator is recycled back to the reactor by the circulating catalyst. One important condition for the operation of the FCC is that the heat generated by combusting all the coke generated equals the heat absorbed in heating and cracking the oil feed.

There are many variables in the FCC that need to be controlled to a greater or lesser extent. These include the fractional conversion of the oil feed to products, the gasoline yield and its composition, the N_{ox} and CO levels of the flue gas leaving the regenerator, and

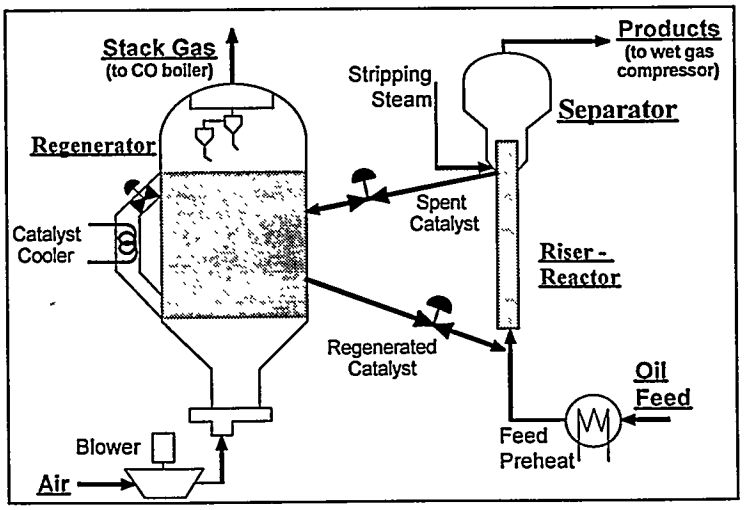


Figure 2: Schematic Diagram of FCC

olefin content of the product gas which is used as a feed for alkylation, to name just a few. There are but a few variables available with which to exert this control on the appropriate time scale, primarily the feed rate of air to the regenerator, F_{air}, the catalyst circulation rate between the two units, F_{cat}, and the temperature of the feed, T_{feed}

Some units have a catalyst cooler to allow processing of heavy feedstocks, for coke make in the FCC is tied to conversion. For highly coking feeds, excess heat has to be removed. Such a cooler decouples the reactor from the regenerator and changes the control.

In addition to U_{cd}, modern FCCs have a range of slowly manipulated variables available, such as catalyst activity and other properties which are controlled by adding and withdrawing catalysts as well as additives that impact on Y_{pj}. [7] reviews how design changes in FCC developed and how they impact on FCC operation.

There is another aspect of design that changes the control of the FCC. Historically, most units operated at regenerator temperatures of 1180 to 1250°F with a CO_2/CO ratio between 1.2 to 1.5. This reduces the amount of heat evolved, allowing higher conversion. The CO is combusted in a CO boiler, but old units often simply emitted it to the atmosphere. Environmental laws forced operations of such units to change the operation to complete CO combustion and combustion promoters allowed to practically eliminate CO emission. Today many units operate in full CO combustion. This changes the control as can be seen from Fig. 3 where we plot T_{rgn} (regenerator bed temperature) versus air rate at constant T_{ris} (riser top temperature). For a given air rate and T_{ris} there are two steady states, one in partial and the other in complete CO combustion.

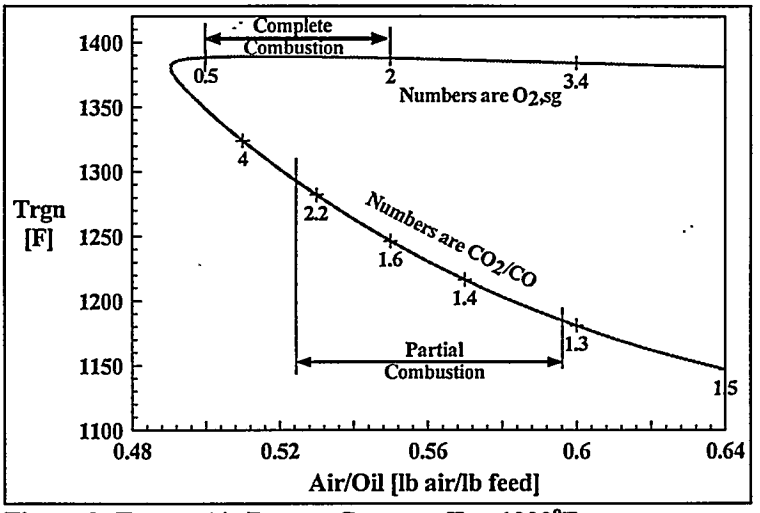


Figure 3: T_{rgn} vs. Air Rate at Constant T_{ris}=1000°F

Fig. 4 gives some options for control matrices and loops that have been used in the partial CO combustion regime. For complete combustion the only option is to use [F_{air} , $O_{2,sg}$] which is essential to keep the unit in complete CO combustion. Excess oxygen varies between 0.5 and 2.0% but aside from keeping the unit stable has no impact on either T_{rgn} or Y_{pj}.

If we look at Fig. 4, the measured variables entering the matrices in partial CO combustion are T_{ris}, T_{rgn}, T_{sg} (stack gas temperature in the regenerator) and ΔT (= T_{sg} - T_{rgn}) which is proportional to the excess oxygen entering the cyclone. While ΔT and T_{sg} have been used in the past and have been preferentially promoted in the control literature [8], they have two disadvantages over T_{rgn} and T_{ris}. They are not dominant and very model

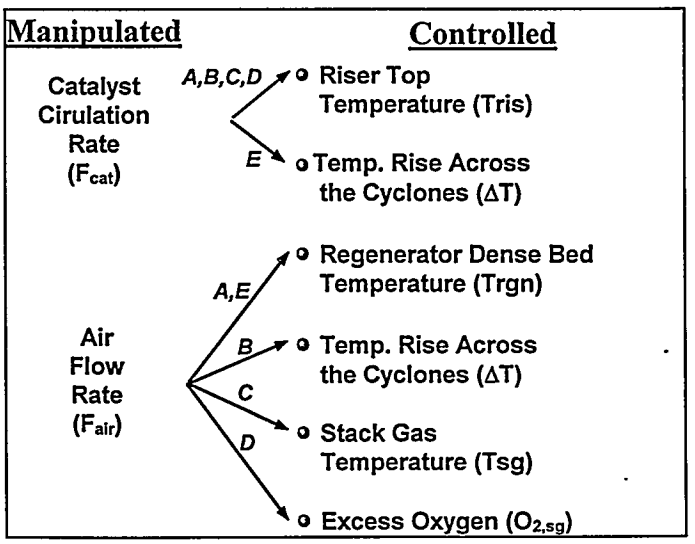


Figure 4: Possible Control Structures for the FCC

sensitive, and therefore not predictable from laboratory experiments. In the FCC the nature of the reactions are such that residence time has a small impact. However, reactor temperature impacts strongly on all reactions. Therefore, reactor outlet temperature T_{ris} as well as inlet temperature T_{mix} strongly impact on Y_{pj}. T_{mix} is a strong function of T_{rgn}. In the regenerator in partial combustion T_{rgn} determines the ratio of CO_2/CO as well as the rate of coke combustion and therefore C_{rgc} (coke on regenerated catalyst). T_{rgn} is therefore dominant. In full CO combustion C_{rgc} is practically zero and CO_2/CO close to infinity. Therefore T_{rgn} is not dominant as long as it is high enough.

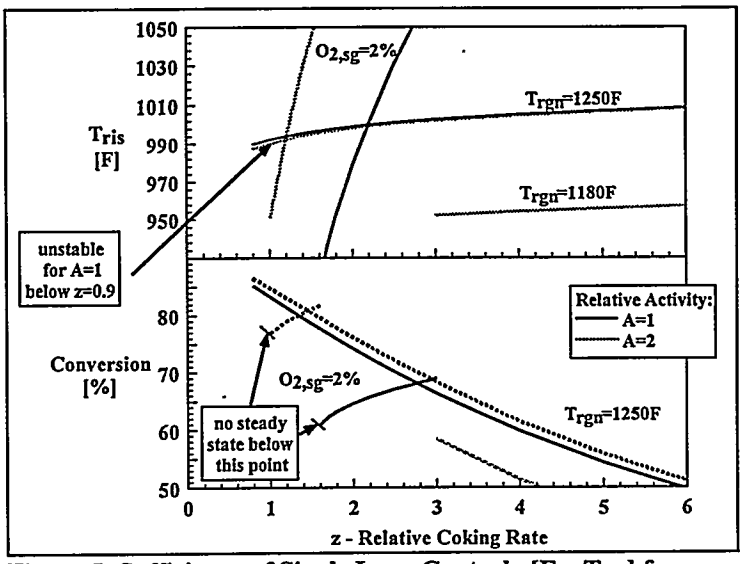


Figure 5: Sufficiency of Single Loop Control. $[F_{air},T_{rgn}]$ for partial combustion; $[F_{air},O_{2,sg}]$ for complete combustion.

The dominant variables are therefore T_{ris} and T_{rgn} in the reactor and T_{rgn} in the regenerator. They are also the only ones of the set that are modellable.

While a discussion of the variables requires far more space than we have, we can use this to illustrate some of the ideas. For example, in partial combustion it is perfectly possible to maintain a stable operation over a wide range of catalyst properties by one single control loop $[F_{air}, T_{rgn}]$. This can be seen from Fig. 5 where we plot T_{ris} vs. relative coking rate (z), as T_{ris} has narrow limits. It has to be maintained between 900 and 1010°F. Below 850°F part of the feed will not vaporize or crack and at about 1030°F thermal cracking becomes the predominant reaction. We note that the impact of F_{air} on T_{rgn} is small, so varying T_{rgn} allows one to keep T_{ris} within the permissible range. At low relative coking rate, z, the unit becomes unstable, but a single PI controller can stabilize it. At very low z it loses the steady state and the only option is a more active catalyst or one that produces more coke.

In full combustion a single control loop $[F_{air}, O_{2,sg}]$ does not allow stable operation. We note that small changes in z will change T_{ris} such that it falls outside the permissible range. The only option is to change catalyst activity. Controlling T_{ris} is here essential. The two loops are equivalent to a single loop in partial combustion.

While a single loop is sufficient to maintain the unit stable in partial combustion (or a 2x2 matrix in complete CO combustion), it is very limiting for controlling Y_{pj}. This is illustrated in Fig. 6 where we show some elements of Y_{pj} for the scheme $[(F_{air}, T_{rgn},), (F_{cat}, T_{ris})]$ for a specific operating point in partial combustion. We note that instead of a line an area of Y_{pj} becomes accessible. Adding slow manipulated variables (U_{csj}) will change the accessible space to a multi-dimensional volume. Similar plots can be made for complete combustion, but here we need three

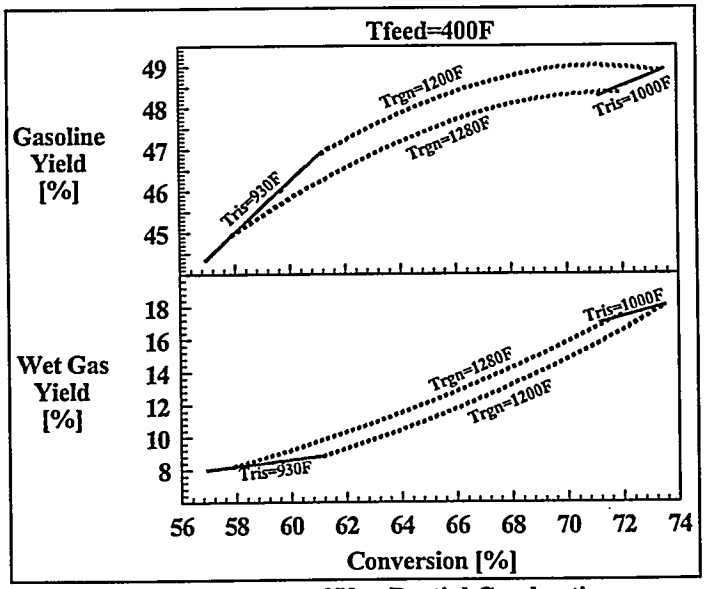


Figure 6: Accessible Area of Yp. Partial Combustion . Control: $[(F_{air},T_{rgn}),(F_{cat}, T_{ris})]$

manipulated variables to increase the accessible space of Y_{pj} from a line to an area.

In complete combustion there is an additional problem. Even three loops do not allow adequate control of T_{rgn} for higher coking rates. The only variable that allows control of T_{rgn} over a wide range of z is catalyst activity. It has a large impact on control in complete combustion whereas its effect is limited in partial combustion. These problems can be completely overcome by the addition of a catalyst cooler which allows T_{rgn} to be adjusted independently both in partial and complete combustion.

Our results show that in partial combustion, in addition to lack of modellability and dominance, ΔT (and therefore T_{sg}) has another problem. If ΔT and T_{ris} are kept constant we can show that increasing z could cause the unit to crash as at constant T_{ris} and ΔT increasing z will decrease T_{rgn} (Fig. 7). If T_{rgn} will drop below 1100°F the unit could lose its uppermost steady state. The same is true for the scheme [(F_{air}, T_{rgn}), (F_{cat}, ΔT)] proposed by [9]. In that case an increase in z will cause T_{ris} to decrease (Fig. 7). If T_{ris} goes below 850°F, the unit also loses its steady state. The exact limits strongly depend on the unit and the catalyst. Thus our results explain the phenomena of crashing due to coking feeds which plagued the industry until the control scheme [(F_{air}, T_{rgn}), (F_{cat}, T_{ris})] was introduced. Crashing occurred when a heavy coking feed was suddenly introduced and did not occur in units with well controlled constant feed composition.

This illustrates an important aspect of our results. Choice of the proper control matrix is in most cases far more important than the multi-variable algorithm itself. Further, while in the FCC linear algorithms are enough for taking care of wide changes in operating conditions and feed perturbations, one must chose the matrix based on nonlinear considerations. The understanding of the nonlinear feature of the system is essential.

One can not even design the linear algorithm itself based solely on the linear model for the matrix. One needs to understand the impact of the specifications. Thus, the matrix [(F_{air}, T_{rgn}), (F_{cat}, T_{ris})] has a negative relative gain array (RGA). In FCC's however one normally designs along the negative diagonal contrary to textbook recommendations. This gives a fast response to T_{ris} and allows tight control of it. As T_{ris} dominates Y_{pj} this advantage outweighs the penalty of operating along the negative diagonal which causes the system to become unstable when one loop opens. Another important aspect of our results that have significance to other systems is that the controllability of a system can totally change if changes occur in the system or operating conditions. This can often occur in reactors due to changes in catalyst properties or changes in Y_{pj} due to market requirements.

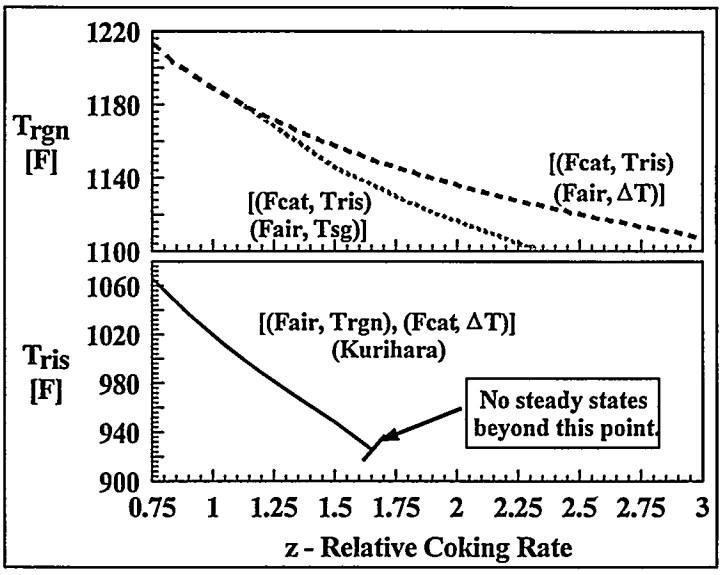


Figure 7: Impact of Coking Rate on Dominant Variables for Different Control Schemes.

Thus our results on the FCC give some important insights as to the design of partial control systems for complex nonlinear systems and raise important challenges for future research. They should also impact on the way we teach process control. The current work focuses on the impact of the design on the control and on the minimum information required to design and operate a successful partial control system. Future work will try to generalize the approach by dealing with other complex systems.

The results are published in a way that indicates that they are useful for a wide range of processes and should have a much wider impact than just FCC control. A paper on how the results impact on economic control was recently submitted for publication. For more details, consult references 1-4.

ACKNOWLEDGMENTS

The authors wish to acknowledge with gratitude the support of the Office of Basic Energy Sciences of the Department of Energy. This research was supported by grant number DE-FG02-91-ER14221.

REFERENCES

1. Arbel, A., Z. Huang, I.H. Rinard, and R. Shinnar, (1995) ,"Dynamics and Control of Fluidized Catalytic Crackers. 1. Modeling of the Current Generation of FCC's," *I&EC Research*, **34**,(4), 1228- 1243.
2. Arbel, A., I.H. Rinard, and R. Shinnar, (1995) , "Dynamics and Control of Fluidized Catalytic Crackers. 2. Multiple Steady States and Instabilities," *I&EC Research*, **34**(9), 3014- 3026.
3. Arbel, A., I.H. Rinard, and R. Shinnar, (1996) , "Dynamics and Control of Fluidized Catalytic Crackers. 3. Designing the Control System: Choice of Manipulated and Measured Variables for Partial Control," *I&EC Research* (accepted for publication).
4. Arbel, A., I.H. Rinard, and R. Shinnar, (1996) , "Dynamics and Control of Fluidized Catalytic Crackers. 4. The Impact of Design on Partial Control," *I&EC Research* (submitted for publication).
5. Shinnar, R. (1981), "Chemical Reactor Modeling for the Purposes of Controller Design," *Chem. Eng. Commun.*, **9**, 73-99.
6. Shinnar, R. (1986), "Impact of Model Uncertainties and Nonlinearities on Modern Controller Design. Present Status and Future Goals," In *Chemical Process Control-CPCIII, Proceedings of the 3rd International Conference on Chemical Process Control, M.* Morari and T.J. McAvoy, Eds., Elsevier, New York, 53-93.
7. Avidan, A. A.; Shinnar, R. Development of Catalytic Cracking Technology. A Lesson in Chemical Reactor Design. *Ind. Eng. Chem. Res.*, **1990**, *29*, 931-942.
8. Hovd, M.; Skogestad, S. Controllability Analysis for the Fluid Catalytic Cracking Process. *Ind. Eng. Chem. Res.*, **1992**, *31*, 1589.
9. Kurihara, H. Optimal Control of Fluid Catalytic Cracking Process. Ph.D. Dissertation, MIT, **1967**

NON-LINEAR SYSTEM IDENTIFICATION IN FLOW-INDUCED VIBRATION

P.D. Spanos and B.A. Zeldin

Department of Mechanical Engineering
Rice University, Houston, Texas 77251

R. Lu

Hudson Engineering Corporation
Houston, Texas 77218

ABSTRACT

The paper introduces a method of identification of non-linear systems encountered in marine engineering applications. The non-linearity is accounted for by a combination of linear subsystems and known zero-memory non-linear transformations; an equivalent linear multi-input-single-output (MISO) system is developed for the identification problem. The unknown transfer functions of the MISO system are identified by assembling a system of linear equations in the frequency domain. This system is solved by performing the Cholesky decomposition of a related matrix. It is shown that the proposed identification method can be interpreted as a "Gram-Schmidt" type of orthogonal decomposition of the input-output quantities of the equivalent MISO system. A numerical example involving the identification of unknown parameters of flow (ocean wave) induced forces on offshore structures elucidates the applicability of the proposed method

INTRODUCTION

Mathematical modeling of wave-induced forces on offshore structures is a complex technical problem. Conventionally, the wave forces on structural elements of size that is small compared to the characteristic wavelength are calculated by Morison's equation [1]. Selecting parameters for the Morison's model which are appropriate for the conditions of a specific site is a delicate procedure. In this context, it is noted that it is rather difficult to account for the effects of waterline, body motion, velocity head, and second-order potential. Thus, the validity of the Morison's force model and its applicability for describing wave forces on flexible offshore structures have been questioned in the literature. In this regard, this paper addresses the issue of determining reliable models of non-linear forces acting on offshore structures by using system identification techniques in conjunction with possibly available experimental data.

Although the identification of linear systems is a well developed subject and is extensively used in engineering practice [2], identification of non-linear structural systems and forces is a less understood scientific area. Several alternative approaches have been proposed in the literature in

conjunction with structural dynamics applications. Non-linear systems can be identified by using the Volterra/Wiener representation theory in which the kernels of the associated integrals are determined from the governing differential equations [3]. In Refs. [4] and [5] the non-linear restoring force was expanded by using a polynomial basis and the coefficients of this expansion were found by minimizing the discrepancy between the observed and simulated responses. By representing the non-linear restoring forces as a set of parallel linear subsystems, the identification of a non-linear system can be developed by adopting a multi-input-single-output (MISO) linear system analysis procedure [6,7]. Using this approach, Vugts and Bouquet [8] verified the accuracy of Morison's equation based on realistic measurement data. Utilizing the representation of a non-linear structural system as a collection of linear subsystems, Spanos and Lu [9] addressed the non-linearity induced by the structure-environment interaction in marine applications; pertinent examples demonstrated the validity of the identification method. Recently, Zeldin and Spanos [10] proposed to use a parametric ARMAX procedure to identify the introduced MISO linear systems. The latter procedure can be especially useful for identification of transfer functions exhibiting sharp peaks; this feature is quite common for structural systems in marine applications.

In this paper, a new procedure is developed for the identification of non-linear models of structural systems for offshore engineering applications. Specifically, the non-linear restoring force is represented as a set of known zero-memory non-linear transformations that are combined with linear subsystems; it reduces a nonlinear system to a linear MISO system [6,7,9]. This method is particularly efficient if preliminary studies of the non-linear system may be used to speculate with regards to the mathematical form of the non-linear force. The identification of the equivalent MISO system is performed by assembling a system of linear equations in the frequency domain with respect to the unknown linear transfer functions of the MISO system. This system is solved by performing the Cholesky decomposition of the associated matrix. The developed identification method can be viewed as a "Gram-Schmidt" type of orthogonal decomposition of the input-output quantities of the equivalent MISO system examined in Refs. [6,7,9]. Pertinent numerical calculations elucidate the implementation of the proposed identification method for marine engineering applications.

MODELING OF NON-LINEAR FORCES

Consider the differential equation

$$m\ddot{q} + c\dot{q} + kq = f(u, \ldots, q, \ldots), \tag{1}$$

where $q(t)$ is the scalar variable representing the structural motion, $u(t)$ is the variable describing the wave kinematics, and f is the random force function which represents the wave induced non-linear force. Further, m, c, and k are the structural mass, damping, and stiffness of the structure, respectively; the dot signifies differentiation with respect to time. To find the unknown structural parameters of Eq. (1), the non-linear force is expanded in terms of base functions of the wave and structural kinematic parameters. In this regard, it is assumed that the system and the excitation nonlinearities can be expressed as a superposition of zero-memory non-linear transformations and linear subsystems. Specifically,

$$f(u, \ldots, q, \ldots) = \sum_{k=1}^{M} H_k\left(\frac{d}{dt}\right) f_k(u, \ldots, q, \ldots), \tag{2}$$

where $H_k(s)$ are polynomials, $f_k(u, \ldots, q, \ldots)$ are non-linear zero-memory transformations, and M is the total number of base functions used in the representation of the non-linear excitation force.

Some preliminary studies may be used to speculate with regards to the mathematical form of the non-linear transformations of Eq. (2). For example, for a Duffing oscillator under random excitation $u(t)$, the non-linear excitation force can be expressed as

$$f(u, \ldots, q, \ldots) = -\varepsilon q^3 + u \tag{3}$$

and the unknown non-linear parameter ε can be determined from measurements by using the proposed identification procedure. For an offshore structure with small structural elements, the non-linear excitation force induced by waves may be represented by Morison's equation. Specifically,

$$f(u, \ldots, q, \ldots) = \alpha_1 |u - \dot{q}|(u - \dot{q}) + \alpha_2 m_a(\dot{u} - \ddot{q}) + m_a \ddot{q}, \tag{4}$$

where α_1 and α_2 are unknown drag and inertia coefficients, respectively. Alternatively, a polynomial expansion of the drag force can be used instead of Morison's equation. That is, the non-linear excitation force can be expressed as

$$f(u, \ldots, q, \ldots) = \sum_{j=0}^{l} \sum_{i=0}^{l-j} \alpha_{ij} u^i \dot{q}^j + \alpha_2 m_a(\dot{u} - \ddot{q}) + m_a \ddot{q}. \tag{5}$$

In this case, the non-linear force is described by $M = (l+1)(l+2)/2 + 2$ parameters.

EQUIVALENT MISO SYSTEM

By substituting Eq. (2) into Eq. (1), the non-linear equation of motion of the structural system can be expressed as

$$H_0\left(\frac{d}{dt}\right) q = \sum_{k=1}^{m} H_k\left(\frac{d}{dt}\right) f_k(u, \ldots, q, \ldots), \tag{6}$$

where $H_0(s) = ms^2 + cs + k$. The proposed identification method rearranges the terms of Eq. (6) by treating them as the input/output quantities of an equivalent MISO linear system. Specifically, the terms q and $f_k(u, \ldots, q, \ldots)$ of Eq. (6) that are associated with the unknown structural parameters are interpreted as the inputs x_k of the MISO system. The remaining terms of Eq. (6) are combined into the system output $y(t)$. Note that this study uses the terms "input" and "output" (x and y) in the context of system analysis terminology, whereas "excitation" and "response" (u and q) denote physically motivated structural analysis terms. The composed MISO system is described by the equation

$$\sum_{k=1}^{n} H_k\left(\frac{d}{dt}\right) x_k(t) = y(t), \tag{7}$$

where n denotes the total number of the input variables used in the equivalent MISO representation of the non-linear structural system.

The time domain description of the MISO system of Eq. (7) leads a frequency domain representation, as well. Specifically, using Fourier transform in Eq. (7) yields

$$\sum_{k=1}^{n} H_k(i\omega) X_k(\omega) + N(\omega) = Y(\omega), \tag{8}$$

where the capital letters denote the Fourier transforms

$$X_k(\omega) = F\{x_k(t)\}, \qquad Y(\omega) = F\{y(t)\}, \tag{9}$$

and the symbol $N(\omega)$ is introduced to represent the extraneous noise. The MISO system with the governing equation (8) is schematically shown in Figure 1.

IDENTIFICATION OF THE TRANSFER FUNCTIONS OF THE MISO SYSTEM

It is assumed that the excitation and the response of the non-linear system are stationary stochastic processes. Correspondingly, the input-output variates of the equivalent MISO linear system are also stationary processes; their correlation is described by the cross-correlation function

$$R_{x_i x_j}(\tau) = E[x_i(t) x_j(t+\tau)] \tag{10}$$

or by the cross-spectral density function

$$S_{x_i x_j}(\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} R_{x_i x_j}(\tau) e^{-i\omega\tau} d\tau. \tag{11}$$

In Eq. (10) $E[\circ]$ denotes the operator of mathematical expectation. Alternatively, the cross-spectral density function for stationary processes $x_i(t)$ and $x_j(t)$ can be defined as

$$S_{x_i x_j}(\omega) = \frac{2\pi}{T} E[X_i^*(\omega) X_j(\omega)], \tag{12}$$

where T denotes the length of observation of the processes $x_i(t)$, and $(\ )^*$ is used to denote complex conjugation. It can be shown that the definitions of Eqs. (11) and (12) are identical in the limit as T tends to infinity. In this context, they can be considered as equivalent for an adequately large, but finite, value of T.

In conjunction with the preceding comments, Eq. (8) governing the input-output relationship of the MISO system can be rewritten in the equivalent form

$$\{\underline{X}^t(\omega), Y(\omega)\}\begin{Bmatrix} \underline{H}(i\omega) \\ -1 \end{Bmatrix} = -N(\omega), \tag{13}$$

where $\underline{X} = \{X_1, \ldots X_n\}^t$, $\underline{H} = \{H_1, \ldots H_n\}^t$, and $(\ )^t$ denotes the transposition of the corresponding vectors. Premultiplying Eq. (13) by $2\pi(\{\underline{X}^t(\omega), Y(\omega)\}^*)^t/T$ and taking mathematical expectation yield

$$\begin{bmatrix} S_{\underline{x}\underline{x}} & S_{\underline{x}y} \\ S_{y\underline{x}} & S_{yy} \end{bmatrix}\begin{Bmatrix} \underline{H} \\ -1 \end{Bmatrix} = \begin{Bmatrix} 0 \\ -S_{yN} \end{Bmatrix}, \tag{14}$$

where it is assumed that the noise process $N(\omega)$ is uncorrelated from the inputs of the MISO system. Further, it follows from Eq. (14) that if the observation noise is zero and all non-linear effects are properly captured by the MISO system, then the vector $\{\underline{H}^t, -1\}^t$ is the eigenvector of the spectral matrix of the vector-process $\{\underline{X}^t(\omega), Y(\omega)\}^t$ corresponding to the zero eigenvalue.

The system of equations (14) can be readily used to determine the unknown transfer functions of the MISO system. Indeed, the first n equations of the system can be used to find the vector of the transfer functions $\underline{H}(i\omega)$. The last equation can be utilized to determine the value of S_{yN} which represents the influence of the measurement noise and the non-linear terms which were not included into the MISO system. In this regard it is pointed out that the value S_{yN} is equal to the spectral density of the noise. Indeed, by squaring Eq. (13) one can derive

$$S_{NN}(\omega) = \frac{2\pi}{T}\{(\underline{H}^t)^*, -1\}E\left[\begin{Bmatrix} \underline{X}^* \\ Y^* \end{Bmatrix}\{\underline{X}^t, Y\}\right]\begin{Bmatrix} \underline{H} \\ -1 \end{Bmatrix} = \{(\underline{H}^t)^*, -1\}\begin{bmatrix} S_{\underline{x}\underline{x}} & S_{\underline{x}y} \\ S_{y\underline{x}} & S_{yy} \end{bmatrix}\begin{Bmatrix} \underline{H} \\ -1 \end{Bmatrix} =$$

$$\{(\underline{H}^t)^*, -1\}\begin{Bmatrix} 0 \\ -S_{yN} \end{Bmatrix} = S_{yN}. \tag{15}$$

Upon evaluating the spectral density function S_{yN}, the cumulative coherence function [6,9] can be evaluated from the equation

$$\gamma_{y:\underline{x}}(\omega) = 1 - \frac{S_{NN}(\omega)}{S_{yy}(\omega)} = \frac{S_{yy}(\omega) - S_{yN}}{S_{yy}(\omega)}. \tag{16}$$

The coherence function $\gamma_{y:\underline{x}}(\omega)$ expresses the cumulative linear contribution of the random vector $\underline{x}(t)$ to the MISO system output $y(t)$. This coherence function reflects, in essence, the "goodness" of the selected model in describing the physical system.

In context with the preceding development, the Gaussian elimination procedure applied to the system of equations (14) provides a useful interpretation of the proposed method for numerical evaluation of $\underline{H}(i\omega)$. Note that the Gaussian elimination procedure transforms the matrix of Eq. (14) to an upper triangular matrix. At the step j of this procedure the components of the matrix located on the j-th column below the main diagonal are eliminated. In this regard, every step of the

Gaussian elimination is equivalent to a premultiplication of Eq. (14) by a lower triangular matrix with non-zero components located on the main diagonal and on the j-th column. Thus, during the first step of the Gaussian elimination, Eq. (14) is premultiplied by the matrix

$$A_1 = \begin{bmatrix} 1 & 0 & \dots & 0 \\ -S_{x_2x_1}/S_{x_1x_1} & 1 & & 0 \\ \dots & & & \dots \\ -S_{yx_1}/S_{x_1x_1} & 0 & \dots & 1 \end{bmatrix}. \tag{17}$$

This yields the system of equations

$$\begin{bmatrix} S_{x_1x_1} & S_{x_1x_2} & \dots & S_{x_1y} \\ 0 & S_{x_2x_2;1} & & S_{x_2y;1} \\ \dots & & & \dots \\ 0 & S_{yx_2;1} & \dots & S_{yy;1} \end{bmatrix} \begin{Bmatrix} H_1 \\ H_2 \\ \dots \\ -1 \end{Bmatrix} = \begin{Bmatrix} 0 \\ -S_{yN} \end{Bmatrix}, \tag{18}$$

where $S_{x_kx_j;1} = S_{x_kx_j} - S_{x_kx_1}S_{x_1x_j}/S_{x_1x_1}$. In this manner, during the j-th step of the Gaussian elimination, the system of equation obtained at the end of the previous $(j-1)$ step is premultiplied by the matrix

$$A_j = \begin{bmatrix} 1 & 0 & \dots & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & & \dots \\ 0 & 0 & & 1 & & 0 \\ 0 & 0 & \dots & -S_{x_{j+1}x_j;j-1}/S_{x_jx_j;j-1} & \dots & 0 \\ \dots & \dots & & \dots & & \dots \\ 0 & 0 & \dots & -S_{yx_j;j-1}/S_{x_jx_j;j-1} & \dots & 1 \end{bmatrix}. \tag{19}$$

Then, the Gaussian elimination procedure can be formalized as a premultiplication of Eq. (14) by the matrix

$$A = A_nA_{n-1}\dots A_1. \tag{20}$$

Since all matrices A_j in Eq. (20) are lower triangular, the assembled matrix A is also lower triangular. Also, note that all elements of the main diagonal of the matrix A are equal to 1. The system of equation obtained after completion of the Gaussian elimination can be written as

$$\tilde{S}\begin{Bmatrix} \underline{H} \\ -1 \end{Bmatrix} = \begin{Bmatrix} 0 \\ -S_{yN} \end{Bmatrix}, \tag{21}$$

where $\tilde{S}$ is the upper triangular matrix.

In this regard it is pointed out that Eq. (21) can be obtained directly from Eq. (13) by pre multiplying it by the vector $2\pi(\{\underline{\tilde{X}}^t, \tilde{Y}\}^t)^*/T = 2\pi(A^*\{\underline{X}^t, Y\}^t)^*/T$ and taking mathematical expectation. Note that the matrix S of Eq. (21) can be viewed as the cross-spectral density matrix of the vector-process $\{\underline{\tilde{x}}^t, \tilde{y}\}^t$ and the vector-process $\{\underline{x}^t, y\}^t$. Therefore, it can be concluded from the triangular form of S that the process $\tilde{x}_l$ is uncorrelated from any of the processes x_j for $j < l$, and the process $\tilde{y}$ is uncorrelated from the vector-process $\underline{x}$. Further, the spectral density matrix of the vector-process $\{\underline{\tilde{x}}^t, \tilde{y}\}^t$ can be found from the equation

$$S_{\{\underline{\tilde{x}}^t, \tilde{y}\}^t} = \frac{2\pi}{T}E\left[\begin{Bmatrix} \underline{\tilde{X}} \\ \tilde{Y} \end{Bmatrix}^* \{\underline{\tilde{X}}^t, \tilde{Y}\}\right] = A\begin{bmatrix} S_{\underline{x}\underline{x}} & S_{\underline{x}y} \\ S_{y\underline{x}} & S_{yy} \end{bmatrix}(A^t)^* = \tilde{S}(A^t)^*. \tag{22}$$

Since both matrices $\tilde{S}$ and $(A^t)^*$ are upper triangular, the matrix $S_{\{\tilde{x}^t, \tilde{y}\}^t}$ is also upper triangular. Moreover, the matrix $S_{\{\tilde{x}^t, \tilde{y}\}^t}$ is a symmetric spectral density matrix of a random vector; it is a diagonal matrix. That is, the random vector-process $\{\tilde{x}^t, \tilde{y}\}^t$ is a vector-process with uncorrelated components. Also, note that since the elements of the main diagonal of the matrix $(A^t)^*$ are all equal to 1, the main diagonal of the matrix $S_{\{\tilde{x}^t, \tilde{y}\}^t}$ is identical to the main diagonal of the matrix $\tilde{S}$. That is, upon implementing Gaussian elimination on Eq. (14), the elements of the main diagonal of the obtained matrix S represent the spectral density functions of the corresponding orthogonalized processes.

In conjunction with the preceding developments, it is pointed out that the Gaussian elimination procedure of the system of equations (14) is equivalent to a "Gram-Schmidt" type of orthogonal decomposition of the random vector-process $\{\underline{x}^t, y\}$; this approach was previously pursued in Refs. [6,7,8,9] for identification of the MISO system transfer functions $H_j(i\omega)$.

Finally, it is noted that the matrix A can be determined in a more efficient manner, from a computational standpoint, by using the Cholesky decomposition of the matrix of Eq. (14). Indeed, Eq. (22) yields

$$\begin{bmatrix} S_{\underline{x}\underline{x}} & S_{\underline{x}y} \\ S_{y\underline{x}} & S_{yy} \end{bmatrix} = \left(A^{-1}\sqrt{S_{\{\tilde{x}^t, \tilde{y}\}^t}}\right)\left(\sqrt{S_{\{\tilde{x}^t, \tilde{y}\}^t}}(A^t)^{-*}\right) = LL^{t*}, \tag{23}$$

where L is the lower triangular matrix of the Cholesky decomposition. Since the Cholesky decomposition is unique, one finds

$$L = A^{-1}\sqrt{S_{\{\tilde{x}^t, \tilde{y}\}^t}}. \tag{24}$$

Note that the elements of the main diagonal of the matrix L represent the square root of the spectral density functions of the corresponding components of the orthogonalized vector-process $\{\tilde{x}^t, \tilde{y}\}^t$.

Upon evaluating the transfer functions of the equivalent MISO system by solving equation (14) with the use of the Cholesky decomposition procedure, the unknown structural parameters of the non-linear model can be obtained by using a curve-fitting procedure over the frequency range of interest. Several computational issues of the developed identification method warrant additional remarks. First, the introduced equivalent MISO and the ordering of its components must be selected based on a physical model which makes use of the available knowledge regarding the non-linear structural behavior. A physically motivated model can provide mathematical forms for the base functions of Eq. (2), or indicate the most expeditious expansion of the non-linear excitation force into a series of functions in a statistically equivalent sense. Note that the components that better describe the system physics correspond to the higher value of the coherence function and should be assigned to the first elements of the vector $\underline{x}$. Second, the proposed identification method can be sensitive to the error of estimating the power spectral density functions. It is a common practice in linear spectral analysis to use an appropriate window to reduce the effect of leakage in spectral estimation. Further, overlapping and zooming can be applied to the estimation process to recover the information lost during windowing [11]. Note that a biased estimation of the corresponding spectral density functions and inadequate MISO system representation may yield a matrix $S_{\underline{x}\underline{x}}(\omega)$ which is not positive definite for certain frequencies. Finally, it should be noted that noise may be induced during the measurement. The proposed method relies on the assumption, which is quite reasonable in most circumstances, that the noise is statistically independent from the structural excitation and response quantities; filtering techniques must be applied to reduce this noise as much as possible.

NUMERICAL EXAMPLE

To illustrate the application of the proposed method, a single-degree-of-freedom system excited by a wave force is considered. The wave force is modeled by Morison's equation. Thus, the

motion of the system is governed by the equation

$$m\ddot{q} + c\dot{q} + kq = \alpha_1|u-\dot{q}|(u-\dot{q}) + \alpha_2 m_a(\dot{u}-\ddot{q}) + m_a\ddot{q}. \quad (25)$$

The numerical values $m = 100$, $c = 1$, $k = 0.25$, $m_a = 10$, $\alpha_1 = 0.9$, and $\alpha_2 = 1.25$ are used. This structure has system natural frequency $\omega_n = \sqrt{k/m} = 0.05$, and damping ratio $\zeta = 0.1$. The Gaussian surface wave elevation is generated consistently with the Pierson-Moscowitz spectrum [9] corresponding to a wind speed of 25.74 m/s, using a superposition of sinusoids at 1000 frequency points uniformly spaced in the interval [0, 1.5] rad/sec. The horizontal wave velocity and acceleration at a depth of 10m in a "deepwater" location are computed using linear wave theory. A steady current of 0.5 m/s is superimposed on the velocity at the same location. Note that the assumption of an infinite water depth and of linear wave theory are adopted for simplicity and the proposed identification algorithm does not hinge upon these assumptions. Figure 2 shows the time histories, one-sided power spectra, and probability density functions of the water particle velocity and the displacement response. The low frequency response and the non-Gaussian distribution of the response are clearly discernible.

For the considered example the values of the parameters α_1 and α_2 are assumed to be unknown and are identified based on the simulated data. The base functions are selected as $|u-\dot{q}|(u-\dot{q})$ and $\dot{u}-\ddot{q}$. The following decomposed system is derived: $x_1 = |u-\dot{q}|(u-\dot{q})$, $x_2 = \dot{u}-\ddot{q}$, $y = (m-m_a)\ddot{q} + c\dot{q} + kq$, $H_1(i\omega) = \alpha_1$, and $H_2(i\omega) = \alpha_2 m_a$. Results of identification are shown in Figure 3. The transfer functions $H_1(i\omega) = \alpha_1$ and $H_2(i\omega) = \alpha_2 m_a$ are found to be constant; the parameters are obtained as $\hat{\alpha}_1 = Real(H_1(\omega))$ and $\hat{\alpha}_2 = Real(H_2(\omega))/m_a$, accordingly. The broken lines represent the true parameters, α_1 and α_2. The averaged parameters are $\hat{\alpha}_1 = 0.896$ and $\hat{\alpha}_2 = 1.231$. Note that beyond the cutoff frequency ($\omega = 1.5$ rad/sec) the identified values are of no interest. Shown in Figure 4 are the cumulative coherence functions associated with the SISO system having the input x_1 and with the MISO system with the inputs x_1 and x_2. Full frequency range is displayed, although only the range of [0, 1.5] rad/sec is to be observed. The coherence functions in Figure 4 provide a measure of the validity of the decomposed system and indicates the confidence with respect to the identified transfer functions. Indeed, a quantitative error analysis can be pursued to relate the confidence limit of an identified parameter to the coherence of its associated transfer functions. Also, it was found that the developed identification method is quite robust, with respect to adding noises to the observed excitation and response processes, within the frequency range containing the frequencies which are significant for describing waves and structural behavior.

CONCLUDING REMARKS

Method of identification of non-linear structural systems has been developed. It is based on the decomposition of system nonlinearities in terms of base functions, and on subsequent treatment of these base functions as input-output variables of an equivalent MISO linear system. The transfer functions of the MISO are determined by solving a system of frequency domain equations with the use of the Cholesky decomposition. The numerical example involving identification of the parameters of a non-linear wave force on the offshore structure demonstrates the applicability and usefulness of the proposed identification method.

ACKNOWLEDGEMENT

The financial support of this work by the grant DE-FG03-95ER14500 from the Department of Energy is gratefully acknowledged.

REFERENCES

1. J.R. MORISON, M.P. O'BRIEN, and J.W. JOHNSON, "The Force Exerted by Surface Waves on Piles," *AIME Petroleum Transactions*, 189, 149 (1950).
2. F. KOZIN and H.G. NATKE, "System Identification Techniques", *Structural Safety*, 3, 269, (1986).
3. S.J. GIFFORD and G.R. TOMLINSON, "Recent Advances in the Application of Functional Series to Nonlinear Structures", *Journal of Sound and Vibration*, 135, 289, (1989).
4. S.F. MASRI, H. SASSI, and T.K. CAUGHEY, "Nonparametric Identification of Nearly Arbitrary Nonlinear Systems", *ASME Journal of Applied Mechanics*, 49, 619, (1982).
5. K.S. MOHAMMAD, K. WORDEN, and G.R. TOMLINSON,. "Direct Parameter Estimation for Linear and Nonlinear Structures", *Journal of sound and Vibration*, 138, 471, (1992).
6. J.S. BENDAT, *Random Data: Analysis and Measurement Procedures*, Wiley-Interscience, NY. (1990)
7. J.S. BENDAT, P.A. PALO, and R.N.COPPOLINA, "A General Identification Technique for Nonlinear Differential Equations of Motion", *Probabilistic Engineering Mechanics*, 7, 43, (1992).
8. J.H. VUGTS and A.G. BOUQUET, "A Nonlinear Frequency Domain Description of Wave Forces on an Element of a Vertical Pile in Random Seas," *BOSS*, London, U.K., 693, (1985).
9. P.D. SPANOS and R. LU, "Nonlinear System Identification in Offshore Structure Reliability", *Journal of Offshore Mechanics and Arctic Engineering*, 117, 171, (1995).
10. B.A. ZELDIN, and P.D. SPANOS, "Identification of Nonlinear Systems under Random Excitation," *Proceedings of the ASCE Engineering Mechanics Conference*, Fort Launderdale, May 19-22, (1996).
11. A.V. OPPENHEIM and R.W. SHAFER, *Digital Signal Processing*, Prentice-Hall, NJ, (1975).

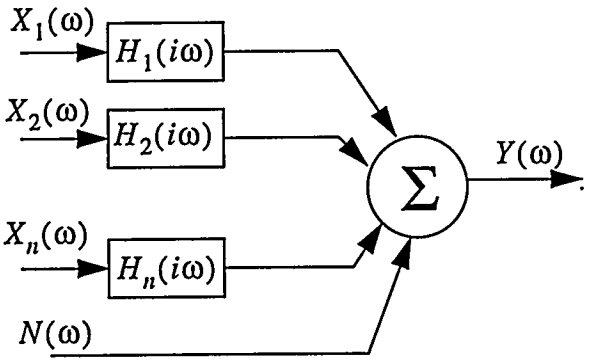


Figure 1. Equivalent MISO system

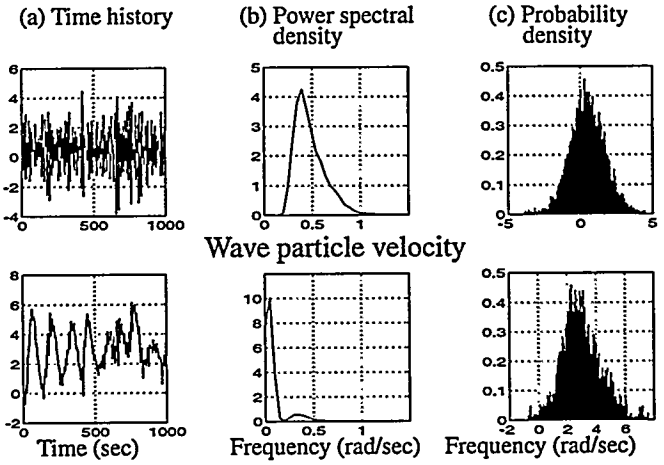


Figure 2. Simulated water particle velocity and structural displacement

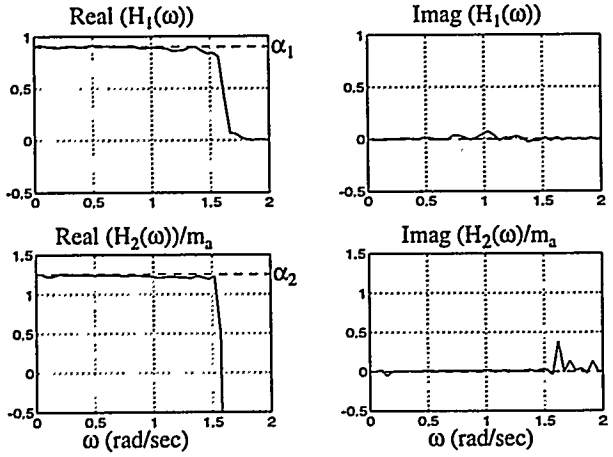


Figure 3. Identified linear transfer functions

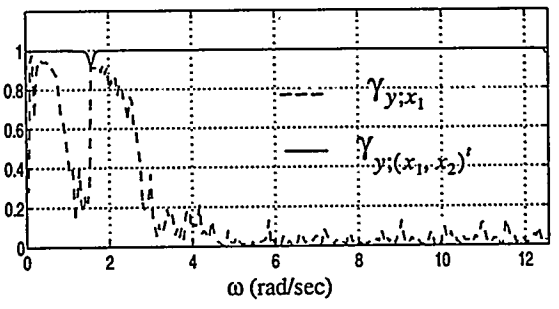


Figure 4. Cumulative coherence functions