EXPERIMENTS ON SCALAR MIXING AND TRANSPORT

Z. Warhaft

Sibley School of Mechanical and Aerospace Engineering
Ithaca, New York 14850, U.S.A.

ABSTRACT

We provide an overview of our recent work on passive (temperature) scalar mixing in both homogeneous and inhomogeneous turbulent flows. We show that for homogeneous grid generated turbulence, in the presence of a linear temperature profile, the probability density function (pdf) of the temperature fluctuations has broad exponential tails, while the pdf of velocity is Gaussian. However in the absence of a scalar gradient the pdf of temperature is Gaussian. This new result sheds insight into the fundamentals of turbulent mixing as well as to the nature of the velocity field. We also show that the spectrum of the temperature fluctuations has a scaling region that is consistent with Kolmogorov scaling although a similar scaling region is absent for the velocity field in this low Reynolds number flow. Finally, we describe our results concerning the mixing and dispersion of scalars in a jet. We show that although initially the scalar mixing is strongly dependent on input conditions, the mixing is rapid and that the correlation coefficient asymptotes to unity by x/D ~ 20.

INTRODUCTION

The understanding of scalar mixing and transport in turbulent flows remains a vital issue because of its fundamental importance, both in its own right and also in the way it sheds light on the basic characteristics of the velocity field itself, and because of its obvious practical significance in combustors, chemical mixers and the environment. At the fundamental level, a full statistical description has not emerged, even for simple flows although very recently much progress has occurred in our understanding of the nature of the probability density function (pdf) of the scalar fluctuations. Our contribution to this will be described in part 1 below. There are, however, still complex problems concerning one of the oldest and most studied statistical descriptions; the one dimensional scalar spectrum. Both experiments (e.g. Warhaft and Lumley 1978) and computation, (Metais and Lesieur 1992) show anomalous scaling regions in the scalar spectrum in isotropic turbulence. They are anomalous since they occur in the absence of such regions in the velocity spectrum. Although the experimental observations are quite old there has been no systematic study of the scalar spectrum as a function of Reynolds number and thus it has not been possible to determine whether these scaling regions are artifacts of the initial conditions in these low Reynolds number flows or whether they are fundamental to a statistical description of the flow. Our recent experiments towards an understanding of this problem will be described in part 2. We will show there is indeed a scaling region of constant slope close to -5/3 and its width increases systematically with Reynolds number. We also continue to be interested in the effect of initial conditions on the rate of mixing of scalar fluctuations. We have recently been studying scalar dispersion from heated wires in a jet. We have examined both single and two scalar mixing. Here we will describe our results and relate them to mixing in grid turbulence.

Thus our paper is concerned with three distinct topics: the pdf of passive scalar fluctuations, the spectrum of scalar fluctuations, and mixing and transport in a jet. Necessarily, because of space, only a brief summary can be provided and the reader is referred to Jayesh and Warhaft (1992, 1993), Warhaft (1992) and Tong and Warhaft (1993) for details, including descriptions of experimental apparatus.

1. THE PROBABILITY DENSITY FUNCTION (PDF) OF A PASSIVE SCALAR IN GRID TURBULENCE

Until our recent study (Jayesh and Warhaft 1991, 1992) there appeared to be no published experimental data on the details of the passive scalar pdf in homogeneous isotropic turbulence, particularly its tails that describe the higher-order moments. Possibly, this is due to it having been assumed that the scalar fluctuations are purely Gaussian, reflecting the velocity field, which early on was shown to have a Gaussian pdf, at least up to the fourth moment. Perhaps more pertinently, there has been no theory (until recently, see below) that suggested universality in the tails of the scalar pdf and thus experimental motivation has been lacking.

The principal impetus for our study came from the theory of Pumir, Shraiman and Siggia (1991). They argued, using a one-dimensional phenomenalogical model for a passive scalar advected by turbulence, that in the presence of a mean scalar gradient, the scalar pdf will have exponential tails but in the absence of the gradient (i.e. with uniform mean temperature) the scalar will have a Gaussian pdf. The two techniques we have developed over the years (the *mandoline*, Warhaft and Lumley (1978) and the *toaster* Sirivat and Warhaft 1983) were ideal to test their theory since the *mandoline* provides temperature fluctuations without a mean temperature gradient while the *toaster* can produce a linear temperature gradient. In our experiment (Jayesh and Warhaft 1991, 1992) we systematically varied the Reynolds number and other flow parameters.

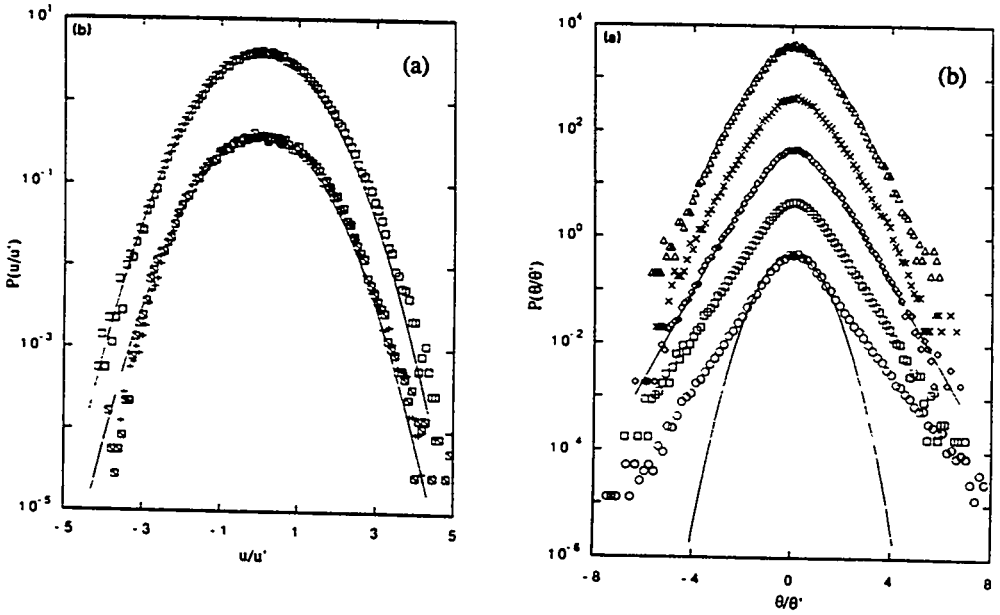


Figure 1 a) Pdf's of longitudinal velocity, u. Lower curves are at x/M = 4, upper curve is at x/M = 62.4. The solid line is Gaussian. The upper curve has been shifted one decade with respect to the lower one. b) Pdf of temperature at various x/M U=8.9 m/sec, β=6.06 K/m. M=2.5 cm. O:x/M=36.4;□ :x/M=62.4;◇ :x/M=82.4;✕ x/M=102.4;△ x/M=132.4. Each curve has been shifted one decade with respect to the lower one. A Gaussian curve is shown at x/M=36.4 and a straight line fit to the tails ($|\theta/\theta'|>2$) is shown at x/M=82.4. Here U is the mean speed, β is the temperature gradient and M is the mesh length. The pdf's have been normalized by the tempeature rms, θ'. Reproduced from Physics of Fluids A.

Figure 1 shows the velocity pdf and the temperature pdf in the presence of a mean scalar gradient. While the velocity pdf is Gaussian deep into its tails (note the logarithmic plot) the scalar pdf is distinctly non-Gaussian, showing exponential tails (linear on the log plot). Our study has shown that these tails weaken slightly with downstream distance (Figure 1b) but always remain broader than Gaussian. They were observed in the integral scale Reynolds number, Re_l, range $60 < Re_l < 1100$. On the other hand, in the absence of a mean gradient, the scalar pdf is close to Gaussian (Jayesh and Warhaft 1992). The qualitative difference between the gradient and no gradient case appears to provide confirmation of the Pumir, Shraiman, Siggia theory and, for the gradient case are consistent with recent work of Gollub et al. (1991) and Kerstein (1991).

We have also studied other statistics such as the conditional scalar dissipation rate, the pdf of the temperature derivative and the effect of filtering on the scalar pdf. These are described in Jayesh and Warhaft (1992).

Our findings should have particular significance in the fields of turbulent mixing and combustion since they show, for the linear profile case, that enhanced thermal dissipation occurs in the presence of the large, rare temperature fluctuations that are responsible for the extended tails of the pdf. Thus more rapid smearing (mixing at the molecular level) will occur, enhancing reaction and combustion rates.

2. TEMPERATURE SPECTRA IN GRID TURBULENCE

The Corrsin-Obukhov (Corrsin 1957, Obukhov 1949) extension of the Kolmogorov (1941) similarity theory shows that for high Reynolds numbers the spectrum of a passive scalar in the inertial subrange has the form

$$E_\theta(k)=\beta\varepsilon^{-1/3}\varepsilon_\theta k^{-5/3} \qquad (1)$$

Here $E_\theta(k)$ is the one dimensional spectrum defined by $\overline{\theta^2} = \int E_\theta(k)dk$ where $\overline{\theta^2}$ is the scalar variance and k is the wave number in the x direction; β is a universal constant and ε and ε_θ are the average dissipation rate of energy and average rate at which $\overline{\theta^2}$ is smeared at the molecular diffusive scale respectively. Measurements indeed show scaling regions but their slope is dependent on the type of flow and the Reynolds number. Sreenivasan (1991) has compiled various data from shear flow experiments (wakes and jets and boundary layers, both in the laboratory and in the field) and shows that the scalar scaling exponent increases from about 1.3 for a micro-scale (Taylor) Reynolds number, Re_λ, of about 200 to about 1.63 at $Re_\lambda = 2000$. The data seem to suggest an asymptotic limit of -5/3 although the Reynolds numbers have not been high enough to properly confirm this. It appears that for these strongly anisotropic shear flows R_λ must be significantly greater than 2000 before a locally isotropic region is sufficiently well established to fulfill the similarity requirements of the Kolmogorov-Corrsin-Obukhov theory.

Recently we (Jayesh and Warhaft 1991, 1992) have employed both the *toaster* and the *mandoline* to study passive scalar fluctuations (principally the scalar probability density function (pdf) and related statistics) in grid turbulence. In that work, for the mean gradient experiment, we varied the mean speed, U, and the mesh length M thereby varying the integral scale Reynolds number, Re_l from 60 to 1,100. Here $Re_l \equiv ul/\nu$ where u is the rms longitudinal velocity, l is the turbulence integral scale (close in value to the mesh length, M) and ν is the kinematic viscosity. This corresponds to a significant (but modest) variation of the micro-scale Reynolds number, R_λ, ($\equiv u\lambda/\nu$, where λ is the Taylor micro-scale) from approximately 30 to 130. These experiments have provided a broad data set from which passive temperature spectra, as a function of Reynolds number, can be studied.

Figure 2(a) shows four temperature spectra, for different Reynolds numbers, using the *toaster* to generate the passive thermal fluctuations (i.e., in the presence of a mean temperature gradient). The Reynolds numbers are given in the figure caption. The spectra show that there is a scaling region (a region of constant slope on a log-log plot) and that it increases with width as the Reynolds number increases.

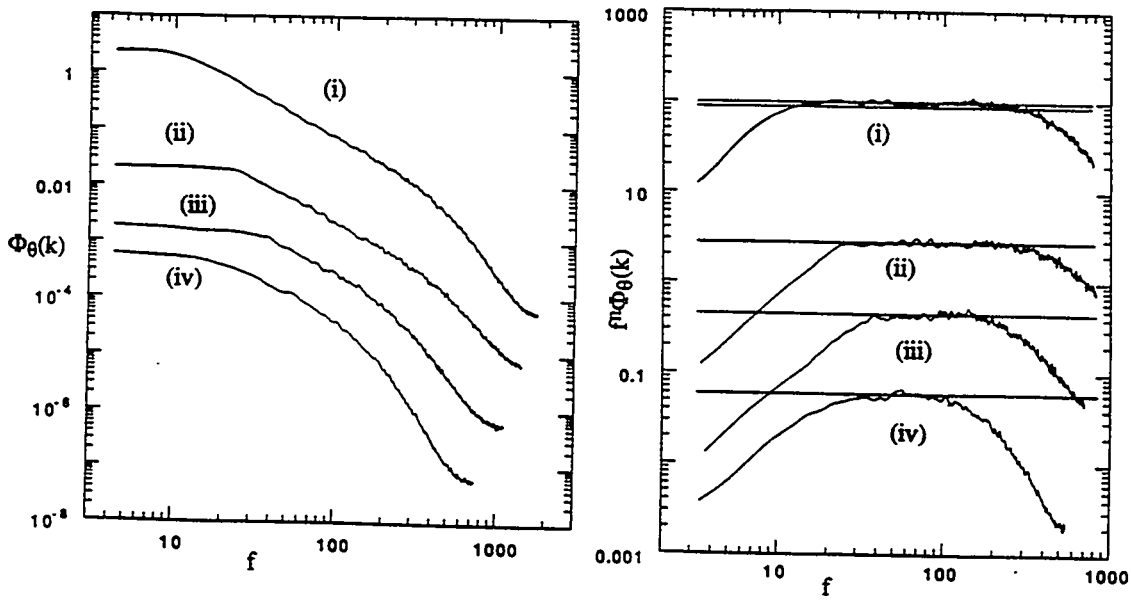


Figure 2 Temperature spectra in grid turbulence with a linear (passive) tempeature gradient. a) "Raw" spectra, b) spectra of a) multiplied by f^n where n is the slope of the scaling region in a). The Reynolds number for the four spectra are (i) $Re_l = 856$, (ii) $Re_l = 282$, (iii) $Re_l = 67$, (iv) $Re_l = 59$.

In order to more clearly display the scaling region we multiplied the ordinate of the raw spectra of Figure 1(a) by f^n (where f is the frequency and n is the scaling exponent). These spectra are shown in Figure 2(b). We determined n by fitting a least square best fit straight line to the scaling region of the raw spectra of Figure 2(a). The scaling region must be horizontal in the plot $f^n\phi_\theta(f)$ if the choice of n is correct. Figure 2(b) shows a clear scaling region of more than a decade for the high Re_l cases and no scaling region for the lowest Re_l case.

Figure 3 shows n as a function of Reynolds number for all of our experiments. Although there is quite a bit of scatter within each experiment (for a fixed Re_M), it is quite apparent that n does not have an overall variation with Reynolds number: its value for all the data was found to be 1.58 with a standard deviation of 0.07. Given the scatter the result is not inconsistent with the Kolmogorov scaling value of 1.67.

The data set of Figure 3 is mainly from the *toaster* experiments, for which there is a linear temperature profile. However we also obtained one data set for the *mandoline* and this gives the same scaling exponent as the toaster data (Figure 3). Note that although the velocity field is isotropic in both cases, the thermal field is not; for the mean gradient there is a heat flux and thus large scale anisotropy in the thermal field (Sirivat and Warhaft, 1983) while for the *mandoline* there is no heat flux (no mean gradient), suggesting approximate isotropy for the large scale thermal field (Warhaft and Lumley 1978). The same value of n obtained from these two different ways of creating the thermal field suggest it is the large scale structure of the velocity field (rather than the thermal field) that is relevant in determining the slope in the scaling region.

The width of the scaling region for all of our data as a function of Reynolds number is shown in Figure 4. The monotonic increase of the width with Reynolds number is consistent with fundamental notions of scaling (e.g. Tennekes and Lumley, Chapter 8) and implies that our data are not "anomalous" as was earlier thought from measurements done at a single Reynolds number. Note that here too, the *mandoline* data are consistent with the *toaster* data. Kolmogorov scaling shows that the width of the scaling region should increase as $Re_l^{3/4}$ since $l/\eta \sim Re_l^{3/4}$ where l is the integral scale and η is the dissipation scale (note for our experiment of Prandtl number 0.7, $\eta_\theta \sim \eta$ where is the thermal smearing scale). We have plotted a line of slope 0.75 on Figure 4 and it reasonably represents the trend in the data, given its scatter.

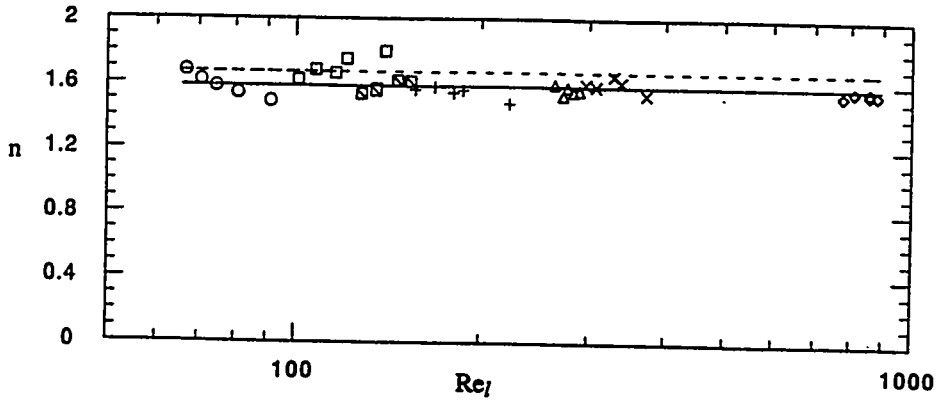


Figure 3. The slope of the scaling region for all experiments as a function of Re_l. The mean value is 1.58 (solid line). It is within one standard deviation of the Kolmogorov value of 5/3. The ⊠ symbols are for no tempeature gradient (mandoline), all other symbols are for the toaster.

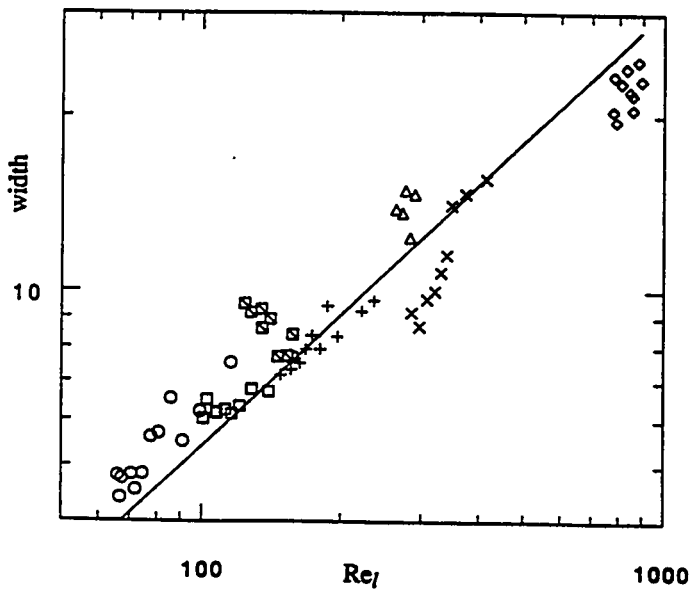


Figure 4. The width of the scaling region as a function of Re_l. The solid line has a slope of 0.75, the value predicted using Kolmogorov scaling: $l/\eta \sim Re_l^{0.75}$.

3. THERMAL DISPERSION AND MIXING IN A JET

Despite the importance of turbulent mixing in a jet, there appears to be no previous work on diffusion and dispersion from point or line sources; all previous experiments have introduced the scalar evenly throughout the flow (e.g. heating the jet (Corrsin and Uberoi 1980) or having a jet of pure species A mixing with the surroundings of pure species B (Dowling and Dimotakis 1990)). Thus there has been no information on how quickly one or two species mix in such a flow, an issue of fundamental importance. Here we extend the inference method of Warhaft 1984, which consisted of placing two line sources in grid turbulence, to placing two circular thin heated wire rings in a jet (Figure 5a). A single wire ring is analagous to a single line source; from it we can determine how long the fluctuations take to smear and fill the whole jet. Two line sources (or rings) enable us to determine, by inference, the cross correlation between the fluctuations, thereby providing information on the mixing rate of two independent species (Fig. 5b).

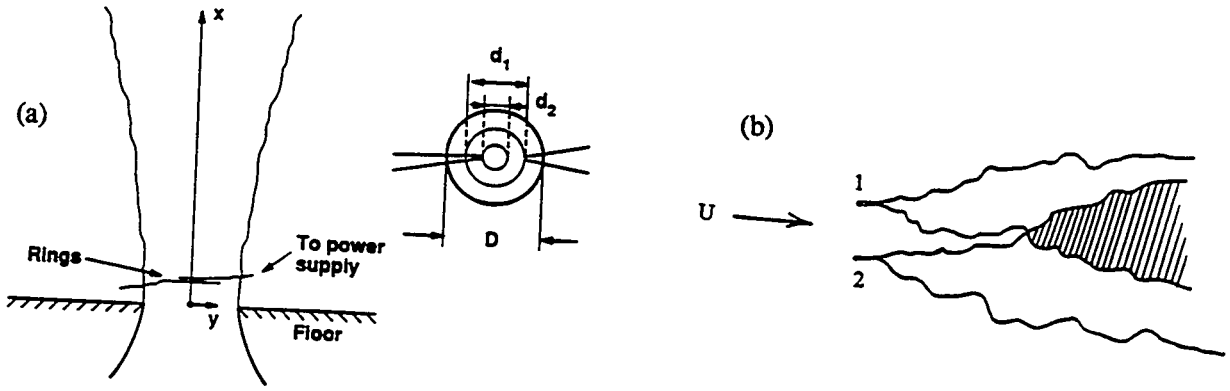


Figure 5a) Side and plan view of the jet of diameter D showing the fine wire rings (of diameters d_1 and d_2). The rings were placed concentrically above the jet, both in the same plane. They were suspended in the flow by means of their leads which were held by a clamp, outside the jet. The wires for the rings were 0.127 mm nichrome. b) Instantaneous thermal fields for two thermal line sources in a turbulent flow showing the region of overlap or interference (hatched region). This situation was studied by Warhaft 1984. Here we present the jet analogue using two fine wire rings.

As in our previous studies (Warhaft 1981, 1984) the cross correlation, ρ, between the thermal fields coming from each ring is determined by operating ring 1 and ring 2 separately and then operating them together. Under the assumptions of statistical stationarity, and that the scalar fields are passive, we determine the cross correlation from the relation

$$\rho = (\overline{\theta_b^2} - \overline{\theta_1^2} - \overline{\theta_2^2}) \Big/ 2\,(\overline{\theta_1^2})^{1/2}\,(\overline{\theta_2^2})^{1/2} \qquad (2)$$

where $\overline{\theta_b^2} = \overline{(\theta_1 + \theta_2)^2}$ is the variance with both rings operating and $\overline{\theta_1^2}$ and $\overline{\theta_2^2}$ are the variances of each ring operating separately. Figure 5b shows a sketch of a region of overlap (mixing of 2 scalars) for 2 line sources. Our interest is in determining ρ for this region, when two rings are used to generate the

thermal field (Figure 5a). A practical realization of our experiment would be two species in concentric pipes, mixing in a jet formation (Kerstein 1990).

We have carried out a systematic investigation of ring placement which was varied relative to the jet exit, and of Reynolds number. We have also studied the effect of the rings on the flow since they tend to slightly suppress the velocity fluctuations by inhibiting vortex paring. Our studies will be reported in Tong and Warhaft 1993. Here we show the effect of initial conditions on the evolution of ρ and compare it to grid turbulence.

Figure 6a shows ρ vs. x/D (where D is the jet diameter) for rings placed concentrically at x/D=9, i.e. at the beginning of the turbulent region. D for the jet was 30mm and the Reynolds number was 18,000. Notice that in spite of the large early differences in ρ, they all asymptote to unity very quickly, by about x/D = 20, which is equivalent to an eddy turn over time $S \left(\equiv \int_{x_0}^{x} \frac{dx/U}{l/u} \right)$ of about 2. (Here U and u are the mean and fluctuating velocity respectively l, is the integral scale and x_0 is the virtual jet origin). On the other hand, in grid generated turbulence, without mean shear, the evolution time for ρ is very long. Figure 6b compares the jet to the grid turbulence. Notice for comparable wire spacing complete mixing has not occurred by 4 eddy turn over times for the grid turbulence showing the important role of mean shear (and anisotropy) in the mixing process.

A systematic study of the evolution of ρ as well as temperature spectra and cospectra, pdf's and conditional dissipation is given in Tong and Warhaft 1993.

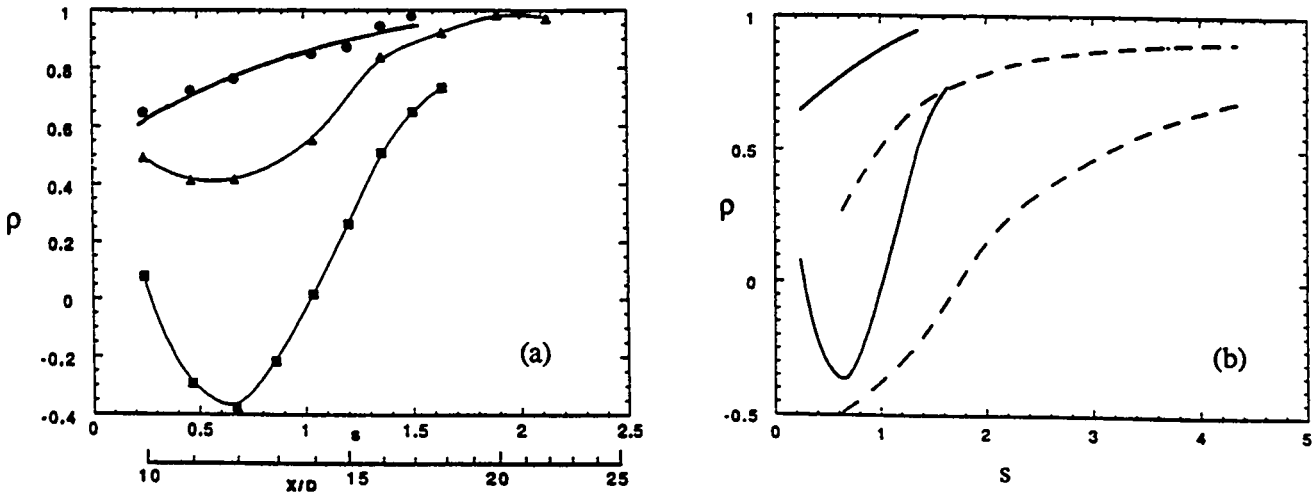


Figure 6a) ρ vs. x/D and S (the eddy turnover time) for the rings placed at x/D=9 for a 3 cm jet, Re_j= 1.8 x 10^4. The ring diameters for the circles are 35 and 40 mm, for the triangles they are 30 and 40 mm, and for the squares they are 20 and 40 mm. b) The data for the circles and squares of part a) compared with experiments carried out in grid turbulence for comparable wire spacing i.e. for comparable d/l where d is the distance between the wires and l is the integral scale of the turbulence (Warhaft 1984).

ACKNOWLEDGEMENTS

I am grateful to my students Jayesh and Chenning Tong and to Edward Jordan for technical assistance. The work was supported by the Department of Energy, Basic Energy Sciences under Grant DE-FG02-88ER13929.

REFERENCES

S. CORRSIN, "On the spectrum of isotropic temperature fluctuations in isotropic turbulence," *J. Appl. Phys.* 22, No. 4, 469-473 (1957).

S. CORRSIN, "Heat transfer in isotropic turbulence," *J. Appl. Phys.*, 23, 113 (1952).

S. CORRSIN, and M.S. UBEROI, "Further experiments on the flow and heat transfer in a heated turbulent air jet," NACA Rep. 998 (1950).

D.R. DOWLING and P.E. DIMOTAKIS, "Similarity of the concentration field of gas-phase turbulent jets," *J. Fluid Mech.*, 218, 109 (1990).

J.P. GOLLUB, J. CLARKE, M. GHARIB, B. LANE, and O.N. MESQUITA, "Fluctuations and transport in a stirred fluid with a mean gradient," *Phys. Rev. Lett.* 67, 3507 (1991).

JAYESH and Z. WARHAFT, "Temperature spectra in grid generated turbulence," (to be submitted to *Physics of Fluids A*).

JAYESH and Z. WARHAFT, "Probability distribution, conditional dissipation and transport of passive temperature fluctuations in grid-generated turbulence," *Phys. Fluids A* 4, 2292-2307 (1992).

JAYESH and Z. WARHAFT, "Probability distribution of a passive scalar in grid-generated turbulence," *Phys. Rev. Lett.* 67, 3503 (1991).

A.R. KERSTEIN, "Linear-eddy modelling of turbulent transport, Part 6, Microstructure of diffusive scalar mixing fields," *J. Fluid Mech.* 231, 361 (1991).

A.R. KERSTEIN, "Linear eddy modelling of turbulent transport, Part 3. Mixing and differential molecular diffusion in round jets," *J. Fluid Mech.*, 216, 411 (1990).

A.N. KOLMOGOROV "Local structure of turbulence in an incompressible fluid at very high Reynolds numbers," *Dokl. Akad. Navk. SSSR,* 30, No. 4, 299-303 (1941).

O. METAIS and M. LESIEUR, "Spectral large-eddy simulation of isotropic and stably stratified turbulence," *J. Fluid Mech.*, 239, 157 (1992).

A.M. OBUKHOV, "Structure of the turbulent temperature field in a turbulent flow," *Izv. Akad. Nauk SSSR, Ser. Georgr. i Geofiz.*, 13, No. 1, 58-69 (1949).

A. PUMIR, B. SHRAIMAN, and E.D. SIGGIA, "Exponential tails and random advection," *Phys. Rev. Lett.* 66, 2984 (1991).

A. SIRIVAT and Z. WARHAFT, "The effect of a passive cross-stream temperature gradient on the evolution of temperature variance and heat flux in grid turbulence," *J. Fluid Mech.*, 128, 323 (1983).

K.R. SREENIVASAN, "On local isotropy of passive scalars in turbulent shear flows," *Proc. Roy Soc. Lon.* A 434, 165-182 (1991).

H. TENNEKES and J.L. LUMLEY, *A First Course in Turbulence*, MIT Press, Cambridge, MA, (1972).

C. TONG and Z. WARHAFT, "Thermal dispersion and mixing in a turbulent jet," (to be submitted to *Physics of Fluids A*).

Z. WARHAFT, "The interference of thermal fields from line sources in grid turbulence," *J. Fluid Mech.*, 144, 363 (1984).

Z. WARHAFT and J.L. LUMLEY, "An experimental study of the decay of temperature fluctuations in grid-generated turbulence," *J. Fluid Mech.*, 88, 659 (1978).

PARTICLE INTERACTIONS IN CONCENTRATED SUSPENSIONS

Lisa A. Mondy
Sandia National Laboratories, Division 1512, Albuquerque, New Mexico 87185

Alan L. Graham
Los Alamos National Laboratory, MEE- 9, Los Alamos, New Mexico 87545

Howard Brenner
Dept. of Chem. Eng., Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

James R. Abbott
Los Alamos National Laboratory, MEE-9, Los Alamos, New Mexico 87545

ABSTRACT

An overview is presented of research that focuses on slow flows of suspensions in which colloidal and inertial effects are negligibly small. We describe nuclear magnetic resonance imaging experiments to quantitatively measure particle migration occurring in concentrated suspensions undergoing a flow with a nonuniform shear rate. These experiments address the issue of how the flow field affects the microstructure of suspensions. In order to understand the local viscosity in a suspension with such a flow-induced, spatially varying concentration, one must know how the viscosity of a *homogeneous* suspension depends on such variables as solids concentration and particle orientation. We suggest the technique of falling ball viscometry, using small balls, as a method to determine the effective viscosity of a suspension without affecting the original microstructure significantly. We also describe data from experiments in which the detailed fluctuations of a falling ball's velocity indicate the noncontinuum nature of the suspension and may lead to more insights into the effects of suspension microstructure on macroscopic properties. Finally, we briefly describe other experiments that can be performed in quiescent suspensions (in contrast to the use of conventional shear rotational viscometers) in order to learn more about boundary effects in concentrated suspensions.

INTRODUCTION

Many industrial processes include the transport of suspensions of solid particles in liquids, such as coal and other solid feedstock slurries. Oil, gas, and geothermal energy production rely on the transport of suspensions such as muds, cements, proppant, and gravel slurries in the drilling and completion of a well. Suspensions are not only ubiquitous in energy production, but also in high-energy-consumption industrial processes such as found in pulp and paper manufacturing. The complex rheological response of suspensions often limit the efficiency of the design of such processes, causing loss of productivity, increased cost, and increased energy consumption. Because of the importance of particulate two-phase flows in the applications described above, the study of suspension rheology remains an important component of the technical research directed by a national energy policy.

This overview of our recent research supported by the Department of Energy, Office of Basic Energy Sciences, will focus on slow flows of suspensions of relatively large particles, in which colloidal and inertial effects are negligibly small. There is growing evidence that even in this restricted range of flows, the rheology of a suspension with a nondilute particle concentration cannot be characterized by a single viscosity. Instead, the microstructure of the suspension determines the overall macroscopic properties, and the flow

history of the suspension determines aspects of the microstructure. (A good overview of flow-induced microstructural changes can be found in an article by Acrivos [1].) Hence, conventional viscometers, which impose macroscopic flow fields, may not measure the viscosity of the homogeneous suspension originally introduced into the viscometer, but rather may represent a property governed by the nonhomogeneous structure created by the flow. Such structures may be intrinsically different for various classes of flow fields associated with different viscometers.

Advances in the ability to predict the rheological response of concentrated suspensions depend on answering three broad questions: 1) How does the macroscopically imposed flow field affect the microstructure of a suspension? 2) How does the microstructure of a suspension affect the rheological properties? 3) How do boundaries, such as walls, affect the microstructure and properties? Aspects of these questions are being addressed in our work.

In the following section we will illustrate the existence of flow-induced microstructural changes with data on the time evolution of concentration and velocity profiles in suspensions undergoing flow between counter-rotating concentric cylinders (similar to the geometry found in "cup and bob" or "Couette" viscometers). We will show that the resultant profiles in these one-dimensional flows can be predicted well by the expressions describing "hydrodynamic diffusion" originally developed by Leighton and Acrivos [2, 3, 4]. However, additional phenomena arise in more complex flows, such as the two-dimensional migration of particles seen in the eccentric annular gap of a "journal bearing" flow. This illustrates that the complex interaction of particles cannot be adequately described by the one-dimensional theory originally proposed by Leighton and Acrivos, and it suggests that other avenues be taken to relate the macroscopic behavior to the evolution of microstructure. One such avenue recently suggested in the literature is to use a kinetic theory approach, which has been applied successfully in granular flows [5, 6]. In this theory the intensity of the velocity fluctuations, caused by particle interactions, is characterized by a "granular temperature" analogous to the temperature in classical kinetic theories and governed by a balance of fluctuation energy. Kinetic theory approaches emphasize the importance of obtaining experimental data not only on average behavior of suspensions but also on the fluctuations about those averages.

In the third section we will discuss the use of falling ball rheometry as a means to circumvent the problems encountered with using conventional rotational devices to measure suspension viscosity. If the size of the falling ball is of the order of the characteristic length of the suspended particles, the ball disturbs the original microstructure of the quiescent suspension only slightly as it falls. Hence, one can use this technique to determine the viscosity of a homogeneous suspension (or likewise one with any set microstructure). One can then incorporate this information into a constitutive relationship to determine the *local* viscosity in a flow field, given that the local concentration is known [4]. Furthermore, falling ball rheometry is not limited to the measurement of macroscopic average viscosities. The velocity fluctuations experienced by the falling ball can also be measured and can give insights into the importance of particle interactions.

We have also proposed use of quiescent suspensions in other apparatus to provide further insights into the rheological behavior of concentrated suspensions, especially the effects of boundaries. Rolling ball rheometry could be explored as a means to determine the effect on the local viscosity of the microstructure imposed by the wall. Measuring the torque on a ball spinning in an otherwise quiescent suspension has been proposed as a sensitive measure of slip at the wall. These ideas will be addressed in the fourth section of this article, and the results of preliminary measurements will be discussed. In the final section we will summarize our conclusions.

EFFECTS OF FLOW ON THE MICROSTRUCTURE OF SUSPENSIONS

Flow-induced migration and ordering of suspended particles have been hypothesized to create viscosity measurements that vary with the total strain to which a given suspension has been subjected [7, 3]. This migration is thought to occur whenever particle interactions are stronger or more frequent in one part of a flow field than in another, as could occur in the presence of spatially varying shear rate, concentration, or viscosity fields. A Newtonian fluid in the annular domain between rotating concentric cylinders (i.e. wide-gap Couette apparatus) possesses perhaps the simplest flow field of any realizable *nonhomogeneous*

shear flow. As such, this is a useful device in which to study the effects of nonhomogeneous shear on the microstructure of a concentrated suspension.

The spatial distribution of suspended particles present in concentrated suspensions is difficult to measure because most suspensions are opaque even at relatively low particle concentrations. However, under the auspices of the Department of Energy, Office of Basic Energy Sciences, noninvasive techniques based on nuclear magnetic resonance (NMR) imaging have been developed by Fukushima and coworkers to study both concentration and velocity profiles in multiphase flows [8, 9]. We have employed these NMR imaging techniques to study the flow-induced migration of particles in a suspension undergoing flow in a wide-gap Couette apparatus. The details of the experiments can be found elsewhere [10, 11]. However, some results of these studies will be briefly discussed here in order to illustrate how dramatically a suspension's microstructure can be affected during flow.

The primary data obtained from these experiments are NMR images of the concentration and velocity fields. Representative examples of the concentration images are shown in Figure 1. As shown on the left, the initial image of a bimodal suspension (60 vol% neutrally buoyant spheres, of which 65% are 3.175 mm and 35% are 780 μm in diameter, in a viscous Newtonian liquid) is essentially uniform. Dark areas represent areas of high solid concentration. Individual large spheres can almost be distinguished, although the thickness of the imaged volume (2.4 cm) results in a blurring of the particles. After 3600 revolutions of the inner cylinder, the final image was taken, shown on the right. In this image the fluid fraction is significantly higher near the inner rod (the region of highest shear rate) and lower near the outer cylinder. It is apparent that the particles have migrated from the region of highest shear to the region of lowest. Furthermore, distinct shells of larger spheres, interspersed with fluid and smaller spheres, can be seen in this image. From visual observations of the particles near the outer wall of the apparatus, the larger spheres appear approximately hexagonal close packed within the outermost shell. That is, the arrangement of the large spheres is two-dimensional hexagonal close packed in concentric sheets. This structure begins to appear very quickly: significant migration can be detected after only 50 revolutions of the inner cylinder. It is important to note that this migration does not appear to be caused by inertial effects, which are negligible at the rates of rotation, the viscosity of the suspending liquid, and the particle sizes involved here. Subsequent experiments with suspensions of unimodal distributions of large spheres revealed that this shear-induced structure was not unique to bimodal suspensions.

The concentration can be quantified in any region of the image by recognizing that the fluid in the imaged slice gives a full-intensity signal and the particles give no signal. The normalized value of the image intensity is proportional to the density of the liquid phase protons in a volume element. By using an imaging sequence with a slice selective refocusing pulse, as proposed by Cho and coworkers [12], the relative phase shift can be made proportional to the velocity. By using such a technique we can also find the velocity in each region of an image of a flowing suspension. Figure 2 shows velocity measurements for a suspension of 50 vol% of 675 μm spheres undergoing flow in a wide-gap Couette apparatus after a steady-state microstructure has formed. The velocity profile falls off much more rapidly than in a Newtonian fluid (shown for comparison by the solid curve). The particle concentration approaches maximum packing near the outer wall, and the velocity profiles reveal that the suspension is almost stagnant in this region.

In addition to expanding our general understanding of the micromechanics of shear-induced migration, the primary purpose of the NMR imaging studies was to determine the dependence of the particle migration on a number of experimental parameters. These parameters included strain, shear rate, and viscosity of the suspending fluid, as well as concentration and diameter of the suspended particles. The results of a constitutive model based on Leighton and Acrivos' scaling arguments compared very favorably to these experimental results [4].

This constitutive model consists of both a Newtonian constitutive equation, in which the viscosity depends on the local particle volume fraction, and a diffusive equation that accounts for shear-induced particle migration. Two adjustable parameters arise in the diffusive equation, which describe the relative strength of the mechanisms for particle migration. These two rate parameters were taken to be constants and were evaluated by comparison to the experimental measurements of velocity and concentration profiles in the wide-gap Couette apparatus for one suspension at one strain. With these parameters fixed, predictions for particle concentration profiles were then compared to the experimental results for suspensions with a wide

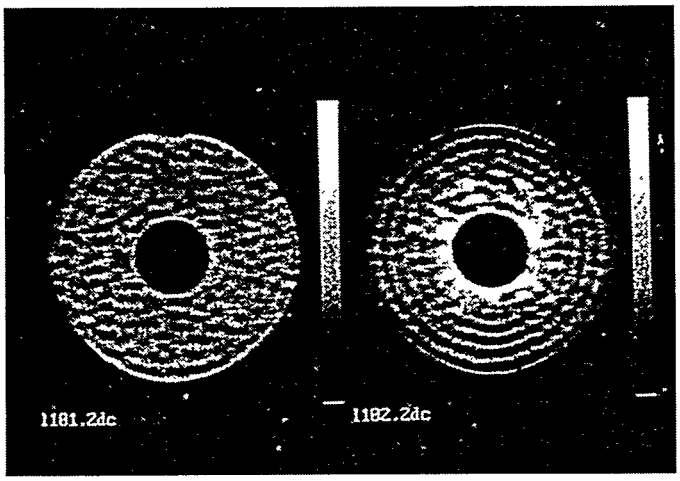


Figure 1: NMR images of a cross section of a suspension of 60 vol% bidisperse spheres between concentric cylinders. The image on the left represents the initially well dispersed state of the suspension. The image on the right was taken after rotating the inner cylinder until a steady state was achieved. The bright area near the inner cylinder represents a higher fluid fraction, indicating that the particles have migrated away from this area of higher shear rate.

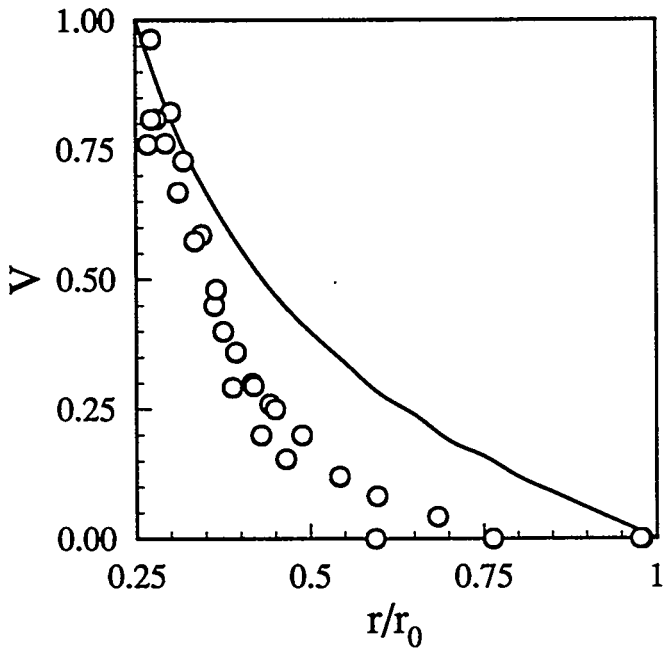


Figure 2: Steady-state velocity profile for a suspension of 50 vol % spheres with a mean diameter of 675 μm. The azimuthal velocity was measured along one diameter of the image. The solid curve is the velocity profile for a Newtonian fluid.

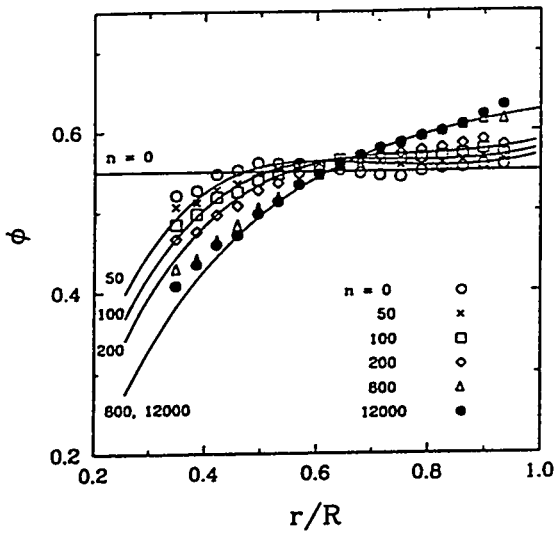


Figure 3: Transient profiles of the particle concentration for a suspension of 55 vol% particles with a mean diameter of 675 μm compared to predictions. Results are shown for the initial profile (o) and after the number of revolutions of the inner cylinder (n) equals 50 (×), 100 (□), 200 (◇), 800 (△), and 12000 (•). Model parameters used are those that best fit the data at n=200. Both model and data show that steady state is reached by the time the inner cylinder has rotated about 800 times.

range of particle sizes and concentrations.

Figure 3 shows the remarkably good comparison between the predictions of this model and the transient concentration profiles obtained for a suspension of 675 μm spheres at a volume fraction of 0.55. Figure 4 shows the steady-state concentration profiles for suspensions of either 100 or 675 μm particles compared with the predictions. The agreement between model and experiment is again excellent, with the theory fitting the experimental data for the suspension of 100 μm particles nearly as well as it did for the suspension of much larger particles used to calculate the two rate parameters.

Excitement generated by these results must be tempered by the results of ongoing research in more complicated two-dimensional flows. NMR imaging has also been used to study the flow of concentrated suspensions in the gap between a rotating inner cylinder placed eccentrically within an outer fixed cylinder (a journal bearing). With a Newtonian fluid, the majority of the flow will be in a cell concentric with the inner cylinder; however, with certain placements of the inner cylinder, a second cell, which rotates in the opposite direction, forms near the region of the outer wall furthest from the inner cylinder [13]. We have used NMR imaging to confirm that similar behavior occurs in concentrated suspensions. Here, particle migration creates a region of maximum solids concentration in the low-shear-rate region of the second cell (away from the wall).

The constitutive expression previously described was subsequently expanded to two-dimensional flows by describing the flow in terms of the strain rate tensor $\mathbf{D}$ and the migration in terms of gradients in the generalized shear rate $\dot{\gamma}=\sqrt{2\mathrm{tr}\mathbf{D}^2}$. The equation set was then again solved numerically and the predictions compared to the NMR imaging data. Unfortunately, this model failed to predict that the steady-state maximum concentration is not always at the outer wall, but in many cases is at a location within the gap.

The failure of the simple expression for one-dimensional hydrodynamic diffusion to capture the qualitative nature of this two-dimensional flow suggests that it has not been appropriately generalized to multiple dimensions and that other avenues should be explored in attempting to relate the macroscopic behavior to the evolution of microstructure. One such avenue recently suggested is to use a kinetic theory approach, which has been applied successfully in granular flows [5, 6]. In this theory the intensity of the velocity fluctuations, caused by particle interactions, is characterized by a "granular temperature" analogous to the

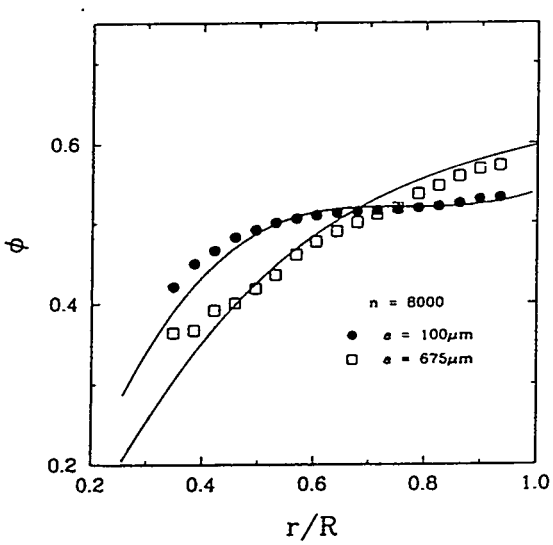


Figure 4: Measurements and predictions of the concentration profiles for suspensions of 50 vol% of either 100 μm (•) or 675 μm (□) particles. Both results are shown for 8000 rotations of the inner cylinder. Model parameters are unchanged from those used in Figure 3.

temperature in classical kinetic theories and governed by a balance of fluctuation energy. This approach emphasizes the importance of measuring not only average behavior of suspensions but the details of the fluctuations about those averages.

FALLING-BALL RHEOMETRY IN SUSPENSIONS

In previous work, we have shown that falling ball rheometry is an excellent tool to probe the rheological properties of a suspension without changing the properties through the very act of measuring them. Unlike conventional viscometers, which employ flow fields that tend to influence the microstructure of the suspension, falling ball rheometry can be used to determine the macroscopic viscosity of a suspension with little effect on the microstructure [14].

The discrete nature of the suspension is readily apparent in falling ball experiments. One expects a very large ball to fall smoothly through a suspension of tiny particles and its velocity to appear fairly constant. However, when we observe the passage of a ball of the same diameter as large suspended particles, we see that actually the velocity is not constant. Periods of almost no motion, as the falling ball approaches and "rolls off" suspended particles, alternate with periods of almost free fall in the interstices between suspended spheres. However, a statistical analysis reveals that the *average* terminal velocity of the ball, measured over a distance usually between 100 and 1000 suspended particle diameters, is reproducible. Furthermore, if this average terminal velocity, corrected for Newtonian wall effects, is translated into a viscosity, this viscosity is independent of the diameter of the falling ball relative to the diameter of the suspended particles over a wide range of falling-ball sizes.

We are also exploring the possibility of using the fluctuations in the terminal velocity, as the ball interacts with individual suspended particles or clusters of suspended particles, to give information about the suspension microstructure. Whereas the mean settling velocity predicts the continuum behavior of the suspension, the dispersivity around the mean velocity allows insight into the non-continuum behavior of the suspension caused by the presence of the macroscopic suspended spheres.

We performed experiments which focused on the three-dimensional dispersion of a single ball settling through a suspension of neutrally buoyant particles. The detailed paths of falling balls were recorded from

direct observations in transparent suspensions (in which the refractive index of the suspending fluid was matched to that of the suspended particles) and using real-time radiography in opaque suspensions. The primary experimental parameters were the relative size of the settling ball and suspended particles, and the concentration and geometry of the suspended particles.

A principal objective of these experiments was to test whether the observed variations in the ball's settling velocity were the result of a Fickian (random) process. For sufficiently long times, the variances grew linearly with time, as predicted for a Fickian process. Because the falling ball in these experiments was similar in size to the suspended particles, it was possible to see the transition between the deterministic effects of a sphere settling past a particular arrangement of particles and the random process associated with a sphere settling past many such arrangements. The deterministic effects resulted in a quadratic growth in the variances for short times. The short-time nonlinear variances of Brownian tracer particles can be described in terms of particle inertia [15]. However, the short-time behavior of a ball falling in a suspension was not caused by the inertia of the falling ball, but rather by the time needed for the ball to change its local environment. This conclusion was supported experimentally by the observed insensitivity of the dispersive behavior to the falling ball's density. To determine the dispersivity when the variances were deterministic, we estimated the time scale on which it takes the settling ball to change its local neighborhood. (Because the settling ball tends to drag the suspension along with it, this time scale is greater than the time to travel just a few ball diameters.) This time scale was then used in a model, similar to that for Brownian tracers, relating the measured short-time variances to the dispersivity.

The resulting dimensionless vertical dispersivities are shown in Figure 5 as functions of falling ball size and volume concentration of suspended spheres. For moderately concentrated suspensions, the vertical dispersivity decreased with increasing ball size, but always less rapidly than predicted by Davis and Hill for dilute suspensions [16]. This effect decreased with concentration, until, for a suspension with a solids volume concentration of 50%, the dispersivity was independent of ball size. At a constant size of falling ball relative to the suspended spheres, the vertical dispersivity increased approximately linearly with concentration. For example, for falling balls the same size as the suspended spheres, the dimensionless vertical dispersivity D^* was observed to depend on the volume fraction of solids, ϕ, as $D^* = 0.60\phi^{1.08}$. The measured horizontal dispersivity was at least 25 times smaller than its vertical counterpart (and below the experimental resolution for suspensions with only 15 vol% solids).

Vertical dispersivity was also measured in suspensions of randomly oriented rods. Here, the dispersivity was always virtually independent of ball size (however, one should note that the balls tested were always significantly larger than the rod diameter). The vertical dispersivity in these suspensions increased linearly with the specific viscosity. Because the viscosity was approximately linear with volume fraction for the suspensions tested [17], this parallels the theoretical behavior in dilute suspensions of spheres [16].

We have recently begun to observe not only the velocity of, but also the pressure drop across a ball falling in a quiescent suspension. For a ball falling in a single-phase Newtonian liquid, this pressure is constant, independent of viscosity (at low Reynolds numbers), and can be described analytically [18]. Although the pressure drop is independent of viscosity in a Newtonian liquid, it is reasonable to speculate that in a suspension the pressure drop may be dependent on the microstructure. Therefore, like the local viscosity, it may vary due to the discrete nature of the material.

We first tested this speculation by modeling a falling ball rheometer numerically with the boundary element method [19]. In this technique, the boundary integral equations for Stokes flow, coupled with the equilibrium equations for the solid particles, are discretized and solved numerically. Fully three-dimensional simulations of suspensions of spheres in a Newtonian liquid bounded by cylindrical walls were performed. The number and size of individual suspended spheres were varied to give volume concentrations ranging from zero to 5%. In these simulations, the pressure drop was influenced only weakly, if at all, by the introduction of neutrally buoyant particles. Furthermore, the arrangement of particles affected the pressure drop negligibly. In contrast, the relative viscosity in these simulations varied by over 10%. These results were consistent with the hypothesis that, despite the noncontinuum nature of these suspensions, each suspension could be treated as a Newtonian liquid with an effective relative viscosity. It also implied that the fluctuations in the velocity of the falling ball would be far more indicative of microstructural variations than would the corresponding fluctuations in pressure drop. However, the suspensions studied were all relatively dilute, and

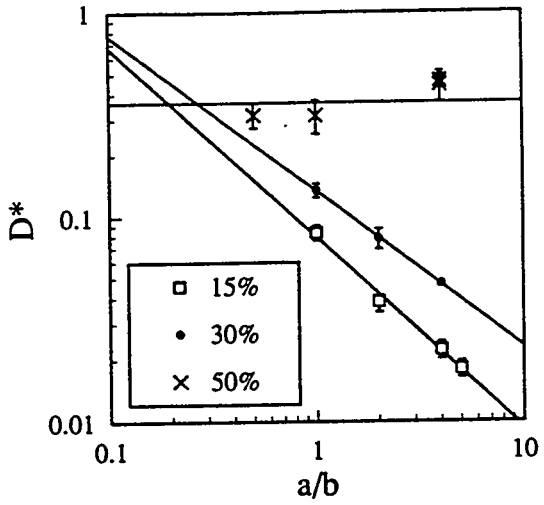


Figure 5: The dimensionless dispersivity as a function of the ratio of the radius of the settling ball (a) to the radius of the suspended spheres (b) for various volume concentrations of solids: 15 vol% (□), 30 vol% (•), or 50 vol% (×). The error limits shown are based on the nonlinear analysis which assumes that the time scale is known exactly, so they represent a minimum for the actual 95% confidence limits.

this behavior may not be present at high concentrations.

In laboratory experiments we have begun to measure the pressure drop across balls of various sizes falling in more concentrated suspensions. Preliminary data indicate that the pressure drop occurring in a suspension made with 30 vol% uniformly sized spheres is identical to that predicted to occur in a single-phase Newtonian liquid. Like the viscosity measurements in moderately concentrated suspensions, the pressure drop behavior is independent of the relative sizes of the falling ball and suspended spheres. (However, the absolute measure of pressure drop in both single- and two-phase fluids is directly proportional to the weight of the falling ball.) Experiments are planned to look at suspensions of higher concentrations, as well as at falling balls much smaller than the suspended spheres.

OTHER EXPERIMENTS IN QUIESCENT SUSPENSIONS

In the falling ball experiments described in the section above, the drag on the ball appeared to be that found in a Newtonian liquid with *no slip* at the boundaries. Instead of measuring the mean velocity of a falling ball, we could instead measure the mean torque on a spinning ball. This geometry is more sensitive to slip at the ball boundary. Whereas the force F on a ball moving slowly through an unbounded Newtonian liquid without slip can be described as $F = 6\pi\mu a v$ (where μ is the viscosity of the liquid and a and v are the radius and velocity of the ball, respectively), the force with perfect slip is $4\pi\mu a v$. In contrast, the torque T on a ball spinning slowly in a Newtonian liquid is $8\pi\mu a^3\Omega$ (where Ω is the angular velocity of the ball); however, the torque on a ball with perfect slip at the boundaries is zero [20].

Kunesh and coworkers studied the torque on balls spinning in single-phase Newtonian liquids, verified the formula above, and quantified the effects of the free surface [21]. We propose similar experiments to measure the torque on balls spinning in otherwise quiescent suspensions. We will analyze the data for any apparent slip at the balls' boundaries.

Figure 6 is an illustration of the apparatus we have built with this goal in mind. A calibrated torque wire holds a ball on a rod while the suspension is rotated on a motorized platform. The number of rotations

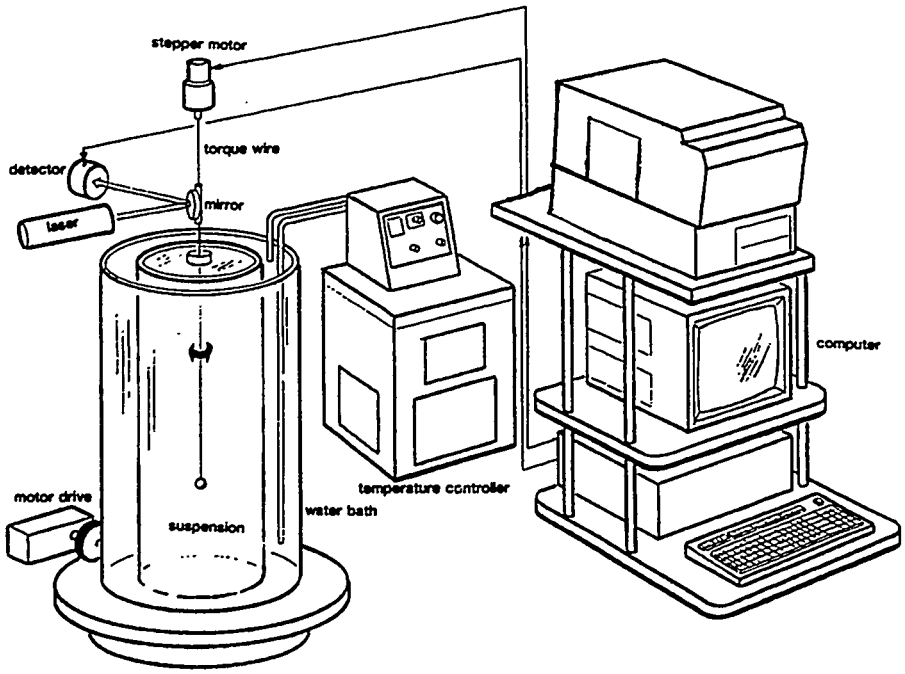


Figure 6: Apparatus to measure the torque on a ball spinning in a suspension.

of the platform is recorded automatically. A mirror at the bottom of the torque wire reflects a laser beam to a detector. The data from the detector is fed to a stepper motor holding the torque wire from above. The controller on the stepper motor uses this information to correct the position of the stepper motor and torque wire so that the reflected laser beam is maintained at the center of the detector. The corrections are recorded so that one knows how much the torque wire has twisted as a function of time. The twist can be directly related to the torque experienced by the ball and rod. We plan to measure this torque as a function of Ω on balls of various sizes relative to the suspended particles and in suspensions of various concentrations.

One must note that this flow geometry has the potential to induce structure over time. This will result in a decrease of the measured torque over time. However, the apparatus will allow us to record any time history associated with the torque, and the short-time behavior should still be an indication of any apparent slip at the ball's surface. The time history obtained will also allow us to study the fluctuations of the torque about the mean, which again may be indicative of the suspension microstructure.

The presence of the walls of a container can also induce structure. Another proposed study is of the behavior of a ball rolling down the wall of a inclined container holding a concentrated suspension. In rolling ball viscometry, a dense ball is allowed to roll/slide down an inclined surface, and its rate of travel is compared to that in a fluid of known viscosity. In a suspension, we can ratio the time it takes for the ball to travel a known distance in the suspension to that in the suspending liquid alone and, from this ratio, estimate the apparent relative viscosity. This procedure is similar to estimating the apparent viscosity with falling ball viscometry; however, the immediate region of the suspension seen by the moving ball is not uniform but has structure determined by the proximity of the bounding wall.

Preliminary studies have been performed with a suspension of 30 vol% 0.318-cm-diameter spheres neutrally buoyant in a viscous Newtonian liquid. The mean velocity of balls of three sizes ranging from 0.238 to 1.905 cm were first measured as they rolled down an 11° incline in the suspending liquid alone. Then the suspended particles were added, the suspension mixed well, and the measurements repeated. Again, as in the falling ball study, the moderately concentrated suspension behaved as a single-phase Newtonian liquid with an effective viscosity. The viscosity implied by the mean velocities of the rolling balls was independent of the ball size and was statistically indistinguishable from that measured in the falling ball experiments.

If one assumes that the suspension microstructure closest to the ball influences most the ball's velocity, then these results imply that in a moderately concentrated suspension, the microstructure near the bounding walls is similar to that in the bulk suspension. (The assumption of nearest-neighbor domination has been shown to be a good one in boundary element method calculations of the effect of neutrally buoyant particles

various distances from a falling ball. Here, particles beyond about 5 ball diameters away exerted negligible influence on the ball's velocity. In other words, the ball fell at the same velocity whether or not the far-field particles were present [19].) Further experiments in very concentrated suspensions are planned. Here, the walls are more likely to induce structure (as seen with NMR imaging near the outer walls of the wide-gap Couette apparatus after the particles have migrated and concentrated to near maximum packing).

CONCLUSIONS

We have performed a variety of experimental and numerical studies to elucidate the linkage between the microstructure and the macroscopically observed responses of suspensions of particles in liquids. NMR imaging studies and visual observations have confirmed that a suspension's microstructure can change dramatically during flow: large concentration gradients can be formed from regions of low shear rate to regions of high shear rate, ordered structure can form at the walls in regions of high concentration, particles of aspect ratio greater than 1.0 can align, etc. Conventional rotational viscometers may induce such changes in the microstructure over time, and, therefore, the data from them may not be accurate measurements of the viscosity of the suspension originally introduced into the viscometer. In fact, a suspension cannot be simply described by a single effective viscosity.

Falling ball viscometers, on the other hand, can be used (with small falling balls) to determine an apparent viscosity of a *homogeneous* suspension, without significantly affecting the microstructure during the measurement. Such a measurement can then be combined with information about the evolving microstructure in a flow to predict the spatial variations in viscosity and the global behavior. However, further studies of the details of particle interactions are needed before definitive predictive capabilities can be developed. Measurement of the detailed fluctuations of the velocity of a ball falling through a suspension is an example of one such study.

Quiescent suspensions can also be used to examine effects of boundaries. We propose to complete two such studies: measurements of the torque on a spinning ball and the drag on a ball rolling down a wall. The former should be a more sensitive measure of apparent slip at the ball boundaries. The latter may elucidate the effect of structure induced by the proximity of walls.

ACKNOWLEDGMENTS

This work was sponsored by the U. S. Department of Energy, at Los Alamos National Laboratory under Contract W-7405-ENG-36 with the University of California and at Sandia National Laboratories under Contract DE-AC04-76DP00789. The authors would like to acknowledge gratefully support for this work by the U.S. Department of Energy, Division of Engineering and Geosciences, Office of Basic Energy Sciences. Partial funding for the work done at Los Alamos and Sandia was also provided by SDIO/IST and managed by the Office of Naval Research.

References

[1] A. Acrivos. The rheology of concentrated suspensions of non-colloidal particles. In M. C. Roco, editor, *Particulate Two-Phase Flow*, chapter 5, pages 169–189. Butterworth-Heinemann, Boston, 1993.

[2] D. T. Leighton and A. Acrivos. Measurements of shear-induced self-diffusion in concentrated suspensions of spheres. *J. Fluid Mech.*, 177:109–131, 1987.

[3] D. Leighton and A. Acrivos. The shear-induced migration of particles in concentrated suspensions. *J. Fluid Mech.*, 181:415–439, 1987.

[4] R. J. Phillips, R. C. Armstrong, R. A. Brown, A. L. Graham, and J. R. Abbott. A constituative equation for concentrated suspensions that accounts for shear-induced particle migration. *Phy of Fluids*, 35(5):773–795, 1992.

[5] J. T. Jenkins and S. B. Savage. A theory for the rapid flow of identical, smooth, nearly elastic, spherical particles. *J. Fluid Mech.*, 134:401–433, 1983.

[6] D. F. McTigue and J. T. Jenkins. Channel flow of a concentrated suspension. In H. H. Shen et al., editor, *Advances in Micromechanics of Granular Materials*, pages 381–390. Elsevier Science Publishers, New York, 1992.

[7] F. Gadala-Maria and A. Acrivos. Shear-induced structure in a concentrated suspension of solid spheres. *J. Rheo.*, 24:799–811, 1980.

[8] P. D. Majors, R. C. Givler, and E. Fukushima. Velocity and concentration measurements in multiphase flows by nmr. *J. Magnetic Resonance*, 85:235, 1989.

[9] A. Caprihan and E. Fukushima. Flow measurements by NMR. *Physics Reports*, 198(4):195, 1990.

[10] A. L. Graham, S. A. Altobelli, E. Fukushima, L. A. Mondy, and T. S. Stephens. NMR imaging of shear-induced diffusion and structure in concentrated suspensions under going Couette flow. *J. Rheo.*, 35:191–201, 1991.

[11] J. R. Abbott, N. Tetlow, A. L. Graham, S. A. Altobelli, E. Fukushima, L. A. Mondy, and T. S. Stephens. Experimental observations of particle migration in concentrated suspensions: Couette flow. *J. Rheo.*, 35(5):773–795, 1991.

[12] Z.H. Cho, C.H. Oh, C. W. Mun, and Y. S. Kim. Some results of high-flow-velocity NMR imaging using selection gradient. *Magn. Reson.*, 3:857, 1986.

[13] G. H. Wannier. A contribution to the hydrodynamics of lubrication. *Quarterly of Applied Mathematics*, 8(1):1–32, 1950.

[14] J. R. Abbott, H. Brenner, A. L. Graham, and L. A. Mondy. Sedimentation of particles through quiescent suspensions. In *Proceedings of the Eighth Symposium on Energy Engineering Sciences*, Argonne, Illinois, May 1990. Argonne National Laboratory.

[15] G. E. Uhlenbeck and L. S. Ornstein. On the theory of the brownian motion. *Phys. Rev.*, 36:823–841, 1930.

[16] R. H. Davis and N. A. Hill. Hydrodynamic diffusion of a sphere sedimenting through a dilute suspension of neutrally-buoyant spheres. *J. Fluid Mech.*, 236:513–533, 1992.

[17] W. Milliken, M. Gottlieb, A. L. Graham, L. A. Mondy, and R. L. Powell. The viscosity–volume fraction relationship of a randomly oriented suspension of rods. *J. Fluid Mech.*, 202:217–232, 1989.

[18] I. Pliskin and H. Brenner. Experiments on the pressure drop created by a sphere settling in a viscous liquid. *J. Fluid Mech.*, 17:89–96, 1963.

[19] L. A. Mondy, M. S. Ingber, and S. E. Dingman. Boundary element method simulations of a ball falling through quiescent suspensions. *J. Rheo.*, 35(4):825–848, 1991.

[20] H. Lamb. *Hydrodynamics.* Dover Publications, New York, 6th edition, 1945.

[21] J. G. Kunesh, H. Brenner, M. E. O'Neill, and A. Falade. Torque measurements on a stationary axially positioned sphere partially and fully submerged beneath the free surface of a slowly rotating viscous fluid. *J. Fluid Mech.*, 154:29–42, 1985.

MOLECULAR MIXING IN TURBULENT FLOW

Alan R. Kerstein

Sandia National Laboratories
Livermore, CA 94551-0969

ABSTRACT

The evolution of a diffusive scalar field subject to turbulent stirring is investigated by comparing two new modeling approaches, the linear-eddy model and the clipped-laminar-profile representation, to results previously obtained by direct numerical simulation (DNS) and by mapping-closure analysis. The comparisons indicate that scalar field evolution is sensitive to the bandwidth of the stirring process, and they suggest that the good agreement between DNS and mapping closure reflects the narrowband character of both. The new models predict qualitatively new behaviors in the wideband stirring regime corresponding to high-Reynolds-number turbulence.

INTRODUCTION

The advection of a passive, diffusive scalar field, whether by a deterministic or a stochastic stirring mechanism, is a process whose richness becomes increasingly apparent as various configurations are investigated. Diverse models of this process have been formulated, motivated by their interesting mathematical properties or by their implications for turbulent mixing.

Here, results previously obtained by two methods, direct numerical simulation and mapping-closure analysis, are reassessed in the context of two new models that are proposed. One of the new models is based on the linear-eddy approach, in which mixing is simulated in one spatial dimension by means of a stochastic process that emulates turbulent advection [1]. The other new model is the clipped-laminar-profile representation (CLAPR), a geometrical construction that subsumes the results of mapping-closure analysis as a special case and generalizes the class of advection processes that is treated [2]. The comparison of the two previous and two new methods leads to new insights concerning both the mathematical properties of diffusion-advection and the phenomenology of turbulent mixing.

Before introducing the new models, the DNS and mapping closure results are summarized. The particular DNS configuration that is considered [3] involves a statistically steady advection process based on numerical solution of the Navier-Stokes equation with stochastic low-wavenumber forcing. Both the advection process and the initial scalar field are homogeneous and isotropic, to a sufficient approximation, within a periodic box. The probability density function (pdf) of the initial scalar field approximates the double-delta-function form $f(\phi;0) = (1/2)[\delta(\phi-1)+\delta(\phi+1)]$ corresponding to two initially unmixed streams. Under the influence of advection and diffusion (with Schmidt number $Sc = 0.7$), the pdf $f(\phi;t)$ evolves toward the large-t asymptote $\delta(0)$. The evolution of the pdf is the principal focus of the present investigation.

Mapping-closure analysis [4] has been used to predict the family of pdf shapes obtained during this evolution, parametrized by a time variable whose relation to physical time is undetermined [5]. The method of analysis is not discussed here, but a new geometrical interpretation of the result that is obtained is outlined shortly.

Figure 1 shows families of pdf shapes obtained by DNS and by mapping closure. Prior to the present investigation, no mechanistic interpretation of the good agreement between these families had been offered. The new results presented here suggest an interpretation and provide additional insights.

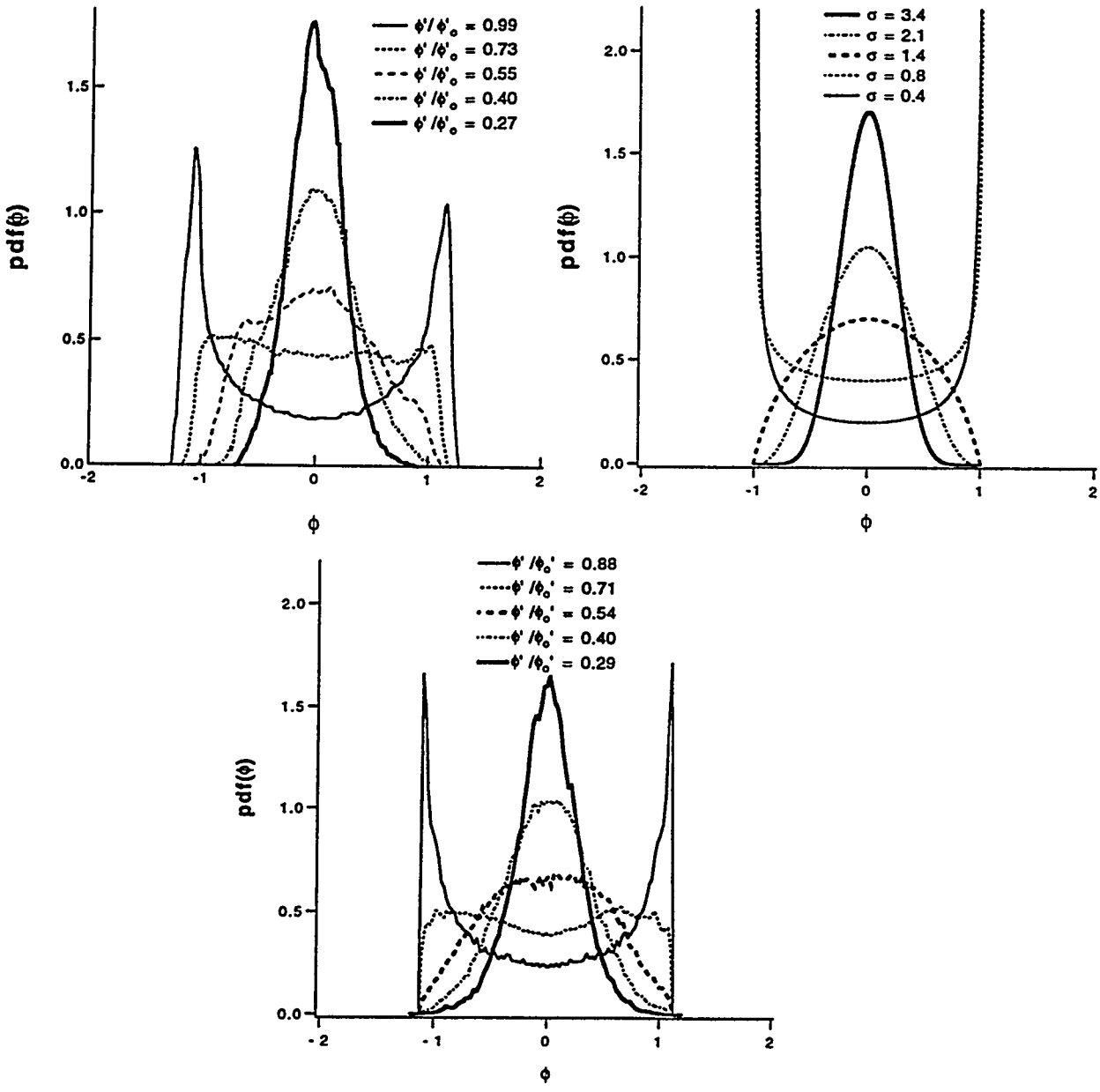


Figure 1. Scalar Pdf Families from Direct Numerical Simulation (Top Left; Parametrized by Rms Scalar Fluctuation ϕ' Normalized by its Initial Value ϕ'_o), Mapping Closure (Top Right; Parametrized by σ) and Single-Scale Linear-Eddy Simulation (Bottom; Parametrized by ϕ'/ϕ'_o)

CLIPPED-LAMINAR-PROFILE REPRESENTATION

A purely geometrical construction is used to obtain a family of pdf shapes. The mechanistic significance of this construction is inferred by comparing the results to those of other models.

A pdf of the scalar ϕ is obtained by taking the spatial domain to be a collection of line segments. The length $2w$ of a given segment is obtained by randomly sampling w from a pdf $h(w)$. Denoting spatial location within a segment by x, where $-w \le x \le w$, the spatial profile of ϕ on each segment is assigned to be

$$\phi(x) = \mathrm{erf}\left(\frac{x}{\sqrt{2}\sigma}\right), \tag{1}$$

where σ has the same value for all segments. For given $h(w)$ and σ, the pdf of ϕ for the scalar field thus defined is given by [6]

$$f(\phi;\sigma) = \frac{dx}{d\phi}\frac{1-H(x)}{2\int_0^1 wh(w)\,dw}, \tag{2}$$

where $H(w)$ is the cumulative distribution function (cdf) corresponding to $h(w)$ and the functional dependence $x(\phi)$ is determined implicitly by Eq. (1).

In Eq. (2), the pdf is parametrized by σ. Variation of σ generates a one-parameter family of pdf's, indicating that σ plays a role analogous to the time parameter of the mapping-closure analysis.

For $\sigma = 0$, the spatial profile on each segment reduces to $\phi(x) = \mathrm{sign}(x)$, so $f(\phi;0)$ is of double-delta-function form for any $h(w)$, corresponding to the initial condition for the mixing problem under consideration. In the limit $\sigma \to \infty$, $\phi(x)$ becomes identically zero, indicating that the physically correct final state is reached. For finite σ, the concentration profile on each segment corresponds to a solution of the diffusion equation on an infinite one-dimensional domain. This does not in itself justify the adoption of this functional form on clipped (finite) domains because it neglects interactions between neighboring interfaces, not to mention the effect of advection.

The motivation for the error-function *ansatz* is that it allows a result of mapping-closure analysis to be reproduced. Namely, if the cdf of the segment-length parameter w is taken to be $H(w) = 1 - \exp(-w^2)$, then Eqs. (1) and (2) yield a family of scalar pdf's that is identical to the mapping-closure family [6]. Moreover, by parametric variation of the functional form of $H(w)$, other families are obtained that may be compared to the mapping-closure family in order to identify the mechanistic origin of particular behaviors.

For this purpose, cdf's of the form

$$H(w) = 1 - \exp(-w^n) \tag{3}$$

are considered, where the case $n = 2$ corresponds to the mapping-closure result. The key property of Eq. (3) is that the bandwidth of $H(w)$ increases with decreasing n. The family of scalar pdf's corresponding to $n = 1$ is shown in Fig. 2. Unlike the pdf's in Fig. 1, this family exhibits a regime of trimodality during its transient evolution.

This behavior can be understood as follows. A wideband segment-size distribution corresponds, for intermediate σ values, to a large number of segments much smaller than σ and a few segments much larger than σ. The many small segments have ϕ profiles entirely contained within a small neighborhood of $\phi = 0$, therefore contributing a peak in the vicinity of $\phi = 0$ to the pdf of ϕ. The large segments, though few in number, contain a large enough fraction of the total scalar domain so that their contribution to the pdf is significant. On these segments, σ is so much smaller than w that

the scalar on these segments is predominantly unmixed, corresponding to the double-delta-function form of the scalar pdf. The combination of the two contributions accounts for the intermediate trimodality.

Though this reasoning relates properties of the scalar pdf to geometrical properties of CLAPR, it does not assure that these considerations are pertinent to turbulent mixing. A qualitative justification of this extension is as follows. High-Reynolds-number turbulence consists of eddies spanning a wide dynamic range, including large, slow eddies responsible for initial length-scale breakdown and numerous small eddies that rapidly complete the homogenization process. Due to intermittency, different localities within the flow will be at different stages of length-scale breakdown at a given instant. Some fluid parcels will have achieved a degree of breakdown such that rapid homogenization occurs, while others will be in a largely unmixed state.

Thus, a connection between trimodality and eddy dynamic range is plausible in the turbulent mixing context. The quantification of this connection by means of a mixing model is now considered.

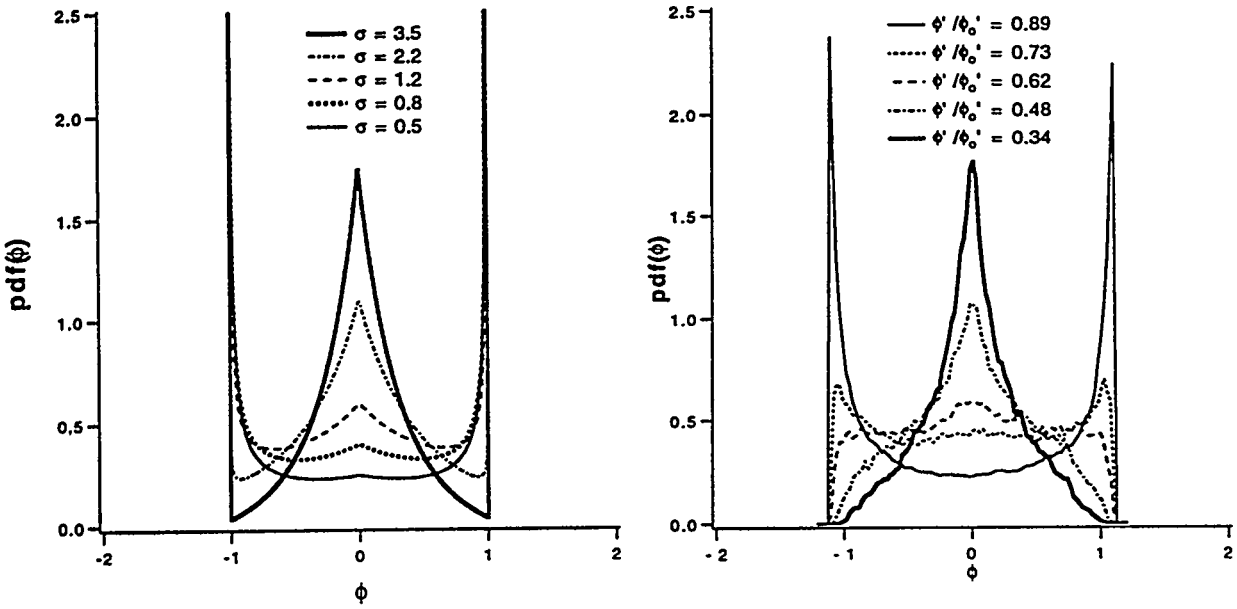


Figure 2. Scalar Pdf Families from CLAPR with $n = 1$ (Left) and Multi-Scale Linear-Eddy Simulation (Right)

LINEAR-EDDY MODEL

The linear-eddy model is a stochastic simulation of advection-diffusion formulated so as to capture the essential mechanistic features of the process in a computation that is affordable at high Reynolds number (Re). The modeling strategy is to maintain full spatial resolution, so that molecular diffusion can be correctly implemented, but to simplify the computation by reducing the problem to one spatial dimension. Since continuum flow cannot be implemented in one dimension (except for trivial Galilean transformations), advection is represented by a random sequence of instantaneous events. Each event involves spatial rearrangement of a randomly selected interval of the spatial domain. The rearrangement is a mapping, denoted the "triplet map," of the interval onto itself.

This mapping is conveniently defined as a two-step process. First, the scalar field within the chosen interval is replaced by three compressed copies of the original scalar field within the interval.

Each copy is compressed spatially by a factor of three so that the three copies fill the original interval. Second, the middle copy is spatially inverted ("flipped").

Graphical illustrations and a formal mathematical definition of this map, as well as a mechanistic rationale for this formulation, have been presented [1]. The triplet map captures, in one spatial dimension, the essential properties of the baker's map as commonly applied to mixing problems. In particular, a spatially homogeneous, statistically stationary sequence of triplet-map events induces exponential growth of material-surface area. In this and other respects, the rearrangement process emulates compressive strain effects in turbulent flow [1].

Operationally, the linear-eddy model is implemented as a Monte Carlo simulation. Molecular processes evolve in a conventional manner based on deterministic finite-difference solution of the governing equations. This deterministic evolution is punctuated by the stochastic rearrangement events.

In the simulations considered here, the size of the mapping interval either is fixed or is randomly selected for each rearrangement event based on a size-versus-frequency distribution corresponding to the eddy-size distribution in inertial-range turbulence. These are denoted the single-scale and multi-scale formulations, respectively. Viewing each mapping event as the model analog of an eddy, comparison of these formulations allows an assessment of the impact of eddy dynamic range on mixing properties.

Scalar field statistics are gathered from the simulation by running an ensemble of realizations for a given initial condition and mapping interval size-versus-frequency distribution. Each realization constitutes a scalar field time-history. Thus, the time evolution of quantities such as the scalar variance and scalar dissipation, as well as the family of pdf shapes, can be extracted.

For comparison to the DNS results, the initial scalar field and molecular diffusivity are assigned in conformance with the DNS calculation. For the multi-scale formulation, the size-versus-frequency distribution is determined by matching the turbulent diffusivity and dynamic range of the DNS flow field. For the single-scale formulation, the eddy size is chosen so that kurtosis of the linear-eddy scalar pdf converges to the same large-t asymptote as the DNS scalar pdf. The eddy frequency is chosen to match the DNS turbulent diffusivity. The rationale for this procedure, and additional details, are presented elsewhere [2,7].

The scalar pdf evolution obtained in this manner for the single-scale formulation is shown in Fig. 1. It is in close correspondence with the pdf evolution indicated by DNS and mapping closure. The multi-scale formulation yields the pdf evolution shown in Fig. 2. That formulation is seen to be in close correspondence with the CLAPR result for $n = 1$. Several aspects of these comparisons are noteworthy.

First, the linear-eddy results exhibit the relationship between trimodality and eddy dynamic range that was discussed in the previous section. This supports the mechanistic interpretation of trimodality as an intermittency effect arising in flows with disparate times scales governing initial length-scale breakdown and mixing completion, respectively. Moreover, the agreement between CLAPR and linear-eddy pdf families indicates that pdf evolution may be insensitive to other aspects of scalar field structure or evolution mechanisms.

Second, the connection between pdf evolution and dynamic range, in conjunction with the good agreement between DNS and single-scale linear-eddy results, suggest that the DNS calculation corresponds to a narrowband rather than a wideband mixing process. This inference is supported by additional comparisons of DNS and single-scale linear-eddy results. For example, scalar variance and scalar dissipation time-histories obtained from the single-scale linear-eddy simulations are in good agreement with corresponding time-histories obtained from the DNS calculation [2]. The multi-scale linear-eddy results do not agree as well [7], and the nature of the discrepancies suggests that

this is due to higher stirring bandwidth in the linear-eddy simulations than in the DNS calculations [2]. This implies that, for the moderate Re at which the DNS is implemented, it does not capture the qualitative features of mixing at high turbulence intensity. In particular, it does not exhibit intermittency effects associated with wideband mixing processes. Linear-eddy simulations indicate that stirring bandwidth influences not only the transient evolution but also the large-t asymptotic form of the pdf [2,7].

DISCUSSION

The comparisons among diverse representations of turbulent mixing indicate the likelihood of high-Re effects that are not apparent at Re values presently accessible using DNS. This raises the following question concerning the interpretation of DNS results. DNS may be viewed as a representation of a small region within a large flow, and as such, should capture the fine-scale properties of that flow (e.g., molecular mixing), however large the flow Re may be. This is true especially when scalar fluctuation length scales are small compared to the DNS box size, so that large-scale entrainment intermittency does not play a role in the mixing process. What, then, is the difference between mixing in a small region within a large flow and the DNS representation of mixing in that region?

The difference is a consequence of the statistics of the low-wavenumber forcing adopted in DNS computations. The low-wavenumber energy input rate corresponds to the mean energy dissipation rate of the flow, incorporating order-unity random fluctuations of the instantaneous input rate relative to the mean. In high-Re turbulence, however, the variability of the energy dissipation rate at a small scale corresponding to the DNS box size is much larger than the mean energy dissipation rate. This extreme variability is a manifestation of the concentration of energy in a small, ostensibly fractal subset of the flow volume [8].

To represent this extreme variability within DNS, it would be necessary to adopt a low-wavenumber forcing consisting of long periods of near quiescence interspersed with brief periods of intense activity. Though it is impractical to implement this computationally, a thought experiment readily indicates its consequences for mixing. If an ensemble of initial-value problems were computed in this manner, at intermediate times (i.e. several large-eddy turnover times based on the *mean* energy dissipation rate), most realizations within the ensemble would experience little advection and therefore would remain unmixed, but a few realizations would experience intense advection and become well mixed. Gathering the pdf over this postulated ensemble of DNS realizations, it is plausible that the pdf would more closely resemble the wideband results obtained using CLAPR and the linear-eddy model than the results of DNS as implemented to date.

As yet, there is no direct experimental or computational evidence of the validity of this conjecture. If this picture of turbulent mixing holds true, then it is necessary to incorporate the influence of fluctuations at all length scales in order to obtain a turbulent mixing model with the physically correct fluctuation properties. The linear-eddy model is an economical formulation that satisfies this requirement. For flows whose multidimensional large-scale structure cannot be represented within the linear-eddy framework, a modeling approach that is presently being pursued involves the use of the linear-eddy model as a subgrid model within large-scale flow computations such as large-eddy simulations [9,10].

ACKNOWLEDGEMENT

This research was supported by the U.S. Department of Energy, the National Science Foundation, and the Office of Naval Research.

REFERENCES

1. A. R. KERSTEIN, "Linear-Eddy Modeling of Turbulent Transport. Part 6. Microstructure of Diffusive Scalar Mixing Fields," *J. Fluid Mech.* 231, 361 (1991).

2. M. A. CREMER, P. A. MCMURTRY, and A. R. KERSTEIN, "Effects of Turbulence Length-Scale Distribution on Scalar Mixing in Homogeneous Turbulent Flow," *Phys. Fluids* A (submitted for publication).

3. V. ESWARAN AND S. B. POPE, "Direct Numerical Simulations of the Turbulent Mixing of a Passive Scalar," *Phys. Fluids* 31, 506 (1988).

4. H. CHEN, S. CHEN, AND R. H. KRAICHNAN, "Probability Distribution of a Stochastically Advected Scalar Field," *Phys. Rev. Lett.* 63, 2657 (1989).

5. S. B. POPE, "Mapping Closure for Turbulent Mixing and Reaction," *Theoret. Comput. Fluid Dyn.* 2, 255 (1991).

6. A. R. KERSTEIN, "Flapping Model of Scalar Mixing in Turbulence," *Phys. Fluids* A 3, 2838 (1991).

7. P. A. MCMURTRY, T. C. GANSAUGE, A. R. KERSTEIN, and S. K. KRUEGER, "Linear Eddy Simulations of Mixing in a Homogeneous Turbulent Flow," *Phys. Fluids* A 5, 1023 (1993).

8. U. FRISCH, P. SULEM and M. NELKIN, "A Simple Dynamical Model of Intermittent Fully Developed Turbulence," *J. Fluid Mech.* 87, 719 (1978).

9. P. A. MCMURTRY, S. MENON, and A. R. KERSTEIN, "A Linear Eddy Sub-Grid Model for Turbulent Reacting Flows: Application to Hydrogen-Air Combustion," in *Proc. 24th Intl. Sympos. on Combustion*, Combustion Institute, Pittsburgh (1992), p. 271.

10. S. MENON, P. A. MCMURTRY, and A. R. KERSTEIN, "A Linear-Eddy Mixing Model for Large-Eddy Simulation of Turbulent Combustion," in *Large-eddy Simulations of Complex Engineering and Geophysical Flows* (B. Galperin and S. A. Orszag, eds.), Cambridge Univ. Press, Cambridge, England (1993), Chapter 14.

ENSEMBLE PHASE AVERAGING FOR DISPERSE TWO-PHASE FLOWS

A. Prosperetti and D.Z. Zhang
Department of Mechanical Engineering, The Johns Hopkins University, Baltimore MD 21218

ABSTRACT

The ensemble-averaging methods used in this paper differ from conventional ones in the use of phase averages (in which only configurations such that the point of interest is in a specified phase are considered) and of averages over quantities defined at particle centers. The first feature leads to rigorous equations of motion of the general two-fluid type, and the second one renders a straightforward derivation of the disperse momentum equation possible even in the presence of degenerate constitutive relations for the particle material. We present numerous applications of this method to: (*i*) rigid particles in inviscid incompressible flow; (*ii*) spherical bubbles in inviscid incompressible flow; (*iii*) rigid particles in Stokes flow; (*iv*) heat transfer in high-diffusivity fluids. While most of the results are for the dilute limit, some numerical results for finite volume fractions are presented for the linear problem.

INTRODUCTION

In Ref. [1] we have developed a method to derive disperse two-phase flow equations by ensemble averaging over the individual phases. The results are useful to obtain approximate closures in the dilute limit, to interpret numerical simulations at finite volume fractions, and possibly to derive approximate closures. Here we give a very brief summary of the method, and illustrate its capabilities by means of several applications, most of which are the result of work in progress.

Some our results coincide with others available in the literature while others are new. In any event, an important point that we wish to stress is that our approach provides a *rigorous unifying framework* by which a large number of disparate results can be systematically obtained without *ad hoc* approximations.

Among the advantages of the present approach one may mention the fact that the phasic averaging leads directly to a two-fluid formulation of the type generally used in the engineering literature. The disperse-phase equations are derived without recourse to the artifices usually needed when dealing with degenerate equations of state, such as negligible density. The closure problem presents itself in the form of *computable* quantities, for which approximate methods or direct numerical simulation are suitable. Finally, as compared with the calculation of *averaged properties*, such as the effective viscosity or thermal conductivity, the present approach has the virtue of systematically providing *all the terms* in the equations.

AVERAGING RELATIONS

A detailed exposition of the averaging techniques and theorems used in this paper is given in I. Here we summarize the essential results omitting the proofs.

Consider an ensemble of two-phase flows in which each realization contains N particles, drops, or bubbles arranged in a time-dependent configuration $\mathcal{C}^N$. We use this word and symbol in a technical sense as short-hand for the set of position vectors of the particles' centers, $\mathbf{y}^\alpha$, $\alpha = 1, 2, \ldots, N$, their translational velocities $\mathbf{w}^\alpha$, their radii a^α, and any other set of parameters necessary for a complete specification. Thus, for each member of the ensemble, with suitable initial conditions and dynamical equations for the particles and the continuous phase, and conditions "at infinity" for the latter, the exact microscopic problem is uniquely specified given the initial configuration.

For brevity here we develop the equations only for equal rigid particles and therefore a configuration of the system is given by a set of positions $\mathbf{y}^\alpha$ of the particle centers, a set Ψ^α of orientations of a system of particle axes with respect to the laboratory frame, a set of translational velocities $\mathbf{w}^\alpha$, and a set of rotational velocities Ω^α. We refer to the first two sets as generalized positions, denoted by the symbol $\mathbf{q}^\alpha$, and to the second pair of sets as generalized momenta, denoted by $\mathbf{p}^\alpha$. The case of

variable radius will be dealt with in the examples. To simplify the formulae the rotational degrees of freedom are not explicitly indicated although, of course, they are accounted for where necessary.

Let $P(N;t)$ be the probability of a specific configuration $\mathcal{C}^N$ at time t. This quantity evolves in time according to

$$\frac{\partial P}{\partial t} + \sum_{\alpha=1}^{N} [\nabla_{\mathbf{q}^\alpha} \cdot (\dot{\mathbf{q}}^\alpha P) + \nabla_{\mathbf{p}^\alpha} \cdot (\dot{\mathbf{p}}^\alpha P)] = 0. \tag{1}$$

We take the particles to be indistinguishable, so that $N!$ different arrangements of the N particles correspond to each physically distinguishable state of the system. It is convenient therefore to normalize the integral of P to $N!$.

Let $\chi_C(\mathbf{x};N)$ and $\chi_D(\mathbf{x};N)$ be the indicator functions of the continuous and disperse phases so that, for example, $\chi_C = 1$ when, with the particles in the configuration $\mathcal{C}^N$, the point $\mathbf{x}$ is in the continuous phase, and equals 0 otherwise. Note that these are geometrical entities that depend on time only indirectly through the time evolution of the configuration. The volume fractions of the two phases are defined by

$$\beta_{C,D} = \frac{1}{N!} \int d\mathcal{C}^N P(N;t) \chi_{C,D}(\mathbf{x},N). \tag{2}$$

The particle boundary is assumed to have zero measure so that $\chi_C + \chi_D = 1$ and, as a consequence, $\beta_C + \beta_D = 1$.

For equal rigid spheres of radius a one has

$$\chi_D(\mathbf{x};N) = \sum_{\alpha=1}^{N} H(a - |\mathbf{x} - \mathbf{y}^\alpha|) = 1 - \chi_C(\mathbf{x};N), \tag{3}$$

with H the Heaviside distribution. With this expression the definition (2) gives

$$\beta_D(\mathbf{x},t) = \int_{|\mathbf{x}-\mathbf{y}|\leq a} d^3y\, n(\mathbf{y},t), \qquad n(\mathbf{y},t) = \int d^2\Psi \int d^3w\, P(1;t). \tag{4}$$

Here n is the particle number density. It will be noted that, in the presence of gradients of the particle distribution, the relation $\beta_D = n\,v$, with $v = \frac{4}{3}\pi a^3$ the particle volume, is not strictly valid.

The phase ensemble averages for a field $f_{C,D}(\mathbf{x},t;N)$ pertaining to the continuous or disperse phase are defined by averaging over all the configurations such that the point $\mathbf{x}$ is in the appropriate phase:

$$< f_{C,D} > (\mathbf{x},t) = \frac{1}{\beta_{C,D} N!} \int d\mathcal{C}^N f_{C,D}(\mathbf{x},t;N)\, \chi_{C,D}(\mathbf{x};N)\, P(N;t). \tag{5}$$

A similar definition is used for the conditional averages, i.e. averages over all configurations such that some particles occupy a specified configuration.

By the same method used in [1] it is readily shown that

$$\nabla(\beta_C < f_C >) = \beta_C < \nabla f_C > + \int_{|\mathbf{x}-\mathbf{y}|=a} dS_y\, \mathbf{n} \int d^3w\, P(1;t) < f_C >_1 (\mathbf{x},t|1). \tag{6}$$

Here where $< f_C >_1$ denotes the conditional average with one particle fixed. Similarly, for the time derivative,

$$\frac{\partial}{\partial t}(\beta_C < f_C >) = \beta_C < \frac{\partial f_C}{\partial t} > - \int_{|\mathbf{x}-\mathbf{y}|=a} dS_y \int d^3w\, \mathbf{w}\cdot\mathbf{n}\, P(1,t) < f_C >_1, \tag{7}$$

and for the Laplacian operator,

$$\begin{aligned} \nabla^2(\beta_C < f_C >) = \beta_C < \nabla^2 f_C > \; &+ \; \nabla \cdot \int_{|\mathbf{x}-\mathbf{y}|=a} dS_y \mathbf{n} \int d^3w P(1;t) < f_C >_1 \\ &+ \; \int_{|\mathbf{x}-\mathbf{y}|=a} dS_y \cdot \mathbf{n} \int d^3w P(1;t) < \nabla f_C >_1 . \end{aligned} \tag{8}$$

In the present approach averaging and differentiation do not commute as in the methods where averaging is carried out irrespective of the phase occupying the position $\mathbf{x}$ [2]. The counterpart for this technical inconvenience is a much greater flexibility in the type of quantities that can be averaged as they do not necessarily have to be defined in both phases.

The integrals in (6) and (7) are over all the particles touching $\mathbf{x}$ for all of which the velocity field satisfies the kinematic boundary condition $\mathbf{n}\cdot\mathbf{w} = \mathbf{n}\cdot\mathbf{u}_C(\mathbf{x}, t|N)$. Therefore, upon combining the two relations, the two integrals cancel so that

$$\frac{\partial(\beta_C < f_C >)}{\partial t} + \nabla\cdot(\beta_C < f_C\mathbf{u}_C >) = \beta_C < \frac{\partial f_C}{\partial t} + \nabla\cdot(f_C\mathbf{u}_C) > . \tag{9}$$

Thus, though not commuting, averaging and convective differentiation satisfy a simple relation that plays a central role in the derivation of the averaged equations of the following section.

The general definition of averaging for the disperse phase is given above in (5). This type of average is however of limited usefulness (see [1]). We make extensive use, rather, of a different kind of average appropriate for quantities $g^{(\alpha)}(t; N)$ defined at particle centers. Examples are the center-of-mass velocity, momentum, radius, and others. For such quantities we define the ensemble average over all the configurations such that the center of one particle is at $\mathbf{y}$ at time t. On the assumption that $g^{(1)}$ does not depend explicitly on the configuration of the other particles, the definition is

$$\overline{g_D}(\mathbf{y}, a, t) = \frac{1}{n}\int d^3w\, P(1;t)\; g^{(1)}(\mathbf{q}, \mathbf{p}, t). \tag{10}$$

As in [1], it is easy to show for such particle-centers averaged quantities that

$$\frac{\partial}{\partial t}(n\,\overline{g_D}) + \nabla_y\cdot(n\,\overline{\mathbf{w}\,g_D}) \;=\; n\,\overline{\frac{\partial g^{(1)}}{\partial t}}, \tag{11}$$

where $\partial g^{(1)}/\partial t$ is the time derivative following the motion of the particles. Since the field $g^{(1)}$ is defined only at particle centers, this result is the same as in kinetic theory, where the finite size of the molecules is disregarded. We now use the previous relations to derive the equations of motion of the phases.

AVERAGE EQUATIONS

With the neglect of compressibility, the equations of motion of the continuous phase are

$$\nabla\cdot\mathbf{u}_C = 0, \tag{12}$$

$$\frac{\partial\mathbf{u}_C}{\partial t} + \nabla\cdot(\mathbf{u}_C\mathbf{u}_C) = -\frac{1}{\rho_C}\nabla p_C + \nu_C\nabla^2\mathbf{u}_C + \mathbf{g}. \tag{13}$$

Upon taking $f_C = 1$ in (9) and using the continuity equation (12) we find the averaged continuity equation for the continuous phase as

$$\frac{\partial\beta_C}{\partial t} + \nabla\cdot(\beta_C < \mathbf{u}_C >) = 0. \tag{14}$$

Similarly, upon taking $f_C = \mathbf{u}_C$, we have from (9) and (13)

$$\begin{aligned}\rho_C\frac{\partial}{\partial t}(\beta_C < \mathbf{u}_C >) + \rho_C\nabla\cdot(\beta_C < \mathbf{u}_C >< \mathbf{u}_C >) = -\beta_C\nabla < p_C > \\ +\;\; \mu_C\nabla^2(\beta_C < \mathbf{u}_C >) + \beta_C\mathbf{A}_C(\mathbf{x},t) + \beta_C\mu_C\mathbf{B}_C + \rho_C\nabla\cdot(\beta_C\mathbf{M}_C) + \beta_C\rho_C\mathbf{g},\end{aligned} \tag{15}$$

where we have introduced the "kinematic" Reynolds stress tensor

$$\mathbf{M}_C = < \mathbf{u}_C >< \mathbf{u}_C > - < \mathbf{u}_C\mathbf{u}_C > = -< (\mathbf{u}_C - < \mathbf{u}_C >)(\mathbf{u}_C - < \mathbf{u}_C >) >, \tag{16}$$

and we have set

$$\mathbf{A}_C(\mathbf{x}, t) = \nabla < p_C > - < \nabla p_C >, \tag{17}$$

or, from (6),

$$\beta_C\mathbf{A}_C(\mathbf{x}, t) = \int_{|\mathbf{x}-\mathbf{y}|=a} dS_y\,\mathbf{n}\int d^3w\, P(1,t)[< p_C >_1 (\mathbf{x}, t|1) - < p_C > (\mathbf{x}, t)]\,. \tag{18}$$

Furthermore,

$$\beta_C\mathbf{B}_C = -\nabla\cdot\int_{|\mathbf{x}-\mathbf{y}|=a} dS_y\,\mathbf{n}\,P(1;t)\; < \mathbf{u}_C >_1 - \int_{|\mathbf{x}-\mathbf{y}|=a} < \nabla\mathbf{u}_C >_1 \cdot\mathbf{n}P(1;t)dS_y. \tag{19}$$

Equation (15) has been written in terms of the gradient of the average pressure as in the standard two-fluid models. This choice introduces the quantity $\mathbf{A}_C$ which must evidently contain all the local phase momentum interactions. This procedure is similar to that of [3] but, rather than relying on separation of scale arguments, has been effected by introducing a clear operational prescription that enables the phase interaction term $\mathbf{A}_C$ to be evaluated. An explicit expression for this quantity valid in the dilute limit is given later.

If we were to follow the same procedure to derive the averaged equations for the disperse phase, the result would involve the constitutive relations of that phase's material. This is undesirable for situations where such a relation is degenerate, as for rigid particles or massless bubbles. In order to avoid this problem, we use the particle-centers averages defined in (10). We start with the number density of the particles. By setting $g^{(1)} = 1$ in (11) we have

$$\frac{\partial n}{\partial t} + \nabla_y \cdot (n\,\overline{\mathbf{w}}) = 0, \tag{20}$$

from which a conservation equation for the particle volume or mass (assumed constant) is readily obtained. In the case of variable v or m, the appropriate procedure is to set $g^{(1)} = v$ or m in (11).

Similarly, by setting $g^{(1)} = m\mathbf{w}$, we find the averaged particle momentum equation in the form

$$\frac{\partial}{\partial t}(n\,\overline{m\,\mathbf{w}}) + \nabla_y \cdot (n\,\overline{m\mathbf{w}\mathbf{w}}) = n\,\overline{m\,\dot{\mathbf{w}}}. \tag{21}$$

The last term in the right-hand side may be rewritten by recalling the equation of motion of the particle center of mass which, neglecting collisions, is

$$m\dot{\mathbf{w}} = \int_{|\mathbf{x}-\mathbf{y}|=a} [-p_C(\mathbf{x},t;N)\,\mathbf{n} + \tau_C(\mathbf{x},t;N)\cdot\mathbf{n}]\,dS_x + m\,\mathbf{g}, \tag{22}$$

where τ_C is the viscous stress tensor. The integral is over the surface of the particle centered at $\mathbf{y}$. Upon substituting this expression into the the right-hand side of (21) we find

$$\frac{\partial n\overline{m}\,\overline{\mathbf{w}}}{\partial t} + \nabla_y \cdot (n\,\overline{m\,\mathbf{w}\,\mathbf{w}}) = \int_{|\mathbf{x}-\mathbf{y}|=a} dS_x \mathbf{n} \cdot \int d^3w < -p_C\mathbf{I} + \tau_C >_1 (\mathbf{x},t|1)P(1;t) + \mathbf{g}, \tag{23}$$

or, with the definitions

$$\mathbf{M}_D = \overline{\mathbf{w}}\,\overline{\mathbf{w}} - \overline{\mathbf{w}\,\mathbf{w}} = -\overline{(\mathbf{w}-\overline{\mathbf{w}})(\mathbf{w}-\overline{\mathbf{w}})}, \tag{24}$$

$$\begin{aligned}\mathbf{A}_D &= -\frac{1}{nv}\int_{|\mathbf{x}-\mathbf{y}|=a} \mathbf{n}\,dS_x \int d^3w\,P(1;t)\quad [< p_C >_1 (\mathbf{x},t|1) - < p_C > (\mathbf{x},t)] \\ &+ \frac{1}{v}\left[v\nabla < p_C > - \int_{|\mathbf{x}-\mathbf{y}|=a} \mathbf{n} < p_C > (\mathbf{x},t)dS_x\right],\end{aligned} \tag{25}$$

and $\rho_D = m/v$,

$$\rho_D \frac{\partial n\overline{\mathbf{w}}}{\partial t} + \rho_D \nabla_y \cdot (n\,\overline{\mathbf{w}\mathbf{w}}) = -n\nabla_y < p_C > + \rho_D \nabla\cdot(n\mathbf{M}_D) + n\mathbf{A}_D + n\rho_D\mathbf{g} + \beta_D\mathbf{B}_D. \tag{26}$$

SMALL-PARTICLE APPROXIMATION

The equations derived in the preceding sections contain several terms involving integration over spheres with a radius equal to the particle radius a. When the macroscopic quantities vary slowly over this scale, these integrals may be approximated by Taylor series expansions. The detailed calculations are given in I. Here we simply note the pertinent results.

The volume fraction β_D and average velocity $< \mathbf{u}_D >$ of the disperse phase are approximately given by

$$\beta_D(\mathbf{x},t) = v\,n(\mathbf{x},t)\,[1 + O(a^2/L^2)], \qquad < \mathbf{u}_D > (\mathbf{x},t) = \overline{\mathbf{w}}(\mathbf{x},t) + O(a/L), \tag{27}$$

where L is the characteristic length for the variation of macroscopic quantities and $v = \frac{4}{3}\pi a^3$ is the particle volume. To the same accuracy, the quantity $\mathbf{A}_D$ introduced in (25) is given by

$$\beta_D\mathbf{A}_D(\mathbf{y},t) = -\int\int d^3w\,P(\mathbf{y},\mathbf{w};t)\int_{|\mathbf{x}-\mathbf{y}|=a} dS_x\,\mathbf{n}[< p_C >_1 (\mathbf{x},t|\mathbf{y},\mathbf{w}) - < p_C > (\mathbf{x},t)]. \tag{28}$$

Notice that, even after this approximation, this quantity remains different from $\mathbf{A}_C$ defined in (17) because the integration is carried out over the surface of a particle centered at $\mathbf{y}$ rather than over the

particles with center at $\mathbf{y}$ such that $|\mathbf{y}-\mathbf{x}| = a$. It can be shown that, with an error of order $\beta_D a^2/L^2$ [1],

$$\beta_C \mathbf{A}_C = -\beta_D \mathbf{A}_D(\mathbf{x},t) + \nabla \cdot \beta_D \mathbf{T}_C, \tag{29}$$

where

$$\mathbf{T}_C(\mathbf{x},t) = -\frac{a}{\beta_D}\int d^3w \int_{|\mathbf{y}-\mathbf{x}|=a} dS_y \, \mathbf{n}\,\mathbf{n}\, P(1;t)[< p_C >_1 (\mathbf{y},t|1) - < p_C > (\mathbf{y},t)]. \tag{30}$$

It will be seen in the following that $\mathbf{T}_C$ affects the average motion of the continuous phase in the same way as an additional stress tensor.

If $\mathbf{A}_C$ and $\mathbf{A}_D$ were to be considered as "internal" forces of the mixture, then, on the basis of a naive interpretation of the action-reaction principle of Mechanics, from the momentum equations (15), (26), one would expect $\mathbf{T}_C$ to be zero. However, in the next sections we give examples of explicit calculations of $\mathbf{T}_C$ that give non-zero results. The reason is that, while the action-reaction principle must of course hold at the microscopic level, there is no reason for it to hold for the averaged quantities $\mathbf{A}_C$ and $\mathbf{A}_D$ which are evaluated by integration over neighboring, but different finite surfaces. For uniform mixtures, however, $\nabla \cdot \mathbf{T}_C = 0$ and $\beta_C \mathbf{A}_C = -\beta_D \mathbf{A}_D$ as expected. Hence the finite size of the particles is seen to play a crucial role in the derivation of (29).

With the previous approximations the averaged equations of the previous section may be written in a slightly more familiar form as follows. For the continuous phase, Eq. (14) is unchanged and will not be repeated. The momentum equation (15) becomes instead

$$\begin{aligned} &\rho_C \frac{\partial}{\partial t}(\beta_C < \mathbf{u}_C >) + \rho_C \nabla \cdot (\beta_C < \mathbf{u}_C >< \mathbf{u}_C >) \\ &= -\beta_C \nabla < p_C > -\beta_D \mathbf{A}_D(\mathbf{x},t) + \nabla \cdot (\beta_C \rho_C \mathbf{M}_C + \mathbf{T}_C) + \beta_C \rho_C \mathbf{g}, \end{aligned} \tag{31}$$

With (27), the number density conservation equation (20) becomes

$$\frac{\partial \beta_D}{\partial t} + \nabla_y \cdot (\beta_D \overline{\mathbf{w}}) = 0, \tag{32}$$

while the momentum equation (26) remains unchanged.

THE DILUTE LIMIT

The set of averaged equations derived before is not closed as the right-hand sides contain averages with one particle held fixed and averages of products that must somehow be related to products of averages. This is the well-known closure problem that is unavoidable in any averaging method.

The simplest closure can be effected in the dilute limit (i.e. with an $O(\beta_D)$ accuracy) as follows. The integrals containing the conditional averages $< \ldots >_1$ in the right-hand sides of the unconditionally averaged equations give a contribution of $O(\beta_D)$. For accuracy to this order, therefore, these conditionally averaged fields are only needed with $O(1)$ accuracy. To this accuracy the conditionally averaged equations are a closed system and can be solved subject to the condition of matching to the unconditionally averaged fields at large distances from the particle held fixed. The solution thus obtained can be substituted into the right-hand sides of the unconditionally averaged equations to give a closed set.

In principle this procedure can be applied to find averaged equations correct to $O(\beta_D^K)$ for arbitrary K, although for most problems $K = 2$ appears to be already the practical limit or possibly beyond. An alternative is to use direct numerical simulation. Finally, one may attempt to find heuristic approximations to calculate the conditional fields.

We now present several examples of dilute-limit and numerical closures. Since from now on all quantities are averaged, we drop the averaging symbols. Furthermore, we write $\mathbf{u}_D$ in place of $\overline{\mathbf{w}}$ which is permissible to $O(\beta_D)$ on the basis of (27).

For all the examples discussed below the continuity equations (14), (32) are necessary to close the system and will not be repeated.

RIGID PARTICLES IN POTENTIAL FLOW

For rigid particles in inviscid flow the rotational degrees of freedom are unnecessary and the particle configuration only depends on $\mathbf{y}^\alpha$ and $\mathbf{w}^\alpha$, $\alpha = 1,2,\ldots,N$. In [1] we have found the results

$$\mathbf{A}_D = \frac{1}{2}\rho_C \left[\frac{\partial \mathbf{u}_C}{\partial t} + \mathbf{u}_C \cdot \nabla \mathbf{u}_C - \frac{\partial \mathbf{u}_D}{\partial t} - \mathbf{u}_D \cdot \nabla \mathbf{u}_D + \frac{1}{\beta_D}\nabla(\beta_D \mathbf{M}_D)\right], \tag{33}$$

$$\mathbf{T}_C = \frac{1}{5}\rho_C \left\{\beta_D \left[2(\mathbf{u}_C - \mathbf{u}_D)^2 \mathbf{I} - \frac{9}{4}(\mathbf{u}_C - \mathbf{u}_D)(\mathbf{u}_C - \mathbf{u}_D)\right]\right\} - \frac{1}{5}\rho_C \beta_D \left[2(\mathrm{Tr}\ \mathbf{M}_D)\mathbf{I} - \frac{9}{4}\mathbf{M}_D\right], \tag{34}$$

$$\mathbf{M}_C = -\frac{\beta_D}{20}\left[3(\mathbf{u}_C-\mathbf{u}_D)\cdot(\mathbf{u}_C-\mathbf{u}_D)\mathbf{I}+(\mathbf{u}_C-\mathbf{u}_D)(\mathbf{u}_C-\mathbf{u}_D)\right]+\frac{\beta_D}{20}\left[3\,(\mathrm{Tr}\;\mathbf{M}_D)\,\mathbf{I}+\mathbf{M}_D\right]. \quad (35)$$

Aside from the Reynolds-stress-like term $\mathbf{M}_D$, the result for $\mathbf{A}_D$ agrees with the expression of the force on a single sphere immersed in a flow obtained by several researchers and, most recently, in [4]. The first group of terms in $\mathbf{M}_C$ is the same as the result obtained in [5] where however the particles were implicitly taken to move with the same velocity so that $\mathbf{M}_D = 0$.

Since the previous results have been obtained on the assumption of potential flow, they cannot contain the effect of a rotation. If frame indifference is postulated, however, they can easily be corrected for this effect as shown in [1] with the result $\mathbf{A}'_D = \mathbf{A}_D + \frac{1}{2}\rho_C(\nabla\times\mathbf{u}_C)\times(\mathbf{u}_D-\mathbf{u}_C)$, where the first term in the right-hand side is given by (33). The corresponding expression of $\mathbf{A}'_C$ follows from (29) with $\mathbf{A}'_D$ in place of $\mathbf{A}_D$. The additional terms are the familiar lift force [4]. This result has been obtained by assuming that a potential flow is viewed from a rotating coordinate frame. Hence, strictly speaking, it is only valid provided $\nabla\times\mathbf{u}_C$ has a spatially uniform value throughout the flow field. However, with the usual assumptions about scale separation, it may be considered as approximately applicable to more general situations as well.

Combining the previous results we have the final form of the momentum equations. For the continuous phase the result is

$$\rho_C\beta_C\left[\frac{\partial\mathbf{u}_C}{\partial t}+(\mathbf{u}_C\cdot\nabla)\mathbf{u}_C\right]+\beta_C\nabla p_C = -\frac{1}{2}\rho_C\beta_D\left[\frac{\partial\mathbf{u}_C}{\partial t}+\mathbf{u}_C\cdot\nabla\mathbf{u}_C-\frac{\partial\mathbf{u}_D}{\partial t}-\mathbf{u}_D\cdot\nabla\mathbf{u}_D\right]$$
$$-\frac{1}{2}\rho_C\beta_D(\nabla\times\mathbf{u}_C)\times(\mathbf{u}_D-\mathbf{u}_C)+\frac{1}{4}\rho_C\nabla\cdot\left\{\beta_D\left[(\mathbf{u}_C-\mathbf{u}_D)^2\mathbf{I}-2(\mathbf{u}_C-\mathbf{u}_D)(\mathbf{u}_C-\mathbf{u}_D)\right]\right\}$$
$$-\frac{1}{4}\rho_C\nabla\left[\beta_D(\mathrm{Tr}\,\mathbf{M}_D)\right]+\beta_C\rho_C\mathbf{g}. \quad (36)$$

Similarly, the final form of the disperse-phase momentum equation is

$$\rho_D\beta_D\left[\frac{\partial\mathbf{u}_D}{\partial t}+(\mathbf{u}_D\cdot\nabla)\mathbf{u}_D\right]+\beta_D\nabla p_C = \frac{1}{2}\rho_C\beta_D\left[\frac{\partial\mathbf{u}_C}{\partial t}+\mathbf{u}_C\cdot\nabla\mathbf{u}_C-\frac{\partial}{\partial t}\mathbf{u}_D-\mathbf{u}_D\cdot\nabla\mathbf{u}_D\right]$$
$$+\frac{1}{2}\rho_C\beta_D(\nabla\times\mathbf{u}_C)\times(\mathbf{u}_D-\mathbf{u}_C)+(\rho_D+\frac{1}{2}\rho_C)\nabla\cdot(\beta_D\mathbf{M}_D)+\beta_D\rho_D\mathbf{g}. \quad (37)$$

The final closure of the system requires an expression for the fluctuating particle volume flux tensor $\mathbf{M}_D$ [6]. This missing information cannot be supplied internally by the theory without a specification of the initial conditions imposed on the particle probability distribution. This point was noted in [7] where it was explicitly assumed that, at each position and time, the particle velocity probability distribution is strongly peaked around its local, instantaneous mean value. In this case $\mathbf{M}_D = 0$. A similar assumption – whether explicit or implict – seems to be present in all of the previous work.

It is shown in [1] that the preceding expressions coincide with those of [8-10] to the present $O(\beta_D)$.

SPHERICAL BUBBLES IN POTENTIAL FLOW

For massless spherical bubbles in inviscid flow the rotational degrees of freedom are again unncessary but the particles instantaneous radii and radial velocities must be added to the set of variables that define a configuration.

In a paper in preparation we study this problem allowing for a spectrum of bubble sizes. Here, for simplicity, we quote only the results applicable when the probability distribution is strongly peaked about an average radius $\overline{a}$ and an average radial velocity $\overline{\dot{a}}$. Again upon dropping the averaging symbols, the momentum equation for the bubbles is

$$\beta_D\nabla p_C = \frac{1}{2}\rho_C\beta_D\left[\frac{\partial\mathbf{u}_C}{\partial t}+\mathbf{u}_C\cdot\nabla\mathbf{u}_C-\frac{\partial\mathbf{u}_D}{\partial t}-\mathbf{u}_D\cdot\nabla\mathbf{u}_D\right]$$
$$+\frac{3}{2a}\rho_C\beta_D(\mathbf{u}_C-\mathbf{u}_D)\left(\frac{\partial a}{\partial t}+\mathbf{u}_D\cdot\nabla a\right)$$
$$+\frac{1}{2}\rho_C\beta_D(\nabla\times\mathbf{u}_C)\times(\mathbf{u}_D-\mathbf{u}_C)+\frac{1}{2}\rho_C\nabla\cdot(\beta_D\mathbf{M}_D). \quad (38)$$

For the continuous phase

$$\rho_C\beta_C\left[\frac{\partial \mathbf{u}_C}{\partial t}+(\mathbf{u}_C\cdot\nabla)\mathbf{u}_C\right]+\beta_C\nabla p_C=\text{same as rhs of (36)}$$
$$+\ \frac{3}{2a}\rho_C\beta_D(\mathbf{u}_D-\mathbf{u}_C)\left(\frac{\partial a}{\partial t}+\mathbf{u}_D\cdot\nabla a\right)$$
$$-\ \frac{1}{4}\rho_C\nabla\cdot\left[\beta_D a\left(\frac{\partial}{\partial t}+\mathbf{u}_D\cdot\nabla\right)\left(\frac{\partial}{\partial t}+\mathbf{u}_D\cdot\nabla\right)a-\frac{5}{2}\beta_D\left(\frac{\partial a}{\partial t}+\mathbf{u}_D\cdot\nabla a\right)^2\right]. \tag{39}$$

The equation of motion for the bubble radius generalizes the well-known Rayleigh-Plesset equation in the form

$$a\left(\frac{\partial}{\partial t}+\mathbf{u}_D\cdot\nabla\right)\left(\frac{\partial}{\partial t}+\mathbf{u}_D\cdot\nabla\right)a+\frac{3}{2}\left(\frac{\partial a}{\partial t}+\mathbf{u}_D\cdot\nabla a\right)^2$$
$$=\ \frac{1}{\rho_C}\left[p_B-\frac{2\sigma}{a}-p_C\right]-\frac{1}{4}(\mathbf{u}_D-\mathbf{u}_C)\cdot(\mathbf{u}_D-\mathbf{u}_C)+\frac{1}{4}\mathrm{Tr}\,\mathbf{M}_D. \tag{40}$$

RIGID PARTICLES IN STOKES FLOW

In this case the probability distribution only depends on the position and angular orientation of the particles. One finds

$$\beta_C\nabla p_C=\nabla\cdot(\mu^*\nabla\mathbf{u}_m)+\frac{9}{2}\frac{\beta_D}{a^2}\mu(\mathbf{u}_D-\mathbf{u}_C) \tag{41}$$

where μ^* is the effective viscosity $1+5/2\,\beta_D$ and $\mathbf{u}_m=\beta_D\mathbf{u}_D+\beta_C\mathbf{u}_C$ the volumetric velocity. The equation of motion for the particles is

$$\rho_D\left(\frac{\partial \mathbf{u}_D}{\partial t}+\mathbf{u}_D\cdot\nabla\mathbf{u}_D\right)=-\nabla p_C-\frac{9}{2}\frac{\mu}{a^2}(\mathbf{u}_D-\mathbf{u}_C). \tag{42}$$

THE ENERGY EQUATION

The same averaging method may be applied to the energy equation for an incompressible fluid $\partial T_C/\partial t+\mathbf{u}_C\cdot\nabla T_C=D_C\nabla^2T_C$, where T_C is the temperature and D the thermal diffusivity. Upon averaging the result is

$$\frac{\partial\beta_C T_C}{\partial t}+\nabla\cdot(\beta_C\mathbf{u}_C T_C)=\nabla\cdot(D_C^*\nabla T_C)+\frac{3\beta_D D_C}{a^2}\left(1+\frac{Pe}{4}\right)(T_D-T_C)$$
$$+\ D_C\nabla^2[\beta_D(T_C-T_D)]+\frac{3}{16}\left(1-\frac{K_D}{K_C}\right)\nabla\cdot[\beta_D(T_D-T_C)(\mathbf{u}_D-\mathbf{u}_C)], \tag{43}$$

where K_D, K_C are the thermal conductivities, D_C^* is the effective thermal diffusivity

$$D_C^*=D_C\left[1+\frac{3\beta_D(K_D-K_C)}{K_D+2K_C}\right], \tag{44}$$

and Pe the particle Péclet number $Pe=2a|\mathbf{u}_D-\mathbf{u}_C|/D_C$.

Upon averaging the microscopic energy equation for the particles $K_D\nabla^2T_D+\dot{q}_D=\rho_D C_{pD}\partial T_D/\partial t$, where $\dot{q}_D$ is the internal volumetric heat source and C_{pD} teh specific heat, one finds the averaged energy equation for the disperse phase in the form

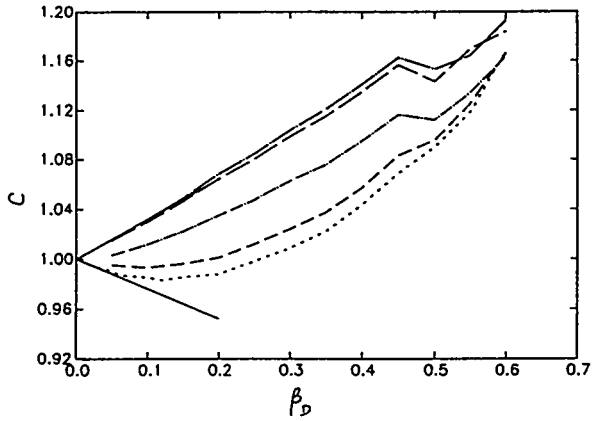
$$\frac{\partial}{\partial t}(\beta_D T_D)+\nabla(\beta_D\mathbf{u}_D T_D)=-\frac{3\beta_D K_C}{a^2C_{pD}\rho_D}\left(1+\frac{Pe}{4}\right)(T_D-T_C)+\frac{\beta_D}{C_{pD}\rho_D}\dot{q}_D. \tag{45}$$

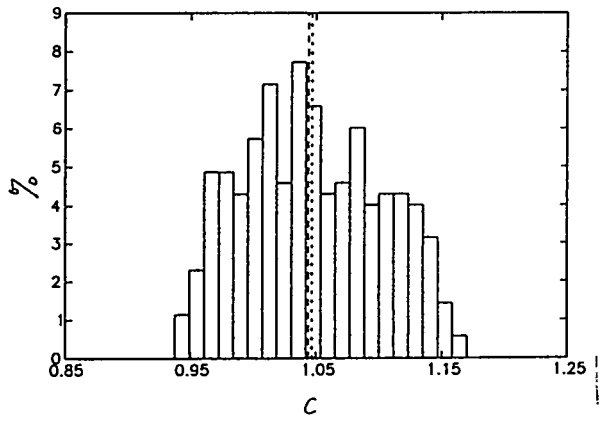
LINEAR PROBLEM AT FINITE VOLUME FRACTIONS

We now consider the case of finite volume fractions for the linear problem for homogeneous suspensions of rigid spheres in potential flow. On the basis of standard Continuum Mechanics argument (Galilean invariance, isotropy, etc.), it is possible to show that $\mathbf{A}_D$ must have the form [1]

$$\mathbf{A}_D = \frac{1}{2}\rho_C \beta_C C(\beta_D, \rho_D/\rho_C) \frac{\partial}{\partial t}(\mathbf{u}_C - \mathbf{u}_D), \tag{46}$$

where C is a scalar coefficient dependent on the ratio of the densities and the volume fraction. The corresponding form of $\mathbf{A}_C$ is given by (29) with $\nabla \cdot \mathbf{T}_C = 0$ due to homogeneity. The well-known Zuber approximation corresponds to $C = 1$ and Wallis's "exertia" E [8] is given by $E = \frac{1}{2}\beta_D C$. In [1] we have calculated C by carrying out ensemble-averaging numerically. Some results are shown in the figure on the left. The figure on the right shows instead an example of the probability distribution of the values of C obtained from different simulations for $\beta_D = 0.4$ and $\rho_D = 0$. It is interesting to note that the spread of this distribution is such as to obscure any dependence of C on its variables.





The present work has been supported by DOE under grant DE-FG02-89ER14043.

REFERENCES

1. D.Z. ZHANG and A. PROSPERETTI, "Pressure forces in disperse two-phase flows," submitted to *J. Fluid Mech.*.

2. E.J. HINCH, "An averaged-equation approach to particle interactions in a fluid suspension," *J. Fluid Mech.* **83**, 695 (1977).

3. A. PROSPERETTI and A.V. JONES, "Pressure forces in disperse two-phase flows," *Int. J. Multiphase Flow* **10**, 425 (1984).

4. T.R. AUTON, J.C.R. HUNT, and M. PRUD'HOMME, "The force exerted on a body in inviscid unsteady non-uniform rotational flow," *J. Fluid Mech.* **197**, 241 (1988).

5. A. BIESHEUVEL and L. VAN WIJNGAARDEN, "Two-phase flow equations for a dilute dispersion of gas bubbles in liquid," *J. Fluid Mech.* **148**, 301 (1984).

6. D.A. DREW, "Effect of particle velocity fluctuations in particle fluid flows," *Physica* **179A**, 69 (1991).

7. A. BIESHEUVEL and S. SPOELSTRA, "The added mass coefficient of a dispersion of spherical gas bubbles in liquid," *Int. J. Multiphase Flow* **15**, 911 (1989).

8. G.B. WALLIS, "The averaged Bernoulli equation and macroscopic equations of motion for the potential flow of a two-phase dispersion," *Int. J. Multiphase Flow* **17**, 683 (1991).

9. A.S. SANGANI and A.K. DIDWANIA, "Dispersed phase stress tensor in flows of bubbly liquids at large Reynodls numbers," *J. Fluid MEch.* **248**, 27 (1993).

10. G.S. ARNOLD, D.A. DREW, and R.T. LAHEY Jr., "Derivation of constitutive equations for interfacial force and Reynolds stress for a suspension of spheres using ensemble cell averaging," *Chem. Eng. Comm.* **86**, 43 (1989).

PARTIAL CONTROL OF COMLEX CHEMICAL PROCESSES I. CONTROL OF FLUIDIZED CATALYTIC CRACKER

R. Shinnar and I. Rinard
Dept. of Chemical Engineering, City College of New York
138th St. at Convent Av., New York, NY. 10031, U.S.A.

SUMMARY

A detailed dynamic model of a fluidized catalytic cracker has been developed that allows evaluation of the impact of different designs, control configurations, catalyst and feed composition and control strategies on the control of a fluid cracker. The present paper deals with the existence and the topology of multiple steady states. It is shown that in some cases five steady states can exist. Further, some of these can be close together in terms of the input operating space. Present trends in operating conditions (higher regenerator temperatures and higher catalyst activities) increase the likelihood that desirable operating conditions are in the regions where such multiplicities occur. It is shown that catalytic combustion promoters can eliminate or reduce this problem. The paper also shows that conventional control structures can lead to input multiplicities and that the choice of additional control variables in the primary matrix should depend on operating conditions.

INTRODUCTION

The goal of our research program is to provide a more rigorous framework for designing control systems for complex chemical processes. First of all such processes are nonlinear. Second, the number of variables that need to be controlled is often much larger than the number of manipulated variables at our disposal. This gives rise to the concept of partial control.

In the conventional sense, controllability means keeping all controlled variables at specified values. For partial control, controllability means that we can keep the vector of controlled variables in a given space. One can use the information obtained from all measurements to adjust simultaneously all manipulated variables via a detailed process model. Alternatively one can use a square matrix by choosing measurements indicating the state of the process. One can then use the information obtained from measurement of the complete output vector including process constraints as well as product specifications to adjust the setpoints of the primary square control matrix. The latter, if properly done, requires less model sensitivity and requires less detailed process information. It is often used in industry and is here our primary concern.

As the first example of our study we chose a fluidized catalytic cracker (see Fig. 1 and Ref. 1). In such a unit fluidized catalyst circulates between a reactor and a regenerator. In the reactor the hot catalyst heats up the high boiling fraction of crude oil and cracks it to gasoline and gaseous products. Coke is

formed on the catalyst during cracking. This catalyst goes to the regenerator in which the coke is combusted with air. This supplies the heat required for the process and the process is adiabatically heat balanced.

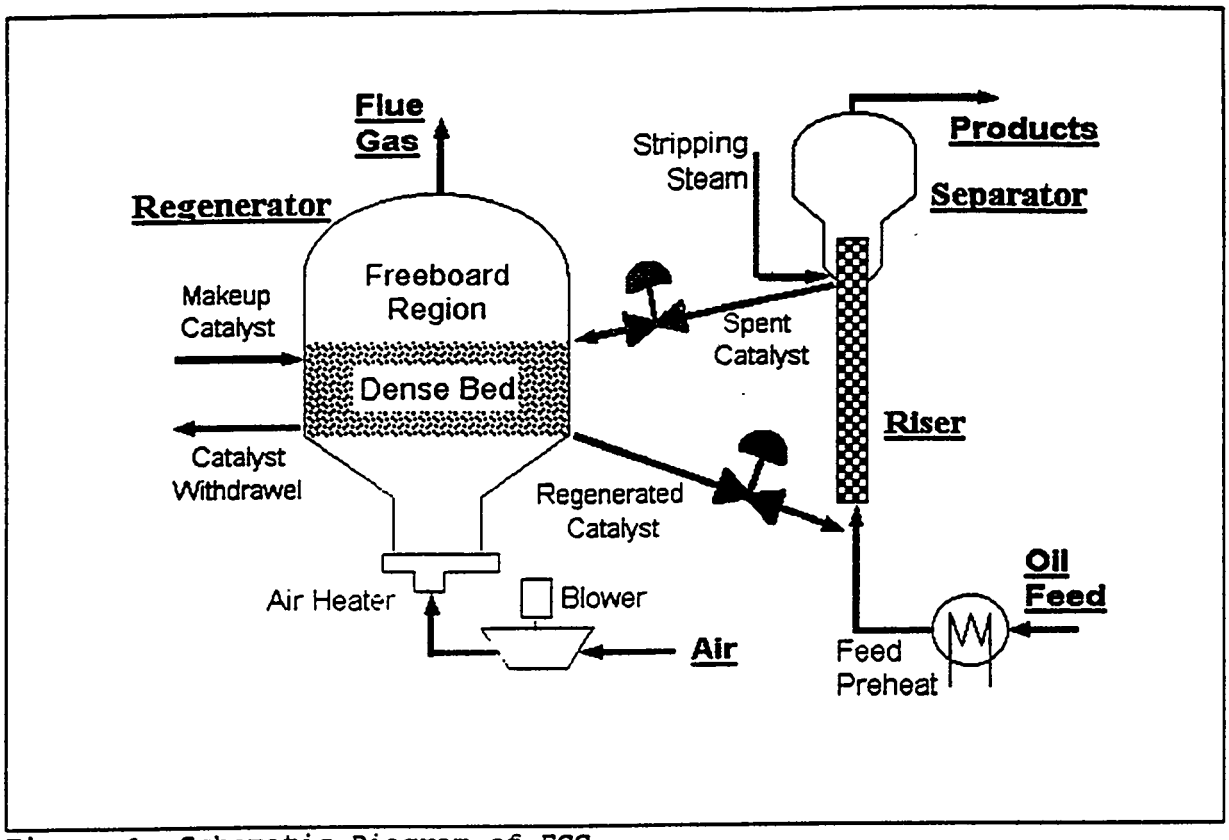


Figure 1: Schematic Diagram of FCC.

The feed can be preheated in a separate furnace and in some units heat can be removed from the regenerator via a catalyst cooler In some older units catalyst circulation is fixed and can only be slowly adjusted by changing catalyst inventory In new designs it can be adjusted. These design changes have a strong impact on controllability

The flue gas from the regenerator normally contains a large fraction of carbon monoxide; many units have a CO boiler to combust the remaining CO in the flue gas. In the last twenty years, catalytic combustion promoters have been introduced which allow complete combustion to CO_2 at temperatures below 1300°F. This increases the heat generated from the coke which, in the absence of a catalyst cooler, either lowers conversion or limits the type of the feed that the unit can handle.

Catalysts and feedstocks have changed strongly over the years which necessitates changes in operating conditions. Feedstock quality and desired operating conditions can change drastically in a given unit during any given year. In recent years, tighter product specifications have also increased the demand for better control as the FCC is the kingpin of a modern refinery. It is the unit in the refinery which is the most flexible and which can operate over the widest range of feed and operating conditions.

In the initial phase of our study we have developed a detailed model of the FCC, and have looked at the basic features of the system such as multiple steady states and the input multiplicities of those control configurations currently used in the industry

FCC MODEL DEVELOPMENT

There are a number of detailed models for the FCC that would be satisfactory for our purposes, but all are proprietary. None of the published models [2, 3, 4] have enough detail, especially on the reactor side, to serve our purposes. In our method we use a detailed nonlinear model to serve as a substitute for the system studied, and our design methods assume that the model is not completely accessible.

All the published models do not allow computation of the complete product specification vector. Some of them do not allow the existence of the observed multiple steady states [8,9]. The model developed up to this point will serve as a basis for future studies, but after further refinements in predicting product properties are added to it.

A detailed description of the model is impossible within the short space allocated and will be published separately. As the senior author has extensive experience with industrial FCC operation, the model was checked as to its capability to predict the observed trends in real operation.

NONLINEARITIES OF CONTROL SYSTEMS

Most chemical reactors are nonlinear systems. Despite that, linearized transfer functions generally serve quite well for the tuning of control loops. There are however, several features of nonlinear systems which are important to understand:

1) Steady state control almost always requires nonlinear models [5].
2) The linearized transfer functions can depend quite strongly on the operating conditions. Steady state gains and other parameters can even change sign as the result of small changes in the operating conditions.
3) The system can have multiple steady states or exhibit stable oscillations [6, 7].
4) A given control matrix may exhibit input multiplicities.

In this paper, we focus on Items 3 and 4.

MULTIPLE STEADY STATES

The existence of multiple steady states as determined by both computation and observation has been discussed in the literature [9, 10]. Ref. 9 however does not give the details of the model used and there has been no thorough discussion of the topology of these multiple states or of the impact of design, catalyst properties and operating conditions on them.

Any adiabatic system such as an FCC has the potential for five steady states. If the feed is not hot enough to cause ignition, there is only one steady state in which the unit remains cold. If the catalyst with a specific feed does not generate enough coke to maintain operation at a high temperature the unit will crash to this state (see Fig 2) Otherwise there are three steady states.

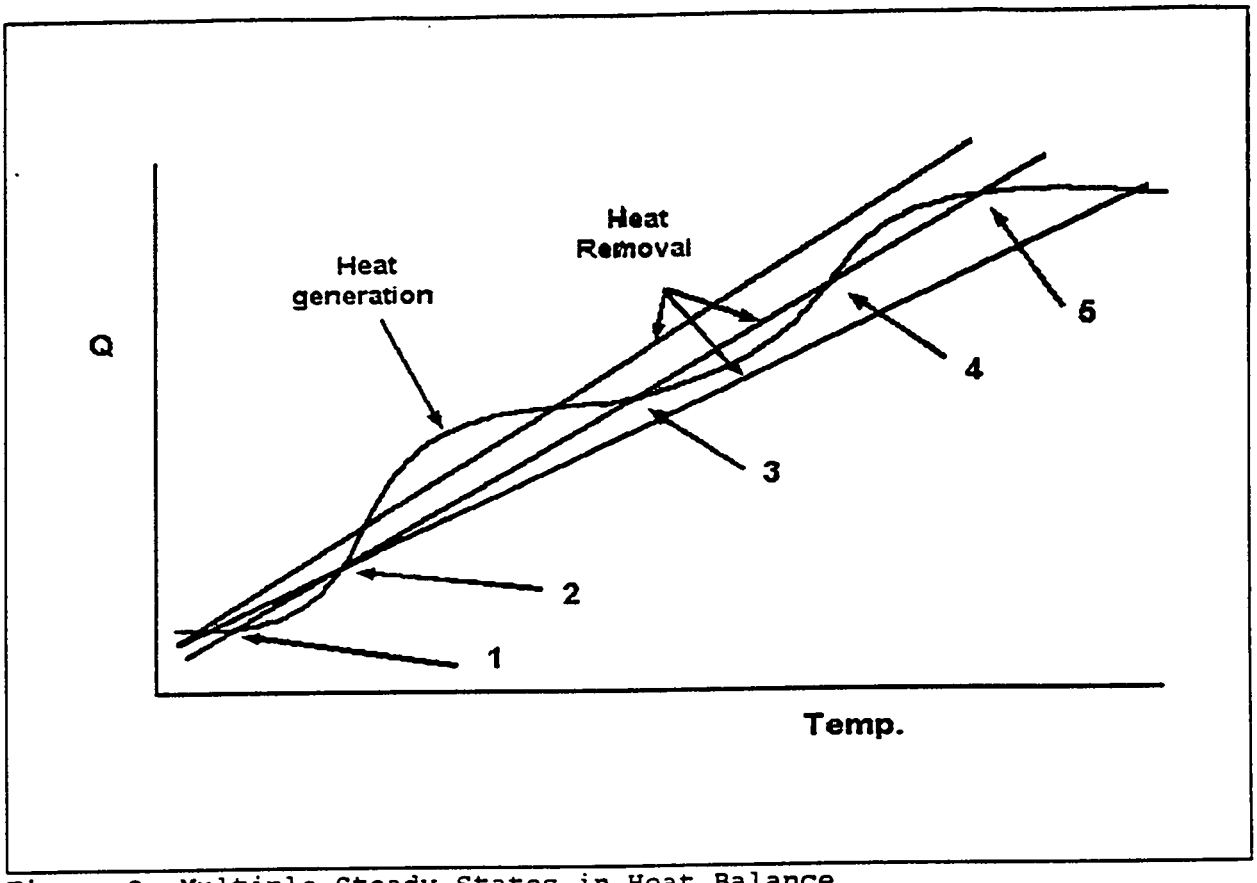


Figure 2: Multiple Steady States in Heat Balance.

The potential for the two additional steady states is created by the consecutive reaction of CO combusting to CO_2 which, in the absence of a promoter, is a homogeneous reaction. At temperatures below 1200°F this reaction is stopped by the presence of catalyst. Therefore, if the flue gas contains oxygen, it will ignite after leaving the cyclone. This will show up as a temperature difference across the cyclone. At higher temperatures (above 1250°F) the inhibition decreases and the reactor can take off even in the presence of a catalyst.

We have approximated this complex relation by using a high activation energy (70 kg cal/g mole) in the presence of a catalyst. This can lead to five steady states. Previously most FCC's operated at temperatures below 1200°F. Modern, more active catalysts require higher regenerator temperatures due to lower circulation rates and the desire for higher octane and more alkylate feed. This increase in temperature makes the existence of five steady states more likely

Fig. 3 gives a plot of the different regimes as a function of open loop inputs of manipulated variables (catalyst circulation rate and air flow rate) In Fig. 4 we show that the impact of feed properties (intrinsic coking rate) on the existence regions of both three and five steady states. Fig. 5 gives the impact of catalyst activity and Fig. 6 gives the impact of a catalyst promoter Decreasing the coking rate or catalyst activity limits the range of stable operation (3 steady states) whereas increasing either one increases the likelihood of five steady states.

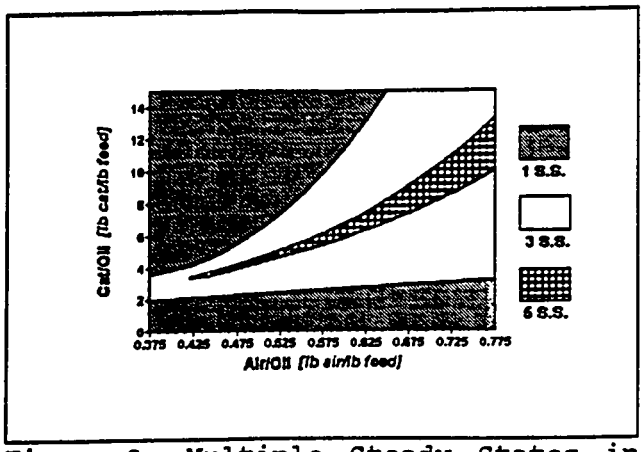


Figure 3: Multiple Steady States in FCC's.

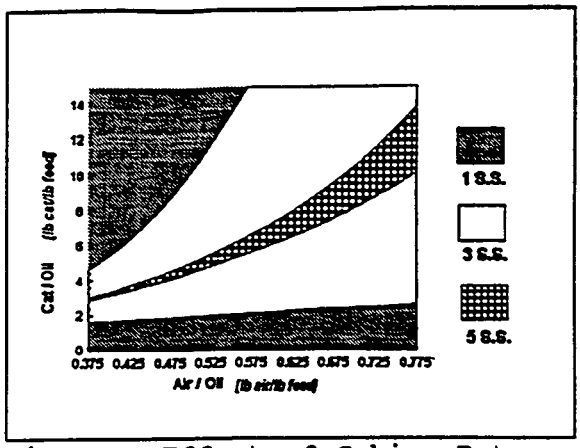


Figure 4: Effect of Coking Rate on Multiple Steady States in FCC's. Coking Rate is x1.25 of base case.

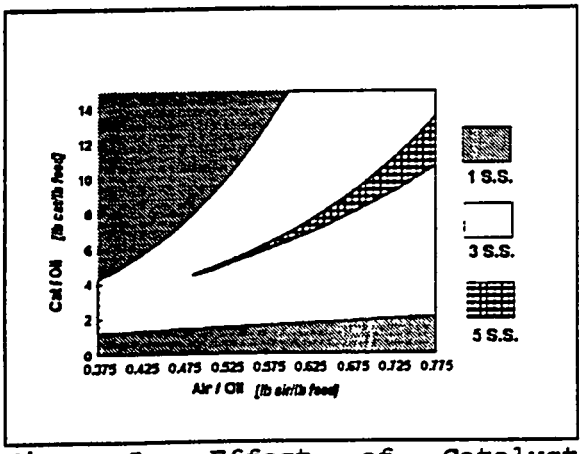


Figure 5: Effect of Catalyst Activity on Multiple Steady States in FCC. Catalyst Activity is x2 of basic case.

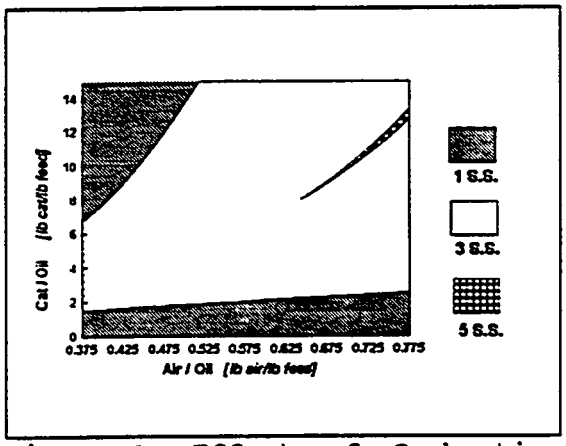


Figure 6: Effect of Combustion Promoter on Multiple Steady States in FCC's. Combustion rate is x18.5 of base case.

In Fig. 6, we show the impact of a catalytic combustion promoter. Promoting the catalytic combustion of CO_2 reduces the area where five steady states exist. It also moves it to higher operating temperatures. This can be seen from Fig. 7 where we superimpose on Fig. 6 the temperature isoclines of the reactor and the regenerator outlining the normal operating range for both cases. We note that the catalytic promoter moves the operating range into the region where only three steady states exist.

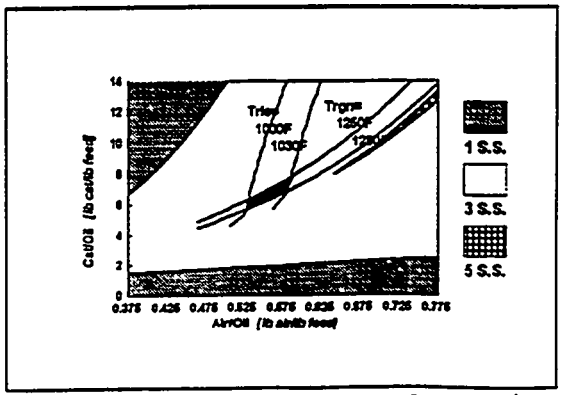


Figure 7: Temperatures Operating Range in FCC with Combustion Promoter.

There is a special characteristic of these multiple steady states in our system. In the desirable operating range, there is a chance that these multiple steady states are close to each other. If they are far apart one still has to know about their existence but their impact on controllability is usually larger when they are close together.

INPUT MULTIPLICILTIES

In Fig. 8, we show some typical control configurations for an FCC. For brevity, we concentrate on the case without a catalyst cooler. Feed temperature is manipulated only for steady state control. Dynamic control focuses on air rate and catalyst circulation rate. Catalyst circulation rate is controlled to adjust the reactor riser top temperature. (All the cracking takes place in the riser which functions as an adiabatic reactor).

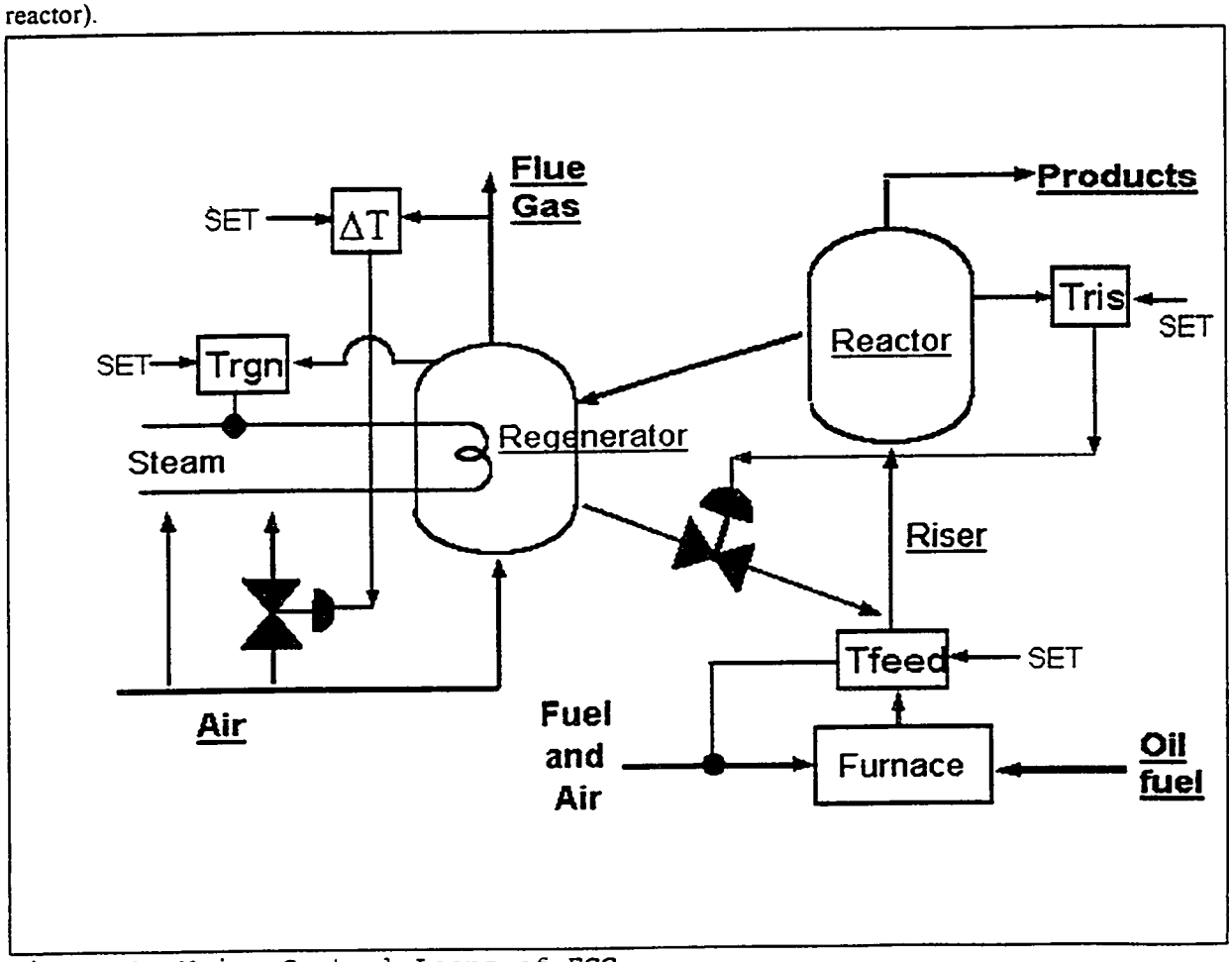


Figure 8: Major Control Loops of FCC.

Air rate is used to adjust one of the following: (1) regenerator bed temperature, (2) the temperature differecne across the cyclones, or (3) the flue gas temperature leaving the cyclones. Regenerator bed temperature is the one of direct importance as it is the only one that can intuitively be used to adjust operating conditions to achieve the right conversion However, temperature difference or flue gas temperature control have a faster response. This is useful to protect the cyclones from metalurgical damage due to overheating. At high temperatures the cyclone temperatuare difference is small and therefore not

suitable for control. One can use either regenerator or flue gas temperatures instead.

In the case where enough promoter is added to achieve complete CO combustion, oxygen in the flue gas is used to control air rate. In Fig. 9, the air rate as a function of the controller settings for these four different cases is shown for a fixed reactor temperature of 980 F. This is shown in the base case in Fig. 3. We note that there are input multiplicities for each of the cases. However, they occur in different operating regimes. This must be taken into account in the design of any control strategy.

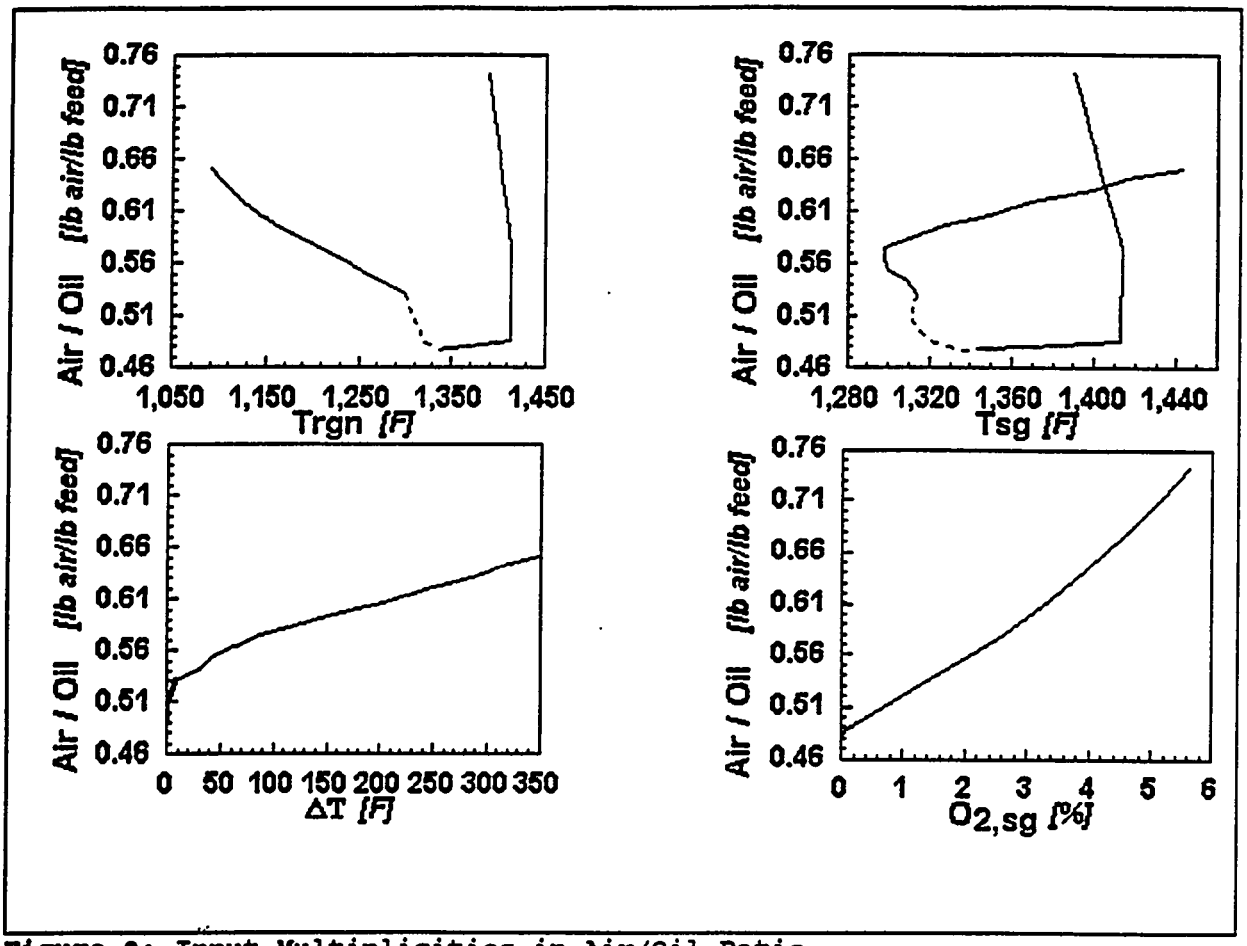


Figure 9: Input Multiplicities in Air/Oil Ratio. Triser is 980F.

Our results also show that for some operating conditions the eigen values of the linearized control matrix can vary strongly over reasonable changes in set points.

Previous control studies [12,13] have emphasized the speed of the response. Our results clearly indicate that in some operating ranges this is less important than the need to avoid large changes in the eigen values for small changes in set points and also the need to avoid input multiplicities.

In a real system, catalyst properties and feed properties as well as operating conditions vary over a wide range. For good control, it is not necessary to be able to reliably model the exact occurrence of multiple steady states and input multiplicities. But a thorough understanding of their nature and potential topology is essential and will be incorporated in our design approach.

REFERENCES

1. Avidan, A. and Shinnar, R., "Development of Catalytic Cracking Technology. A Lesson in Chemical Reactor Design," *Chem. Eng. Res.*, 29 , 931 (1990).

2. Shinnar, R., "Chemical Reactor Modelling for Purposes of Controller Design," *Chem. Eng. Commun.*, 9, 73, (1981).

3. Weekman Jr., V.M. , "Lumps , Models and Kinetics in Practice," *AIChE monograph Series 11*, 75, (1979).

4. McFarlane, R.C., Reinman, R.C., Bartee, J.F. and Georgakis, C.,"Dynamic Simulation for Model IV Fluid Catalytic Cracking Unit," *Computers Chem. Eng.*, 17, 275, (1993).

5. Errazu, A.F., de Lasa, H.I. and Sarı, F., "A Fluidized Bed Catalytic Cracking Regenerator Model. Grid Effect.", *The Canadian J. of Chem. Eng.*, 57, (1979).

6. Shah, Y.T., Huling, G.P , Paraskos, J.A. and McKinney, J.D., "A Kinematic Model for an Adiabatic Tranfers Line Catalytic Cracking Reactor," *Ind. Eng. Chem. Process Des. Dev.*, 16, (1977).

7 Sapre, A.V., Leib, T.M. and Anderson, D.H., "FCC Regenerator Flow Model," *Chem. Eng. Sci.*, 45, 2203, (1990).

8. Edwards, W.M. and Kim, H.N., "Multiple Steady STates in FCCU Operations," *Chem. Eng. Sci.*, 43, 1825, (1988).

9. Arandes, J.M. and de Lasa, H.I., "Simulation and Muliplicity od Steady States in Fluidized FCCUs," *Chem. Eng. Sci.*, 47, 2535 (1992)

10. Luss, D., "Steady State multiplicity Features of Chemcially Reacting Systems," *Chem. Eng. Educ.*, 20, 12 (1986).

11. Lee, W. and Weekman Jr., V M., "Advanced Control Practice in Chemical Process Industry: View from Industry," *AIChE J.*, 22, 27 (1976).

12. Balchen, J.G., Ljungquist. D. and Strand. S., "State Space Predictive Control," Chem. Eng. Sci.,47, 787, (1992).

13 Hovd, M. and Skogestad. S.,"Controlability Analysis for the FCC Process," AIChE Annual Meeting, paper 145g, (1991).

MODELING ASPECTS OF WAVE KINEMATICS IN OFFSHORE STRUCTURES DYNAMICS

P.D. Spanos
L.B. Ryon Professor, Rice University, Houston, TX.

R. Ghanem
Civil Engineering Department, State University of New York at Buffalo, NY.

S. Bhattacharjee
Mobil Research and Development Corp., Dallas, TX.

ABSTRACT

Magnitude and phase related issues of modeling of ocean wave kinematics are addressed. Causal and non-causal filters are examined. It is shown that if for a particular ocean engineering problem only the magnitude representation of wave spectra spatial relation is critical, analog filters can be quite useful models in conjunction with the technique of statistical linearization, for calculating dynamic analyses. This is illustrated by considering the dynamic response of a simple model of a guyed tower.

INTRODUCTION

A rational analysis of offshore structures exposed to wave-induced loads requires the estimation of the fluid kinematics at different spatial locations along the structure. In fact available kinematics information at one point on the water surface, must be propagated on a spatial grid that is commensurate with the structural analysis procedure used. Both vertical and horizontal kinematics propagation are important in this context. However, each of these propagations features a peculiar behavior not present in the other. For example, the vertical propagation is not accompanied by any phase shift, while the horizontal propagation features no attenuation. It is obvious, therefore, that no single filter can be used in addressing both the horizontal and the vertical propagation issues. Examples of filter approaches to wave kinematics representation can be found in other references [1-5]. For the vertical propagation case, a filter should be designed possessing a phase that is independent of depth. This criterion will relinquish the traditional control over the magnitude of the transfer function, in favor of control over the phase. It will be shown that this criterion can only be satisfied if a non-causal filter is used, thus resulting in a symmetric Moving Average (MA) design. In contrast, the horizontal propagation problem can be best accomplished using a filter with constant magnitude, but with a phase that approximates the known phase of the wave propagation. This latter can be obtained based on the dispersive character of horizontal kinematics propagation. The propagation of horizontal wave kinematics is accomplished using an all pass filter with a nonlinear optimization criterion to match the phase response function.

Once filters have been designed for representing the wave kinematics, they can be used efficiently not only for synthesizing time histories, but for conducting, as well, dynamic analyses of offshore structures. Attention to the analytical advantages of filter approaches to ocean wave kinematics representation has been called in references such as [5].

In this paper, the analytical advantages are demonstrated, in conjunction with the technique of statistical linearization [6] for predicting the random response of a guyed tower.

FILTER DESIGN

The analysis of the propagation of wave kinematics at intermediate water depth relies on the Airy, or linearized wave theory. First, an analysis of the vertical propagation of wave kinematics, from a perspective of amplitude and phase, will be considered.

Vertical Propagation of Wave Kinematics

The spectrum of the horizontal velocity of particles at any location z measured from the bottom is given by the

equation [5]

$$S_u(\omega) = \left[\omega \frac{\cosh(\kappa z)}{\sinh(\kappa d)}\right]^2 S_{\eta\eta}(\omega) \ , \tag{1}$$

where $S_{\eta\eta}$ denotes the spectrum of the surface particle elevation. The above equation can be normalized using the following substitutions,

$$b = \kappa d \qquad \bar{\omega}^2 = \omega^2 \frac{d}{g} = b \tanh b \qquad \alpha = \frac{z}{d} \ , \tag{2}$$

where d denotes the local water depth, and g denotes the acceleration of gravity. The resulting normalized equation has the form

$$\bar{S}_u(\bar{\omega}) \equiv \frac{d}{g} \frac{S_{uu}(\omega)}{S_{\eta\eta}(\omega)} = b \tanh b \left[\frac{\cosh(\alpha b)}{\sinh b}\right]^2 \ , \tag{3}$$

Adopting the procedure described by [3,6], equation (3) can be approximated using the rational function

$$\bar{S}_u \approx \hat{S}_u(\bar{\omega}) = \frac{\omega_1^4 \omega_2^4}{[(\bar{\omega}^2 - \omega_1^2)^2 + (2\zeta_1\omega_1\bar{\omega})^2]\,[(\bar{\omega}^2 - \omega_2^2)^2 + (2\zeta_2\omega_2\bar{\omega})^2]} \ , \tag{4}$$

where $\omega_1, \omega_2, \xi_1, \xi_2$ are coefficients dependent on α. The right-hand side of the above equation can then be factorized as

$$\hat{S}_u(\bar{\omega}) = \hat{H}_{uz}(i\bar{\omega})\hat{H}_{uz}(-i\bar{\omega}) \ , \tag{5}$$

where $\hat{H}_{uz}$ denotes the frequency response function of the kinematics propagation filter at depth z and is given by the equation,

$$\hat{H}_{uz}(\bar{\omega}) = \frac{\omega_1^2 \omega_2^2}{[\omega_1^2 - \bar{\omega}^2 + i(2\zeta_1\omega_1\bar{\omega})]\,[\omega_2^2 - \bar{\omega}^2 + i(2\zeta_2\omega_2\bar{\omega})]} \ . \tag{6}$$

The impulse response function corresponding to the function $\hat{H}_{uz}(s)$ is given by

$$\begin{aligned} h_{uz}(t) &= \mathrm{sgn}(R_1) \frac{2\omega_1^2\omega_2^2}{\sqrt{R_1^2 + I_1^2}} e^{-\omega_1\zeta_1 t} \cos(\omega_1\sqrt{1-\zeta_1^2}\,t + \phi_1) \\ &+ \mathrm{sgn}(R_2) \frac{2\omega_1^2\omega_2^2}{\sqrt{R_2^2 + I_2^2}} e^{-\omega_2\zeta_2 t} \cos(\omega_2\sqrt{1-\zeta_2^2}\,t + \phi_2) \ , \end{aligned} \tag{7}$$

where

$$R_1 = 4\omega_1^2 [(\omega_1\zeta_1 - \omega_2\zeta_2)(1 - \zeta_1)] \tag{8}$$

$$I_1 = (2\omega_1\sqrt{1-\zeta_1^2})\,[(\omega_1^2 - \omega_2^2) - 2\zeta_1^2\omega_1^2 + 2\zeta_1\omega_1\zeta_2\omega_2] \tag{9}$$

$$\tan\phi_1 = \frac{I_1}{R_1}; \qquad -\frac{\pi}{2} \le \phi_1 \le \frac{\pi}{2} \tag{10}$$

$$R_2 = 4\omega_2^2 [(\omega_2\zeta_2 - \omega_1\zeta_1)(1 - \zeta_2)] \tag{11}$$

$$I_2 = (2\omega_2\sqrt{1-\zeta_2^2})\,[(\omega_2^2 - \omega_1^2) - 2\zeta_2^2\omega_2^2 + 2\zeta_1\omega_1\zeta_2\omega_2] \tag{12}$$

$$\tan\phi_2 = \frac{I_2}{R_2}; \qquad -\frac{\pi}{2} \le \phi_2 \le \frac{\pi}{2} \ . \tag{13}$$

The vertical propagation of the vertical velocity component can be treated in a similar fashion, starting with the following nondimensional form of the spectrum,

$$\bar{S}_v(\bar{\omega}) \equiv \frac{S_v(\omega)}{S_\eta(\omega)} \left[\frac{d}{\alpha^2 g\bar{\omega}^2}\right] = \left[\frac{\sinh(\alpha b)}{\alpha \sinh b}\right]^2 \ , \tag{14}$$

which is replaced by the following rational approximation,

$$\bar{S}_v(\bar{\omega}) \approx \hat{S}_v(\bar{\omega}) = \frac{\nu^4}{[(\bar{\omega}^2 - \nu^2)^2 + (2\beta\nu\bar{\omega})^2][1 + (c\bar{\omega})^2]} \cdot \tag{15}$$

This spectrum can again be factorized using the transfer function $\hat{H}_{vz}$ which is given by the equation,

$$\hat{H}_{vz}(\bar{\omega}) = \frac{\nu^2}{[\bar{\omega}^2 - \nu^2 + j2\beta\nu\bar{\omega}][1 + jc\bar{\omega}]} , \tag{16}$$

where ν, β, and c are functions of α. Again, the coefficients in the above rational approximation are usually evaluated so as to minimize a measure of the error between the target spectrum and its approximation, and are therefore a function of the fractional depth α. The impulse response function corresponding to the transfer function of this spectrum can be shown to be equal to,

$$h_{vz}(t) = \text{sgn}(R)\, \frac{2\nu^2 e^{-\nu\beta t}}{\sqrt{R^2 + I^2}} \cos(\nu_d t + \phi) + \frac{\nu^2 c^2 e - t/c}{1 - 2\beta\nu c + \nu^2 c^2} , \tag{17}$$

where

$$\nu_d = \nu\sqrt{1 - \beta^2}, \quad I = 2(\nu\beta c - 1), \quad \phi = \tan^{-1}(\frac{I}{R}), \quad |\phi| \leq \frac{\pi}{2} . \tag{18}$$

The above formulations using cascaded single-degree-of-freedom models are based on an optimality criterion that minimizes the error between the magnitude of the target and approximating spectra. Filters developed in this fashion have a number of desirable properties, including a rational form of the spectrum which can be realized in the time domain using cascaded differential operators. By restricting the error norm to the magnitude of the approximating function, however, control is lost over the behavior of its phase. The implication of this fact on the simulated process, is that although the magnitude of particle velocities are adequately reproduced, their phase relations are arbitrary. This phenomenon can be crucial in designing structures to accommodate fluid-structure interaction where the cross-correlation of the exciting force along the length of the structure is important. Based on this argument, it is apparent that a constant phase is a desirable feature of a simulated process along the depth. Phase information can be incorporated into the design of the filter by recasting the problem into a single-input-multi-output (SIMO) framework [7] with the surface elevation being the input, and the kinematics at the various locations along the depth being the outputs. In case the surface elevation is not a white noise process, it can always be considered as the output of a digital filter to white noise input, and the system equations can be augmented accordingly. Based on the kinematics propagation equations presented above, it can be shown that the phase transfer function for the vertical propagation of the horizontal and vertical velocity components can be approximated by the following functions,

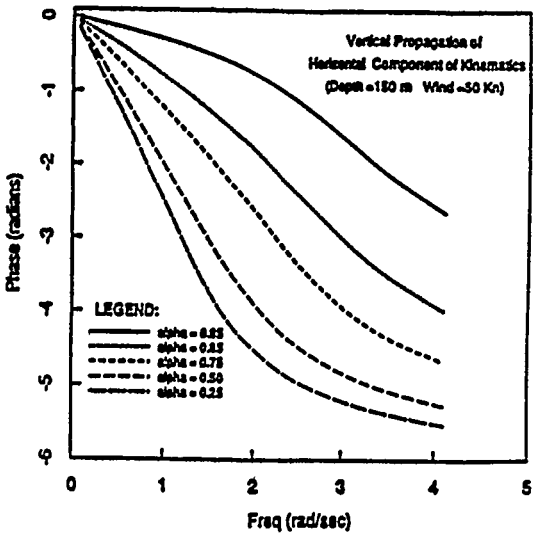


Fig. 1 Phase Response of Oscillator Model for Vertical Propagation of Horizontal Velocity

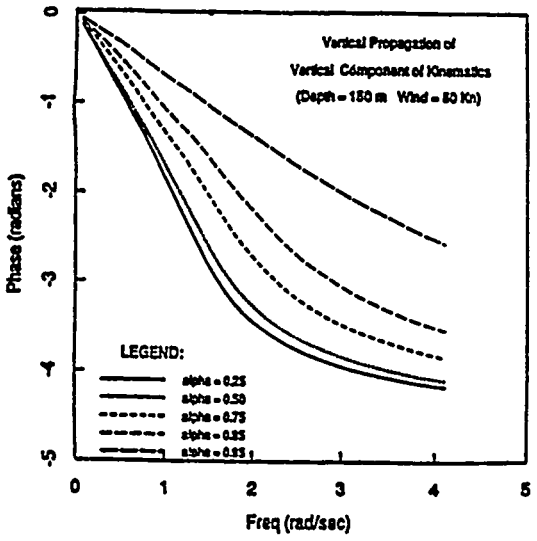


Fig. 2 Phase Response of Oscillator Model for Vertical Propagation of Vertical Velocity

$$\theta_u(\bar{\omega}) = \mathrm{Arg}(\bar{H}_u(\bar{\omega})) = \tan^{-1}\left[\frac{2(\zeta_1\omega_1 + \zeta_2\omega_2)\omega^3 - 2\omega_1\omega_2(\zeta_1\omega_2 + \zeta_2\omega_1)\bar{\omega}}{\bar{\omega}^4 - (\omega_1^2 + \omega_2^2 + 4\zeta_1\zeta_2\omega_1\omega_2)\bar{\omega}^2 + \omega_1^2\omega_2^2}\right] , \tag{19}$$

for the horizontal component, and

$$\theta_v(\bar{\omega}) = \mathrm{Arg}(\bar{H}_v(\bar{\omega})) = \tan^{-1}\left[\frac{c\bar{\omega}^3 - (2\beta\nu + c\nu^2)\bar{\omega}}{-(2\beta\nu c + 1)\bar{\omega}^2 + \nu^2}\right] , \tag{20}$$

for the vertical component. Both of the above expressions for the phase responses of the filter feature nonlinear dependence on frequency and depth. Figures (1) and (2) show these phase spectra for different vertical locations. Note that as much as 2π rd phase difference can occur between values of a given realization at different water depths. This is in contrast with the desired situation of zero-phase shift of the simulated process along the depth coordinate. Such an ideal situation can be accommodated by using the following real transfer functions relating the kinematics at the surface level to those at a specified depth.

$$H_{ux}(\omega;\alpha) = \frac{\cosh(\alpha\kappa d)}{\sinh(\kappa d)} , \tag{21}$$

and

$$H_{vx}(\omega;\alpha) = \frac{\sinh(\alpha\kappa d)}{\sinh(\kappa d)} . \tag{22}$$

The above functions are obtained by considering the square root of the spectrum for the propagation of the associated kinematic, prior to making a rational function approximation. These transfer functions cannot be realized using causal filters since they do not satisfy Kramers-Kronig relationships which require the real and imaginary parts of a transfer function to be Hilbert transforms of one another. Infinite impulse response (IIR) filters cannot therefore be utilized in this case, and only finite impulse response (FIR) filters are adequate for this task. This conclusion is expected since the purpose of the filter is to simulate spatial variability of various kinematic quantities, and causality is a property inherent only in processes with a preferred direction of propagation, this not being the case for the vertical propagation of wave kinematics. As will be demonstrated below, this non-causality does not apply for the horizontal propagation of wave kinematics, which again is expected in view of the preferred propagation in the direction of the wave motion.

Coming back to the vertical propagation of wave kinematics, a two-sided, moving average (MA) filter can be designed, based on an algorithm by Borgman [1]. The algorithm provides a mean for simulating the time evolution of a specified wave kinematic at a specified vertical propagation distance, and a specified water depth and wind velocity. The algorithm can be realized using the following equations,

$$A_n = \frac{1}{\omega_c}\int_0^{\omega_c} \Re\left(H_x(\omega;\Delta x)\right)\cos(\frac{n\pi\omega}{\omega_c})d\omega \tag{23}$$

$$B_n = \frac{1}{\omega_c}\int_0^{\omega_c} \Im\left(H_x(\omega;\Delta x)\right)\sin(\frac{n\pi\omega}{\omega_c})d\omega \tag{24}$$

$$a_0 = A_0, \qquad a_n = A_n + B_n, \qquad a_{-n} = A_n - B_n . \tag{25}$$

In the above equations, the subscript x on H is used to represent either of the horizontal or the vertical component, and $\Re(.)$ and $\Im(.)$ denote the real and imaginary components of a complex quantity, respectively. Once the above coefficients have been calculated, the simulation can be performed using the following equation,

$$y_k = \sum_{n=-N}^{N} a_n x_{k-n}; \quad k = N+1, N+2, \ldots , \tag{26}$$

where n denotes the n^{th} time step. The time step Δ being related to the cut-off frequency ω_c through the Nyquist relation, $\Delta t = \frac{\pi}{\omega_c}$. For the particular case at hand, and since the transfer functions involved are real, the coefficients B_n above are identically zero, and the coefficients of the filter thus obtained are symmetrical about the origin. Figure (3) shows a typical result for the transfer function associated with a filter designed according to the above criteria. Note that the filter features zero phase response over the entire frequency range.

Horizontal Propagation of Wave Kinematics

The above treatment is specific to the vertical propagation of water wave kinematics in intermediate water depth, and it relies on the particular form of the transfer function. The above recommendations are intimately associated

with the fact that along the vertical direction, wave kinematics should be subject to pure attenuation with no phase lag. For the horizontal propagation of wave kinematics, the picture is quite different. In this case, no attenuation is expected, and the propagation is expected to be dispersive. The transfer function for any kinematics of interest in this case, is given by the following equation,

$$H_x(\omega;\Delta x) = e^{-j\kappa\Delta x} = e^{-j\mu b}; \qquad \mu = \frac{\Delta x}{d} \,. \tag{27}$$

In general, the cut-off frequency for the Pierson-Moskowitz wave spectrum is a function of the wind velocity. The transfer function, and therefore, the coefficients of a digital filter will in general be a function of the local water depth and the wind velocity, as well as of the propagation distance Δx.

All-pass filters provide an attractive alternative for the horizontal propagation problem. These filters have a constant transfer function over all frequencies, thus providing no magnitude attenuation, which is commensurate with the horizontal wave propagation phenomenon. These filters have the feature that their poles and zeros are complex conjugate of one another. Stability requirements are satisfied provided $\Re(p_j) < 0$ where p_j are the zeros appearing in the expression for the transfer function,

$$H_{ap}(s) = \frac{(s+\bar{p}_1)(s+\bar{p}_2)\ldots(s+\bar{p}_n)}{(s+p_1)(s+p_2)\ldots(s+p_n)} \,. \tag{28}$$

The phase spectrum corresponding to the all pass filter can be evaluated using the following expression,

$$\tan(\theta(\omega)) = 2\left[\frac{\Im D(\omega)}{\Re D(\omega)}\right] \,, \tag{29}$$

where $D(\omega)$ is the complex denominator in the expression for the transfer function. The filter coefficients are evaluated to approximate the phase of the desired transfer function. This is a significant departure from the more common problem of matching the magnitude of the transfer function. The determination of the coefficients based on this criterion involves a nonlinear optimization problem which can best be solved using the Levenberg-Marquardt algorithm. As an example, the transfer function of a second order all pass filter is given by the equation,

$$H(\omega) = \frac{(\omega_0^2-\omega^2) - j(2\zeta\omega_0)\omega}{(\omega_0^2-\omega^2) + j(2\zeta\omega_0)\omega} \,, \tag{30}$$

where ω_0, ζ, and ω represent the natural frequency, the critical damping, and the independent frequency variable, respectively.

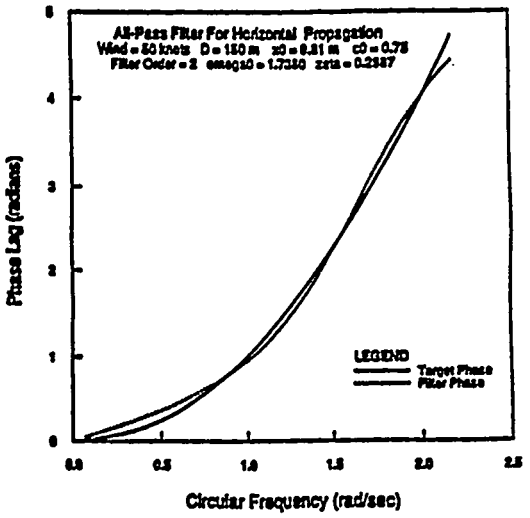


Fig. 4 Phase Approximation of 2nd Order All-Pass Filter for Horz Propagation of Wave Kinematics (Wind = 50 Kn, d=9.81 m)

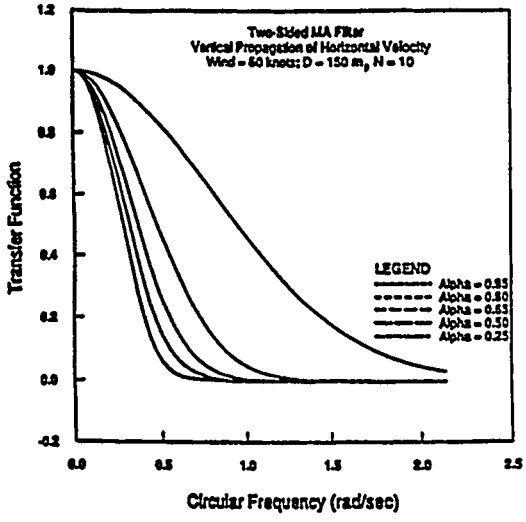


Fig. 3 M-A Filter Approximation of Vertical propagation of Horizontal Velocity Component (MA Filter Order = 10)

The phase of this second order system can be expressed as,

$$\theta(\omega) = -2\tan^{-1}\left[\frac{2\zeta\omega_0\omega}{\omega_0^2-\omega^2}\right] \,. \tag{31}$$

The target phase, based on the transfer function for the horizontal propagation is given by,

$$\theta_t(\omega) = -\kappa(\omega)\Delta x \ , \tag{32}$$

where $\kappa(\omega)$ is a frequency dependent wave number. Thus the optimization problem involves computing the filters coefficients so as to minimize the error given by the following expression,

$$\epsilon = \kappa(\omega)\Delta x - 2\tan^{-1}\left[\frac{2\zeta\omega_0\omega}{\omega_0^2 - \omega^2}\right] \ . \tag{33}$$

A number of properties of this procedure are notable. In particular, the distance over which the filter can successfully propagate the wave kinematics depends on the filter order. An n^{th} order oscillator model can feature a phase shift of up to $n\pi$. Thus the maximum phase shift attainable by the filter imposes limitations on the distance through which the propagation can be carried out. Also, for a given wind velocity, the maximum propagation distance is restricted since the product $\kappa\Delta x$ must lie within the maximum phase delay of the filter, and the value of κ is dependent on the wind velocity. Figure (4) shows the matching between the target and the approximate phase response for a typical set of parameter values.

STOCHASTIC DYNAMIC ANALYSIS OF A GUYED TOWER

It has been shown in the preceding section, that both magnitude and phase characteristics of wave kinematics spectra can be captured by filters. If only magnitude characteristics are important, the analog filters, such as the ones associated with equations (4) and (15) can be useful for expeditious studies of offshore structures. This point will be addressed in this section by considering the dynamic response of a guyed tower to random waves.

The governing equation of motion of the guyed tower considers its rigid body motions about the base hinge. The motion in a plane can be obtained by taking equilibrium about the base. Thus, the governing equation is [8],

$$J_0\ddot{\theta} + cd^2\dot{\theta} + R_x z_{ca} + (F_b z_b - W z_{cg} - R_z z_{ca})\,\theta = F(t)h. \tag{34}$$

In the above equation, J_0 denotes the mass moment of inertia of the platform about the base, including the added mass effects, d is half the depth of the idealized tower, W denotes the combined weight of the platform deck and of the tower. Also, F_b denotes the buoyancy of the platform, z_{cg} and z_b denote the distance from the base of the centers of gravity and buoyancy of the tower alone without the deck, and z_{ca} denotes the distance from the base of the attachement point of the cables. All of the above distances are measured parallel to the tower centerline. The symbol $F(t)$ denotes the total wave force acting on the tower and h is the center of action of this force. Furthermore, R_x and R_z denote the horizontal and vertical restoring forces, and are given by

$$R_x = c_1\theta + c_2\theta^3 \tag{35}$$

$$R_z = \hat{c}_1 + \hat{c}_2\theta^2 \ , \tag{36}$$

where θ denotes the rigid body rotation angle of the tower. Substituting the above expressions results in the following equation of motion,

$$J_0\ddot{\theta} + cd^2\dot{\theta} + (z_{ca}c_1 - W z_{cg} - z_{ca}\hat{c}_1 + F_b z_b)\,\theta + z_{ca}\,(c_2 - \hat{c}_2)\,\theta^3 = F(t)h \ . \tag{37}$$

The above equation of motion is simplified using the technique of equivalent linearization [5,6], and is rewritten in a non-dimensional form as

$$\ddot{\theta} + 2\beta_0\omega_0\dot{\theta} + \omega_{0e}^2\theta = M(t) \ , \tag{38}$$

where ω_{0e} denotes the natural frequency of the equivalent linear system. The wave elevation is considered to be a random process, stationary in time and homogeneous in space. In addition, it is considered to be a zero-mean, Gaussian process, specified by the Pierson-Moskowitz spectrum. The wave field kinematics are assumed to follow the linear wave theory. The force on the structure is estimated by the use of the Morison equation. Using all of the above assumptions and the equivalent linearization technique for the non-linear drag term, an approximate model for the spectrum of the wave-induced moment $M(t)$ about the base can be obtained as [8],

$$\hat{S}_{MM} = \frac{G_0\omega^2}{(\omega^2 - K_0)^2 + (C_0\omega)^2} \ , \tag{39}$$

where the coefficients G_0, K_0, and C_0 are obtained through least-squares approximation to the Pierson-Moskowitz spectrum as discussed by [5]. It can be shown that a process whose spectral density function is a rational function

can be realized as the output of a linear system of appropriate order excited by a generalized derivative of white noise. In this case, the process $M(t)$ is the solution to the following equation, where $W(t)$ denotes a white noise process,

$$\ddot{M} + C_0\dot{M} + K_0 M = \sqrt{G_0}\dot{W}(t) \ . \tag{40}$$

Differentiating the equation of motion twice with respect to time, the equation (40) can be incorporated into equation (38) yielding the following fourth order equation,

$$\theta^{IV} + a_4\theta^{III} + a_3\theta^{II} + a_2\theta^{I} + a_1\theta = \sqrt{G_0}\dot{W}(t) \ , \tag{41}$$

where the various coefficients are given by the following equations,

$$\begin{aligned} va_1 &= K_0\omega_{0e}^2 \\ a_2 &= 2\beta_0\omega_0 K_0 + C_0\omega_{0e}^2 \\ a_3 &= \omega_{0e}^2 + K_0 + 2\beta_0\omega_{0e}C_0 \\ a_4 &= 2\beta_0\omega_0 + C_0 \ . \end{aligned} \tag{42}$$

The above formalism involves the derivative of a white noise process, a quantity that is defined in a generalized sense. From an engineering perspective, the parameter θ is merely the derivative of the response of a fourth order linear system with characteristic equation in the Laplace domain given by

$$\Lambda(s) = s^4 + a_3 s^3 + a_2 s^2 + a_1 s + a_0 \tag{43}$$

and excited by a white noise process with constant power spectrum equal to G_0. Thus, the variance σ_θ^2 can be determined by the equation

$$\sigma_\theta^2 = \int_{-\infty}^{\infty} \frac{\omega^2}{|\Lambda(i\omega)|^2} G_0 d\omega \ , \tag{44}$$

where $i = \sqrt{-1}$, and $|.|$ denotes the modulus of a complex function. Clearly, the power spectrum of θ is

$$S_\theta(\omega) = \frac{G_0\omega^2}{[(\omega_{0e}^2 - \omega^2)^2 + (2\beta_0\omega_0\omega)^2]\,[(K_0 - \omega^2)^2 + (C_0\omega)^2]} \ , \tag{45}$$

from which its variance can be calculated to be

$$<\sigma_\theta^2> = \frac{\pi G_0 a_4}{a_2 a_3 a_4 - a_2^2 - a_1 a_4^2} \tag{46}$$

The analytical form of the integral in equation (44) belongs to the class of integrals which can be calculated analytically as described in [9]; particular attention for its application in offshore engineering has been called in [5].

NUMERICAL EXAMPLE

The approach described in the previous section is applied to an example featuring an idealized guyed tower. The details of the specific values of the various parameters used can be found in [7]. The details of the regression analysis to obtain the polynomial representation for the restoring forces can also be found in that reference. The physical set-up is shown in Figure (5). Table (1) shows the standard deviation at the deck-level of the guy tower for a number of wind speed and associated wave height.

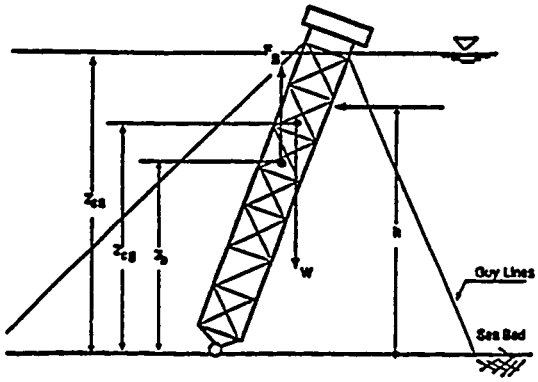


FIG. 5 Model of Idealized Guyed Tower

Wind (ft/s)	H_s (ft)	K_0	C_0^2	G_0 (10^{11})	σ_{X_D} (ft)
30	5.81	0.84	0.11	0.44	1.45
40	10.38	0.52	0.88	1.68	4.45
70	31.84	0.20	0.062	1.41	29.73
80	41.58	0.16	0.041	1.55	39.02

Table 1: Filter Parameters and Response Statistics

CONCLUSIONS

The problem of wave kinematics propagation was reviewed and new related concepts were introduced. Specifically, it was shown that causal filters cannot be effectively used to model the vertical propagation of water particle kinematics. Moving average filters are found to be quite efficient for this situation. Filters were developed for the propagation of both horizontal and vertical kinematics along a vertical line into the water. The horizontal propagation problem is complicated by the dispersive nature of gravity waves. This dependence on frequency of the velocity of the waves was adequately modeled using an all pass filter. This filter featured non-attenuation of magnitude while providing a nonlinear phase to represent the nondispersive nature. The filter was designed so that its phase matches the theoretical phase of the kinematics propagation. The ease with which the various filters introduced can be implemented was demonstrated in conjunction with the dynamic analysis of a simple model of a guyed tower.

ACKNOWLEDGMENT

The support of this work by the grant DE-FG05-91-ER14215 from the Department of Energy is gratefully acknowledged.

REFERENCES

[1] Borgman, L.E., "Ocean wave simulation for engineering design," *Journal of Waterways and Harbor Division, ASCE,* Vol. 95, pp. 557-583, 1969.

[2] Shinozuka, M., and Nai, P., "Digital simulation of short-crested sea surface elevation," *Journal of Energy Resources Technology, ASME,* Vol. 101, pp. 270-275, 1979.

[3] Spanos, P.T.D., and Hansen, J.E., "Linear prediction theory for digital simulation of sea wave," *Journal of Energy Resources Technology,* Vol. 103, pp. 243-249. 1982.

[4] Spanos, P.T.D., "ARMA algorithms for ocean wave modeling," *Journal of Energy Sources Technology, ASME,* Vol. 105, pp. 300-309, 1983.

[5] Spanos, P.T.D., "Filter approaches for wave kinematics approximation," *Applied Ocean Research,* Vol. 8, pp. 2-7, 1986.

[6] Robert J.B. and Spanos, P.T.D., *Random Vibration and Statistical Linearization,* John Wiley and Sons, 1990.

[7] Bhattacharjee, S. *Filter Approaches to Stochastic Dynamic Analysis of Compliant Offshore Structures,* Ph.D. Thesis, Department of Mechanical Engineering, Rice University, Houston, TX March 1990.

[8] Kanegaonkar, H.B., Haldar, A., "Non-Gaussian stochastic response of nonlinear compliant platforms," *Probabilistic Engineering Mechanics,* pp. pp. 38-45, Vol. 2, 1987.

[9] Spanos, P.T.D., "An approach to calculating random vibration integrals," *Journal of Applied Mechanics,* Vol. 54, pp. 409-413, 1987.

INTELLIGENT SENSING AND CONTROL OF GAS METAL ARC WELDING

H. B. Smartt and J. A. Johnson
Idaho National Engineering Laboratory
Idaho Falls, Idaho 83415-2210, U.S.A.

ABSTRACT

Intelligent sensing and control is a multidisciplinary approach that attempts to build adequate sensing capability, knowledge of process physics, control capability, and welding engineering into the welding system such that the welding machine is aware of the state of the weld and knows how to make a good weld. The sensing and control technology should reduce the burden on the welder and welding engineer while providing the great adaptability needed to accommodate the variability found in the production world.

This approach, accomplished with application of AI techniques, breaks the tradition of separate development of procedure and control technology.

INTRODUCTION

Conventional, automated processing generally involves sophisticated sensing and control techniques applied to various processing parameters. In arc welding, for example, these parameters may include current, voltage, welding speed, and various other factors deemed important to process reproducibility. The attributes actually desired in the product, such as properties and quality, are normally controlled by some form of statistical process control. Thus, we can think of conventional practice in terms of real control (implemented by hardware-based systems) applied to the process, and virtual control (implemented by people/paper systems) applied to the product.

The prime objective of intelligent sensing and control is to make a good product the first time. The approach involves integrating off-line inspection into the process via sensors for both process and product state, in combination with appropriate control technology to drive the product state to the desired point. Of course, the objective is the same as for existing conventional technology; the difference is that the time constant of a real control loop should be orders of magnitude shorter than that of a virtual loop. Thus, fewer rejects should be produced. In addition, intelligent sensing and control should be less expensive, mainly due to reduced labor cost.

We can now examine the tools necessary to implement intelligent sensing and control. These include (in no particular order) control theory, process modeling, sensing, and artificial intelligence, in addition to the normal tools of welding engineering and materials science.

Control Theory provides a formal means of developing a strategy to obtain the desired product state and suitable process dynamics. The foundation of control theory is a body of techniques that allow convergence, stability, robustness, frequency response, and other factors to be predicted and obtained. Thus, design of a controller has a mathematical engineering basis.

Process Modeling provides a means of incorporating both first principal and empirical information into a control strategy. Models may be used off-line to evaluate and tune a controller in a simulation. They may also be used to develop transfer functions of a process for use in formal controller design, or to provide maps

between input and output parameters. Process models are an important bridge between what is known and what is desired.

Process Sensing is the necessary means of identifying the state of both the process and the product. Controllers operate by comparing actual process output (in terms, for example, of product attributes) to the desired output. The difference or error is used to calculate the appropriate control input to the process. Thus sensors are normally needed for each of the various parameters or attributes chosen as system outputs in the design of the controller.

Artificial Intelligence (AI) is a body of techniques that attempts to mimic biological intelligence. These various techniques, including expert systems, artificial neural networks, and fuzzy logic, are used for a variety of interesting applications including image and signal processing, selection of nominal parameters, and dynamic control. More is said about this topic in the next section.

Intelligent sensing and control involves application of all of the above tools to control of both the desired operation of a process and the attributes of the product of that process.

THE ROLE OF AI TECHNIQUES

Before turning to the question of what this means to welding, we believe it is worth while to comment on some of the AI techniques mentioned above. Although a well balanced paper would expound on all four of the above tools, we have selected AI techniques for comment because we believe that some simple observations are of sufficient value to include in this short paper. We apologize for the fact that what follows is in no way a comprehensive review of this vast field; there is a story to tell, and we have selected references accordingly.

Papers and presentations on AI methods and applications generally discuss the nuts and bolts of the machine, but gloss over what it is that the machine does. Consider, for example, artificial neural networks. Anderson and Rosenfeld [1] and more recently Simpson [2] provide excellent starting places for the study of artificial neural networks, while MacGregor presents an excellent overview of biological neurons and neural networks [3]. Anderson and Rosenfeld provide an historical view, and Simpson gives possibly the best presentation now available of the most important algorithms being used, based on classification into four groups in terms of feed forward and recursive network structures and supervised and unsupervised learning methods.

Perhaps the most popular algorithm for artificial neural networks is the feed forward network using backpropagation for learning of the interconnection weights. In one of their introductory papers on backpropagation, Rumelhart et al. [4] employ the XOR Boolean logic problem as a test case. In this problem, there are two binary inputs and one binary output; if the inputs are both 1s or 0s, the output will be 0 and if the two inputs are not both the same (i.e.. 1,0 or 0,1) the output is 1. A feed forward network having two binary inputs (X1 and X2), three artificial neurons in the hidden layer, and one output neuron, where all neurons include the nonlinear sigmoid activation function, is taught the XOR problem. This network is able to learn a good solution to the problem in several hundred iterations of supervised learning. The neural network output OUT in the above example may be calculated algebraically from:

$$OUT = f(V1*f(X1*W11 + X2*W12) + V2f(X1*W21 + X2*W22)) \quad (1)$$

where f is the sigmoid function operator.

The resulting output is plotted as a function of X1 and X2 in Figure 1, a plot of the input to output mapping function learned by the network. The function is continuous and is a good solution for binary inputs, but a poor solution for intermediate inputs. The main point to be made is that an artificial neural network is a mechanism for generating an input to output mapping function, given a set of discrete (not necessarily binary) data points. This point is discussed by Gallant and White [5] and more recently Cardaliaguet and Euvrard [6]. As is noted below, fuzzy logic systems are also mechanisms for generating an input to output mapping function, and a strong argument may be made that expert systems are also such mechanisms, though in this latter case the mapping function is not continuous.

Expert systems and fuzzy logic systems both differ from artificial neural networks in that they use conditional logic statements as the input data. The difference between the two methods is that expert systems normally give yes/no or black/white types of output, whereas fuzzy logic systems admit degrees of maybe or levels of gray as inputs and outputs.

Consider the XOR problem discussed above. The conditional logic statements were stated as: **if** the inputs are both 1s or 0s, **then** the output will be 0; and **if** the two inputs are not both the same (i.e.. 1,0 or 0,1), **then** the output is 1. A fuzzy logic solution [7] is shown in Figure 2. This solution has the interesting features that completely ambiguous inputs (X=0.5, Y=0.5) give an ambiguous output (Z=0.5), and the plot is symmetrical. This is a much more logical answer than that obtained by the artificial neural network above. Indeed, it is an enlightening exercise to attempt to reproduce the symmetry of the fuzzy logic solution with an artificial neural network (but one that we will leave to the reader).

At least two additional observations may be made about the above example. One, that method of evaluating fuzzy logic allows one to write the solution as an algebraic function, in the same manner as was possible for the artificial neural network solution. (This is not possible for the standard fuzzy logic paradigm [8], which requires some scheme of defuzzification to calculate a solution. The conventional approach normally generates non-continuous functional solutions.) For this reason, this new method is called "continuous fuzzy logic" to differentiate it from the more standard paradigm.. Two, the solution to continuous fuzzy logic may be represented in a network form, Figure 3, in which the inputs X and Y and the output Z are as discussed above. The elements in the network hidden layer are directly associated with the mathematical evaluation of the individual **if** conditions. The hidden elements-to-output element weights are directly associated with the action to be taken. The output element contains a summation operator. (The reader may want to compare this network with the corresponding structure for conventional fuzzy logic shown in Kosko, figure 11.8, p. 392 [8].) It is probably not appropriate to call this a neural network, but it is proper to refer to it as a connectionist network, thus recognizing that it belongs to a superset that contains artificial neural networks.

In the case of an artificial feed forward neural network, the number of elements required in the hidden layer(s) is determined by the number of degrees of freedom required to adequately map the control law. Determining the actual number required is generally accomplished on a trial and error basis. For a fuzzy

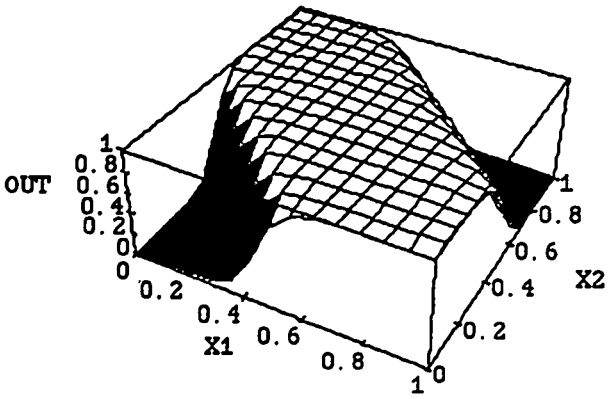


Figure 1. Plot of functional input-to-output mapping learned by feed forward neural network as solution to XOR problem.

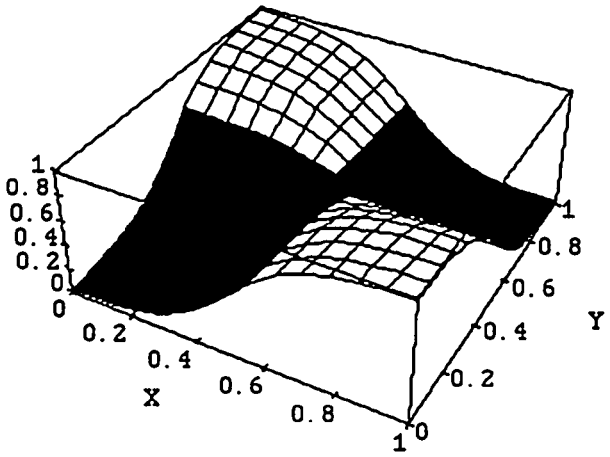


Figure 2. Continuous fuzzy logic solution to the XOR problem.

network, a hidden layer element is associated with each conditional logic rule, so establishing the number of hidden layer elements is no problem.

By now you realize that the fuzzy logic XOR solution is a mapping function. An expert system solution is also easily generated, consisting either of isolated 'peaks and holes' or of a stepped surface, depending upon how you go about it. So we see that all three methods discussed are, in fact, mapping function generators.

The next logic step involves recognition that input to output mapping functions are in fact the same as transfer functions. Transfer function based analysis and design methods form a major portion of dynamic systems analysis [9], signal analysis [10], and control theory; thus, the route to formal integration of AI techniques into these other disciplines is available, but unfortunately few have attempted to exploit it. Fortunately, Ydstie [11], and especially Narendra [12] have, and the reader may start with them to explore this interesting topic.

It may be noted that the mapping function generated by the fuzzy logic system, Figure 2, lacks the faceted appearance normally seen in such functions, for example Kosko figure 9.4a, p. 343 [8]. Figure 2 was generated using a continuous fuzzy logic algorithm [13] that is not normally seen, but which has apparently been derived at least three times by various workers. Continuous fuzzy logic readily allows generation of continuous mapping functions, at least in part, by elimination of the defuzzification step required by the more standard algorithms [8]. An important side benefit is a significant reduction in the amount of computer code required for implementation. But more important from a control standpoint is the effect of using a continuous function as a control law.

Figure 4 shows a mapping function generated using continuous fuzzy logic that describs the control law for a one-dimensional tracking problem similar to seam tracking in welding. The logic used to generate the mapping function is based on the input X being the tracking error, the input Y being the first derivative with respect to time of the tracking error, and the output Z being the control input to the system. Thus, this controller operates as a proportional-derivative controller.

Next, we will examine control of the tracking problem using continuous fuzzy logic.

By making changing the mapping function, the resulting controlled system dynamics may be tuned in a manner that is equivalent to a conventional controller. For example, Figure 5 shows tracking error (gap) as a function of time. As the function's partial derivative with respect to X is increased, the

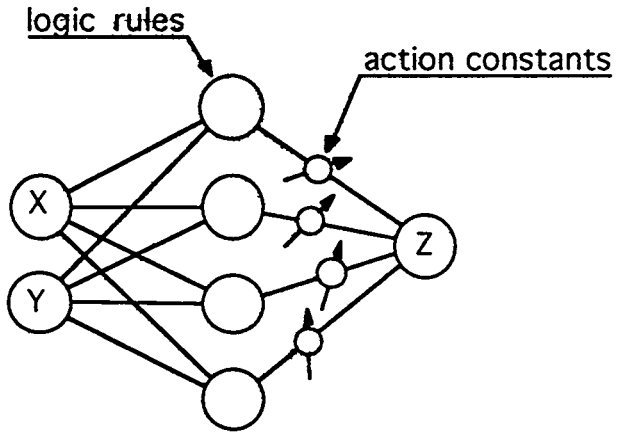


Figure 3. Network representation of continuous fuzzy logic.

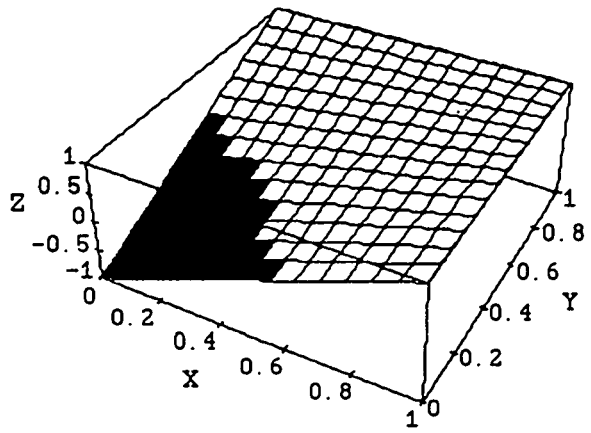


Figure 4. Control law for one-dimensional tracking generated using continuous fuzzy logic.

convergence rate increases. The effect is to increase the time for convergence until overshoot is obtained as shown in Figure 5. Figure 6 shows the steady state solution obtained (in this problem a small positive error (gap) is desired); the steady state value may be varied as the amplitude of the mapping function value is changed for X = Y = 0.

There are lessons to be learned from this simple example. One, all controllers need to be properly tuned, even intelligent ones. There has been considerable hype in the popular press to the effect that fuzzy logic controllers are the panacea for applications where the system transfer functions are not readily available, but "expert" knowledge is available. A recent IEEE video tape on fuzzy logic [14] disclosed that a certain commuter train in Japan that uses a fuzzy logic controller required approximately eight years of controller development including some 350,000 computer simulations for proper tuning. This situation would not be tolerated by industry in this country; the solution is development of engineering tools for controller tuning applicable to fuzzy logic and other AI-based controllers. Two, the partial derivatives of the mapping function with respect to the inputs are equivalent to the gains in a conventional controller. Mapping functions should thus be continuous. Considerable care must be exercised with conventional fuzzy logic controllers to obtain continuous mapping functions. Three, most learning methods for artificial neural networks teach the network the amplitude of the mapping function at a given coordinate. Methods to teach control laws should also be capable of teaching mapping function partial derivatives with respect to inputs at a given coordinate.

This discussion of applying artificial neural networks or fuzzy logic systems as controllers is now summarized. Consider a more generalized formulation of the controller transfer function as:

$$G(s) = K_1(...)_1 + K_2(...)_2 + K_3(...)_3 + ... + K_n(...)_n \quad (2)$$

Replacing such a controller by a connectionist network may be accomplished by formulating a network having an input dimension, D, associated with each of the above n gains. Also associated with each input will be a signal preprocessor $(...)_n$. Using such a formulation, given linear activation or membership functions, the derivatives of the network mapping function, F ,with respect to the network input dimensions D are:

$$\partial F/\partial Dn = K_n \quad (3)$$

where K_n is the local gain associated with the nth input dimension. The connectionist network will have an output dimension associated with each of the conventional controller outputs.

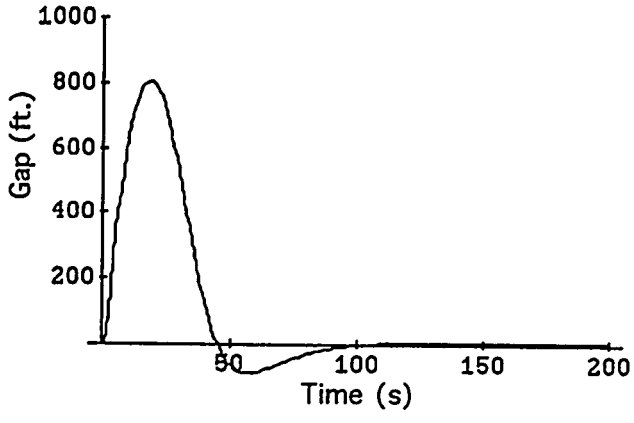


Figure 5 Tracking error (gap) showing overshoot for a high value of mapping function derivative with respect to error input.

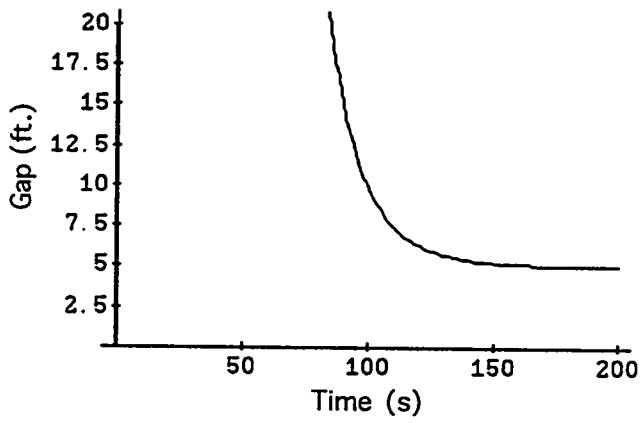


Figure 6 Tracking error (gap) showing small steady state gap for a low value of mapping function amplitude for inputs of zero.

APPLICATION TO WELDING

Recent examples of applications to welding cover the fine technical details [15,16]. We will look instead at what we consider the global problems to be, and speculate on a possible solution.

Typically, welding systems either allow the welder access to process parameters at the expense of placing a significant burden on the welder to program the system, or the system is very easy to program but the welder has limited access to the process parameters. This creates one of two problems. If the welder has easy access to the process parameters, the assumption is made that he/she has the knowledge about the physics of welding necessary to make changes to those parameters during welding to correct an undesirable situation; this is generally not true. On the other hand, systems that are easy to program generally do not provide the adaptability necessary to correct an undesirable situation during welding. The solution to both problems is identical. Do not allow the welder to change anything, but use an off-line statistical process control program to modify machine settings. Which is exactly where this paper started! Now let us suggest an alternative.

The objective may now be restated as that of developing a welding system that is very easy for the welder to program, but which has considerable adaptability to allow in-process corrections to be made for undesirable situations. (More to the point, the system should be able to predict and prevent undesirable situations.)

The alternativee approach is novel, and perhaps disturbing, for the first thing we need to do is to throw away welding procedures.

Consider the approach to development of a welding procedure. If we were doing it for the first time, we would probably want to review fundamental knowledge about the physics of heat and mass transfer in welding, microstructural development during solidification and solid state transformations, effect of thermal gradients and phase transformations on residual stresses and distortion, and also general knowledge about welding processes. From this set of data we could presumably derive a set of conditional logic statements that would at least define the qualitative characteristics of the procedure. We could also develop models describing the physics of the process. If we were able to build these rules and models into the control logic, and we can do so using fuzzy logic and artificial neural networks, then perhaps the welding machine could use this knowledge to actually develop the procedure as the weld was being made.

In order to accomplish the type of control we have been discussing, it is necessary to have a variety of sensors on the welding machine. Selection of the sensors should take into account the source of heat and mass transferred to the base metal -- melting of the base metal, dilution of the filler metal, solidification of the weld bead, microstructural development in the weld bead and heat affected zone, physical properties development, and thermomechanical distortion and residual stresses in the weldment all follow from the heat and mass transferred by the process to the weld.

Unfortunately, with few exceptions, sensors do not exist to detect weld microstructure and properties. This lack is a limit on the ultimate capabilities of sensing and control of arc welding, even for conventional control approaches. The development of advanced sensors is a significant research opportunity.

EXAMPLE

Consider the application in Einerson's paper [16]. For reasons related to the specific end product involved, it is desired to control the weld cooling rate and fill of the weld joint. The conditional logic rules are simple:

1. **If** the reinforcement is too low, **then** increase the ratio of electrode speed to welding speed.
2. **If** the reinforcement is too high, **then** decrease the ratio of electrode speed to welding speed.
3. **If** the cooling rate is too low, **then** decrease the heat transfer rate to the weld.
4. **If** the cooling rate is too high, **then** increase the heat transfer rate to the weld.

It may be shown [17] that the weld bead reinforcement (G), defined as the transverse cross-sectional area added to the weld bead by the addition of filler metal, is given by $G = \pi dS/4R$ where d is electrode diameter, S is electrode speed, and R is welding speed, for 100% deposition efficiency. The amount of heat transferred to the weld per unit length (H) is given by $H = \eta EI/R$ where η is the heat transfer efficiency, E is voltage and I is current.

The sensing requirements are defined by the logic of the problem. Two sensors are required, one to measure the transverse cross-sectional area of the weld joint and a second to measure the cooling rate of the weld bead. It is also necessary to know the welding speed, electrode speed, current, and voltage, the values of which are normally readily available in an automated arc welding system.

The difficult aspect of this example is that reinforcement and heat transfer rates are both functions of welding speed and, in gas metal arc welding, current is a function of voltage and electrode speed. This problem may be handled by deriving a model of these relationships [17], by teaching them to an artificial neural network [16], by using a look-up table, or perhaps other means. We may comment that the relationships between conventional parameters and heat and mass transfer rates are not obvious to the average welder or welding engineer. This is exactly the kind of problem that prevents the welder from adjusting heat input to the weld while maintaining constant fill rate, for example. In this work, the relationships have been used as the training set for a feedforward artificial neural network.

The resulting controller is shown in Figure 9, from [16]. The fuzzy logic controller contains simple rules that specify engineering practice. The resulting control law is tuned, as was discussed above. The artificial neural network contains (experimental) knowledge about the physics of heat and mass transfer in the process. The resulting mapping function effectively linearizes the process with respect to heat and mass transfer rates to the base metal. This system requires only that the welder set the desired weld bead cooling rate. The sensors measure the weld joint ahead of the torch and the weld bead cooling rate behind the weld pool, and the system responds to the measurement as governed by the conditional logic rules given above. This is a simple example in which knowledge of both process physics and engineering practice have been used to develop a control law and linearize a process, but the concept can be extended to include other factors.

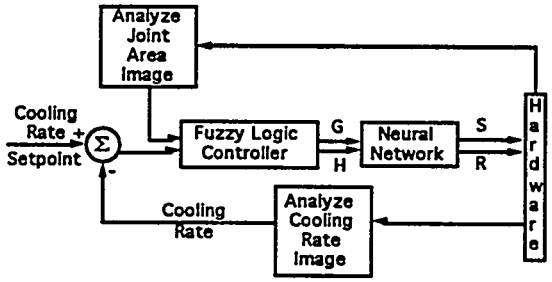


Figure 9. Process control scheme block diagram, where G is the reinforcement, H is heat input, S is electrode speed, and R is travel speed.

CONCLUSIONS

There is much more to this story. However, we will summarize by saying that what is needed is sensing and control technology that reduces the burden on the welder and welding engineer while providing the great adaptability needed to accommodate the variability found in the production world. Conventional approaches to automation of welding have been reasonably successful, but there are still significant opportunities for additional development.

It may be time to consider breaking the traditional approach that separates procedure development methods from control technology. A marriage of these two topics, accomplished with application of AI techniques, may be in order.

Finally, advanced sensor development is still needed for control of weld microstructure and properties.

ACKNOWLEDGMENTS

Appreciation is expressed to C. J. Einerson, P. L. Taylor, U.S. Wallace, and N. M. Carlson for their support. This work was supported by the U.S. Department of Energy, Office of Energy Research, Office of Basic Energy Sciences, under DOE Idaho Field Office Contract DE-AC07-76ID01570.

REFERENCES

1. J. F. ANDERSON and E. ROSENFELD, *Neurocomputing*, MIT Press, Cambridge, MA (1988).

2. P. K. SIMPSON, *Artificial Neural Systems*, Pergamon Press, New York (1990).

3. R. J. MACGREGOR, *Neural and Brain Modeling*, Academic Press, Inc., San Diego, CA (1987).

4. D. E. RUMELHART, G. E. HINTON, and R. J. WILLIAMS, "Learning Internal Representations by Error Propagation," *Parallel Distributed Processing: Explorations in the Microstructures of Cognition*, Vol. 1, ed. D. E. Rumelhart and J. L. McClelland, MIT Press, Cambridge, MA, pp. 318-362 (1986).

5. A. R. GALLANT and H. WHITE, IEEE Int. Conf. on Neural Networks, San Diego, CA, July, pp. 657-664 (1988).

6. P. CARDALIAGUET and G. EUVRARD, Neural Networks, Vol. 5, pp. 207-220 (1992).

7. H. B. SMARTT, J. A. JOHNSON, C. J. EINERSON, and G. A. CORDES, *Intelligent Engineering Systems Through Artificial Neural Networks,* ASME Press, New York, pp. 711-716 (1991).

8. B. KOSKO, *Neural Networks and Fuzzy Systems*, Prentice-Hall, Englewood Cliffs, NJ, 1992.

9. R. H. CANNON, *Dynamics of Physical Systems,* McGraw-Hill Publishing Co. (1967).

10. K. BEAUCHAMP and C. YUEN, *Digital Methods for Signal Analysis*, George Allen & Unwin, London (1979).

11. B. E. YDSTIE, Computers Chem. Engng. Vol. 14, No. 4,5 pp. 583-599 (1990).

12. K. S. NARENDRA and K. PARTHASARATHY, IEEE Trans. on Neural Networks, Vol. 1, No. 1, pp. 4-27 (1990).

13. T. TAKAGI and M. SUGENO, IEEE Trans. Systems, Man, and Cybernetics SMMC 15, pp. 116-132 (1985)

14. Institute of Electrical and Electronics Engineers (1991).

15. K. ANDERSEN, G. E. COOK, J. F. SPRINGFIELD, and R. J. BARNETT, *Intelligent Engineering Systems Through Artificial Neural Networks*, ASME Press, New York, pp. 717-727 (1991).

16. C. J. EINERSON, H. B. SMARTT, J. A. JOHNSON, P. L. TAYLOR, K. L. MOORE, 3rd Int. Conf. on Trends in Welding Research, Gatlinburg, TN, June (1992).

17. H. B. SMARTT, C. J. EINERSON, A. D. WATKINS, R. A. MORRIS, in *Advances in Welding Science and Technology*, ASM, pp. 461-465 (1986).

Welding Process Decoupling for Improved Control

David E. Hardt
Laboratory for Manufacturing and Productivity
Thomas W. Eagar
Materials Processing Center
Jeffrey H. Lang
Lawrence Jones
Laboratory for Electromagnetic and Electronic and Systems
Massachusetts Institute of Technology

Abstract

The Gas Metal Arc Welding Process is characterized by many important process outputs, all of which should be controlled to ensure consistent high performance joints. However, application of multivariable control methods is confounded by the strong physical coupling of typical outputs of bead shape and thermal properties. This coupling arises form he three dimensional thermal diffusion processes inherent in welding, and cannot be overcome without significant process modification. This paper presents data on the extent of coupling of the process, and proposes process changes to overcome such strong output coupling. Work in rapid torch vibration to change the heat input distribution is detailed, and methods for changing the heat balance between base and fill material heat are described.

Introduction

The application of control methods to welding has a long and successful history, yet the development of a fully autonomous welding process that can consistently maintain high quality welding has not been achieved. More importantly, the use of such control methods has not advanced the quality of welding through better process regulation. This failure is linked both to inadequate control approaches and to basic process limitations that control alone cannot overcome.

Welding is actually a locally applied *reprocessing* of the base material with some new alloying elements added through the wire feed. As such it is a process dominated by thermal processes. A highly diffuse energy transport, best represented by a distributed parameter model, characterizes these thermal processes. However, for the purposes of applying feedback control to the process, specific spatially distinct outputs must be identified for measurement, and these can be broken down into geometric and thermal history features as shown in Fig. 1. For example, note that the width, depth and height (W, D and H) of a bead cross section are a parametric description of a general cross-section, and the choice of these measurements is made assuming a well-behaved cross section shape. Likewise the thermal features of heat-affected zone width (HZ) are chosen assuming that its extent in the cross section is well reflected by the width on the surface. Finally, the centerline cooling rate, key to control of thermally activated processes in many material and post weld stress concentrations, is only examined at one point, which is assumed to be the critical region.

With any manufacturing process it is apparent that control can be exerted in several places, depending upon the availability of measurements and the nature of the available process inputs. The most common form of process control is in fact simply *machine control*, which can be defined as feedback control of any machine characteristic. Mechanical power, in the form of force or displacement control, is among the most common of

these and is particularly effective for serial processes where the process trajectory is a primary determinant of output geometry.

In welding, many methods of machine control have been employed, including current-controlled power supplies, automatic voltage control for Gas Tungsten Arc Welding (GTAW), servo-controller wirefeed velocity in Gas Metal Arc Welding (GMAW), and torch trajectory control. In fact, the canonical welding robot is the culmination of efforts to achieve a high level of certainty in the welding machine characteristics.

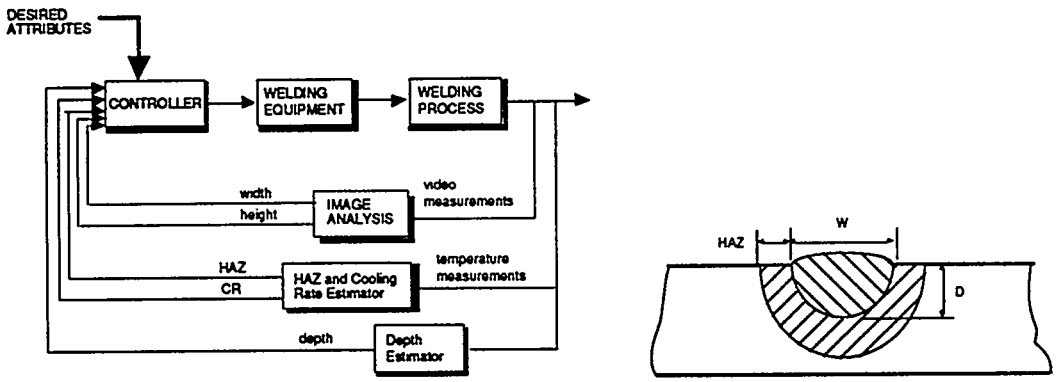


Fig. 1 The Welding MIMO Diagram and Associated Weld Cross Section

When one encounters the entire process output control problem, which can be defined as shown in Fig. 1, several problems appear. The lack of sufficient inputs in Fig. 1 becomes immediately evident. Fig. 2 also indicates that direct measurement of the outputs is often difficult, particularly with regard to the weld depth. Thus when approaching *process* control instead of machine or equipment control, several new problems arise, typically involving process coupling and inaccessible measurements. The problem of measurement has been well documented and researched, and various techniques for video (e.g. Richardson and Gutow, 1986), thermal imaging, (Kahn et al., 1986) and ultrasonic measurement (Lott, 1984; Hardt and Katz, 1984) have been proposed and executed to some degree of success. In addition, the use of on-line calibrated estimators (Song, 1992) has shown acceptable output "measurement" to be permissible. Even with such accomplishments, however, the input-output coupling inherent in this process constrains the total process control problem, and so we introduce an approach to solving this problem below.

Output Coupling in Welding

The outputs shown in Fig. 1 include the bead geometry and the thermal effects of passing a concentrated heat source past the weld joint. All five of these outputs are determined by the heat and mass transfer from the weld torch to the plate. In fact, as a first approximation it is evident that the temperature distribution set up in the weldment by the heat source determines all of these outputs.

Let us consider the simplest model of welding where no mass transfer occurs and the welding torch is modeled as the simple point source welding model first proposed by Rosenthal (1941). This model assumes pure heat conduction and a point source moving at a constant velocity. The solution of this problem takes the form (Carslaw and Jaeger, 1959):

$$T(x,y,z,t)-T_o=\int_{t=0}^{t=t'}\frac{\eta Q}{\rho c(4\alpha\pi)^{3/2}(t-t')^{3/2}}\exp\left(-\frac{(x-vt')^2+y^2+z^2}{4\alpha(t-t')}\right)dt' \qquad (1)$$

where

T = temperature in the material T_0 = initial temperature η = heat source efficiency

α = thermal diffusivity ρ = density v = source velocity

x,y,z = Cartesian coordinates, centered on the point source, with x in the direction of travel

Assuming for the moment no phase transformation in the weldment, this source sets up a set of isotherms in the weldment, that, when viewed in the reference from of the torch, resemble those shown in Fig. 3. Further examination of this model reveals that the point source mandates that (for example) pool half-width and depth will always be the same for a semi-infinite plate. A thinner plate can change this relationship, but no manner of torch modulation can. In other words, the pool geometry simply scales with heat input, but the aspect ratio of the pool cannot be controlled independently.

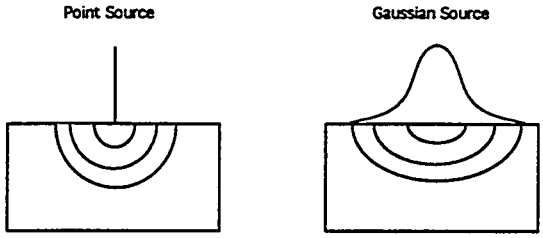


Fig 3. Isotherms Predicted by Eqn 1 (point source) and Eqn 2 (Gaussian source)

If the source is instead modeled as a distributed heat flux

$$q(x,y)=\frac{\eta Q}{\sqrt{2\pi\sigma}}\exp\left(-\frac{x^2+y^2}{2\sigma^2}\right)$$

a solution similar to that of Eqn 1 can be found (Tsai, 1983; Song, 1992):

$$T(x,y,z,t)-T_o=\int_0^t\frac{\eta Q}{\pi\rho c(2\alpha t+\sigma^2)}\exp\left(-\frac{(x-vt')^2+y^2}{4\alpha t+2\sigma^2}-\frac{z^2}{4\alpha t}\right)dt$$

In this case the pool aspect ratio is now changed, as shown in Fig 3. However, the aspect ratio is still fixed for a given heat source distribution function σ.

Despite the approximate nature of the above solutions, this strong output coupling is evident in experimental results for several different welding situations. Doumanidis and Hardt (1990a, b, c), in attempting to control the Heat-Affected Zone width (HZ) and maximum cooling rate (CR) of a weld, found that the basic input-output map of the process showed very little reachability as the inputs were varied. In fact, HZ and CR were shown to be so strongly coupled as to be uncontrollable.

Again the classical thermal conduction model provides a useful tool for examining sensitivity and decoupledness. Under the assumptions of an infinite plate geometry–homogeneous, isotropic, temperature-invariant material properties with no phase transformations–and conductive heat flow with no surface losses, the steady-state temperature field developed by either a line or point moving source can be solved for the maximum width of the T_m and T_h isotherms and centerline cooling rate at T_c. This yields the following expressions of the welding outputs as functions of the heat input Q, torch velocity v and preheat temperature T_o:

$$HZ=c_2\left(\frac{Q}{v^f}\right)^{1/n_b}\left[\left(\frac{1}{T_h-T_o}\right)^{1/n_b}-\left(\frac{1}{T_m-T_o}\right)^{1/n_b}\right] \tag{3}$$

$$CR=c_3\left(\frac{Q}{v^f}\right)^{-n_c}\left(\frac{1}{T_c-T_o}\right)^{-(n_c+1)} \tag{4}$$

where :

Q_1 = heat rate input for torch 1

Q_2 = heat rate for torch 2

v = travel speed
T_m = melting temperature
T_h = HZ critical temperature
T_c = cooling rate critical temperature
T_o = Ambient Weldment Temperature
$f = NB - (n_b - n_2) / (n_3 - n_2) = 1,$

The coefficients c_1, c_2, c_3 and the exponents n_a, n_b, n_c depend on the geometry and material of the plates as well as on the specific environmental and process conditions.

Equations 3 and 4 show that HZ and CR both strongly depend upon the factor $\frac{Q}{vf}$, the "heat per unit length". Accordingly, if only Q and v are available as inputs (which is typically the case), HZ and CR cannot be independently modulated.

In the case of geometric outputs, such simple conduction solutions are less illustrative, but empirical data is no less compelling. In an investigation of input-output modeling for Gas Metal Arc Welding, Hale and Hardt (1990a) demonstrated that the reachability of the process was very small when torch speed v and wirefeed rate (essentially Q) are inputs and when pool width, W, depth, D, or reinforcement height, H, are outputs. As shown in Fig. 4, which is based on empirical data, the input-output map for GMAW shows a very narrow range of operation when pool width and reinforcement height are considered. A similar map exists for width and depth as outputs. Notice also that the figure shows a non-unique input output mapping over a large range of operation. Finally, if we consider the process outputs to be the pool width and the heat affected zone, empirical evidence again indicates little process latitude, as shown in Fig. 5.

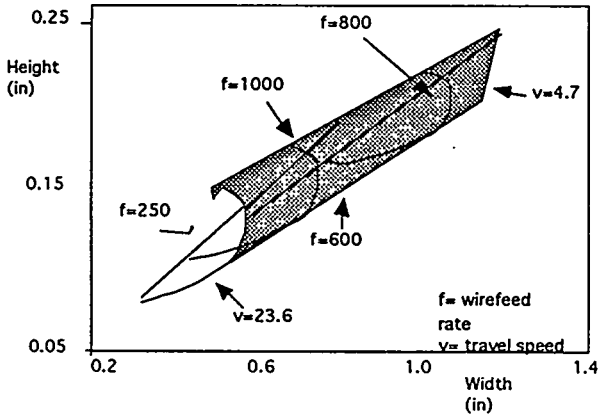


Fig. 4 Steady State Input - Output Map for Gas Metal Arc Welding (From Hale and Hardt, 1990a)

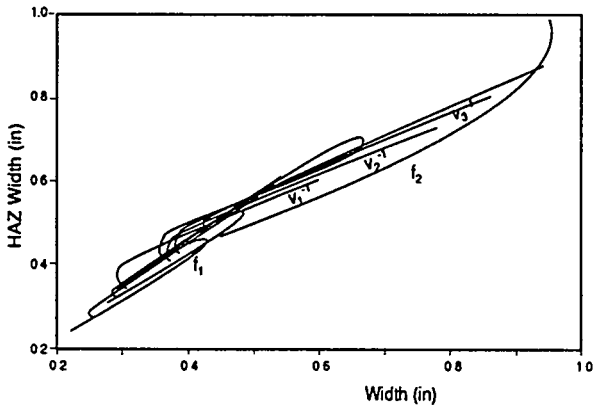


Fig. 5 Input - Output Mapping for Width and Heat Affected Zone in GMAW (Hale, 1990)

Process Control

Despite this strong coupling, some measure of multivariable feedback control can be achieved. For example, Hale and Hardt (1990b) and Song and Hardt (1992) have shown that simultaneous control of W and H and W and D can be achieved., as shown in Figs. 6 and 7.

However, the limited reachability of the process severely limits the range of operations. More importantly, it makes parameter disturbances in the process nearly impossible to eliminate. This happens because a parameter change has the effect of shifting the range of operation to a new location of the output plane. If this moves the range outside the change, as Fig. 8 shows schematically, then when a simple disturbance is introduced (in this case welding over a small void, simulating a "fill" disturbance), the closed-loop system is unable to reach equilibrium, as Fig. 9 shows, since the desired operating point is no longer in the reachable range.

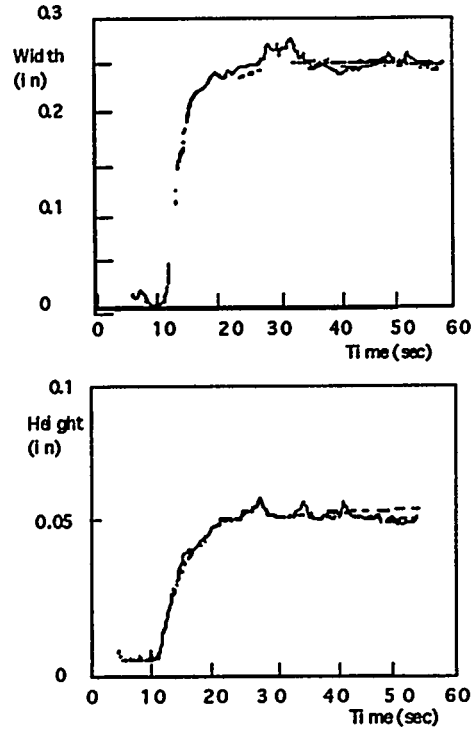


Fig. 6 Experimental and Simulated Step Responses for Zero Tracking Controller Step Commands = ΔW = 0.25 in and Δ H = 0.06in. (From Hale and Hardt, 1990b)

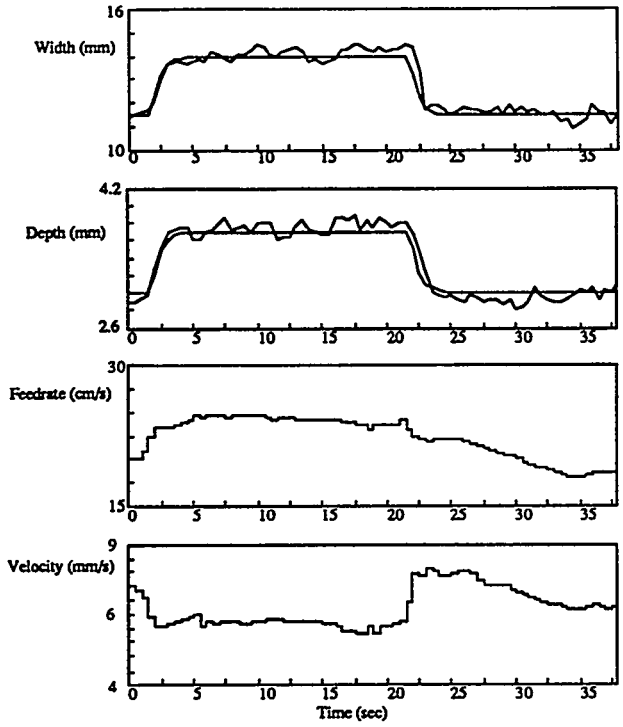


Fig 7 Step Response for a Two Variable Control System for GTAW (from Song, 1992)
(The controller is a one-step ahead direct adaptive system, using a thermally based estimator for depth feedback and video measurement for width feedback)

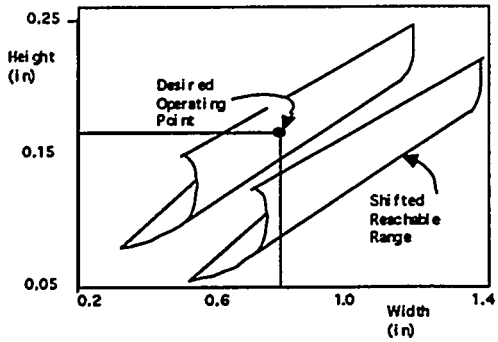


Fig. 8 Shift in Reachable Range Caused by Parameter Disturbance

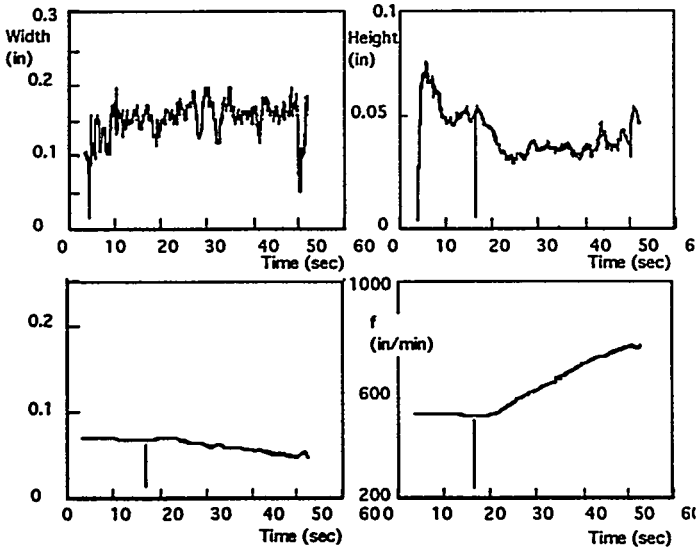


Fig. 9. Effect of Limited Reachability on Disturbance Rejection. Disturbance is a 3/16" x 3/32" groove in a flat plate

Process Modification for Improved Reachability

These results indicate that while the process of welding as currently practiced can benefit from the application of advanced feedback control methods, the true power of multivariable/ adaptive methods cannot be realized because of the highly coupled nature of the process outputs. As the brief discussion above indicated, this coupling directly results from the basic heat diffusion physics of the process. For the case of manual welding, such coupling may in fact be advantageous, since it reduces the demands on the operator. However, for truly improved process control, it is necessary to change the basic design of the process. This could be as exotic as totally separating heat source and mass transfer, as suggested by Singer(1985), or by far more straightforward methods, as suggested by the heat transfer analysis presented above.

As that simple discussion made evident, the pool geometry (or more correctly the weldment isotherms) are of a fixed geometry for a given plate material and a given heat source. However, the shift from a point source to a Gaussian-distributed source does have the effect of altering the isotherm aspect ratio. In turn this implies that the heat input distribution can be *arbitrarily* varied in real-time, and that a greater variety of temperature distributions can be achieved in the weldment. The implication follows that the pool shape and pool-heat-affected zone relationship can be modified.

Case I: Gaining Independence of the HZ and CR

As discussed earlier, if no changes are made the cooling rate (CR) and heat-affected zone (HZ) are so closely linked as to be uncontrollable in a MIMO sense. However, it has been shown (Doumanidis and Hardt, 1989) that by adding additional heat sources trailing the "primary" torch, a measure of decoupling can be achieved. While the details are omitted here, the effect is evident from examination of Eqns 2 and 3. In these, the ambient temperature T_O plays a different role in the two equations. This implies that if T_O can be modulated, a new control input will be achieved. The additional source or sources trailing the torch play this role, by effectively using the primary torch (which creates the molten zone) as a "pre-heating" source of the secondary, trailing source (See Fig. 9). This same effect can be achieved by using a heat distribution that can be varied to the rear of the weld location, as shown in Fig. 9. This is essentially an infinite set of trailing torches, and provides the maximum range of independent modulation of the HZ and CR.

When this two torch scheme was implemented, the new inputs to the process became Q_1 and Q_2.. A new input-output map could then be created, which Fig. 9 shows. Although some range of independent operation is now apparent with the two-torch method, the reachability is still quite limited.

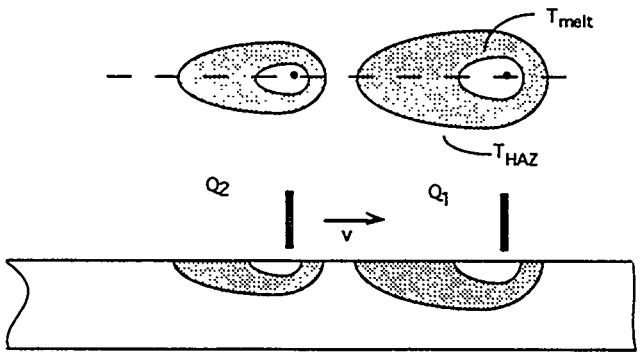


Fig. 9 Schematic of Two Source Heat Input. The Bead size (T_m isotherm) and the HZ (T_h isotherm) are completely determined by Q_1, while Q_2 is used to respond the preheat provided by Q_1 to modulate the CR.

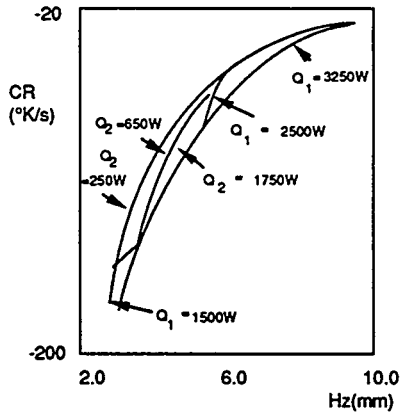


Fig. 10 Reachability for HZ and CR When a Two Source System is Used.

Note that this data was gained by having two distinct torches but also by using the concept of "mechanical multiplexing," wherein a single torch is moved rapidly between the Q_1 location and the Q_2 location. The residence time in each, integrated over time, equates to a relative heat input, provided that the traverse rates and

dwell times are considerably faster than the characteristic heat diffusion times of the weldment. As discussed in Doumanidis and Hardt (1989), this range can be improved by going to a continuously variable source distribution to the rear. Such a distribution could, in fact be attained by either magnetic deflection of an autogemous arc, or through rapid mechanical scanning of a concentrated source, such as a laser or plasma arc.

Case II: Gaining Independence of W and HZ

A desirable condition for control of welding would be to independently regulate the width of the weld bead and the attendant heat affected zone. However, as Fig. 3 points out, the W and HZ are simply different isotherms within the same temperature field. Thus they cannot be expected to be decoupled unless the distribution of heat input is varied. That this is the case is illustrated by data obtained from GMA welding tests (see Hale, 1990) where both W and HZ were measured by etching transverse cross sections. Note here the nearly compete coupling of outputs over a large range of heat input (f) and travel speed (v).

To overcome this problem, the concept of a variable heat input distribution can again be exploited. As discussed in Masmoudi and Hardt (1992), for this problem, it entails a high frequency (~ 6 Hz) *transverse* oscillation of the heat source while the torch progresses along the weld line. Although similar to the common practice of weaving, this action is intended not to weld wide joints, but rather to add variable transverse distribution to the heat source. The results of GTAW experiments (Masmoudi and Hardt, 1992) indicate that this can effectively decouple W and HZ. The change in the surface isotherm with and without weaving is shown in Fig. 11, and the increase in reachability of the process is illustrated in Fig. 12.

Discussion

It is clear that GMAW, while a productive process, is not well designed for used in a multivariable control setting, owing to the highly coupled nature of the heat and mass transfer involved. In the above three cases, the process has been modified incrementally to allow some greater range of temperature distribution variation and to afford some measure of control of the mass transfer independent of the heat input. Both of these problems can be generalized and separated if one departs from the conventional GMAW process. In our current work we are exploring just such options along several fronts:

• Active control of three-dimensional temperature distribution with a scanned point heat source: Exploration of the general distributed parameter thermal control problem, but with system identification and control methods borrowed from self-tuning control theory

• Development of spray or stream welding for independent mass transfer control: Independent creation of the liquid filler material, but at a precisely controlled mass and enthalpy rate.

• High frequency vibration of the electrode to control droplet detachment: Partial decoupling of the droplet volume and heat content from the arc heat through precise timing and feedback control of droplet detachment.

These three methods will then be exploited to create a welding process with far greater reachability and controllability than conventional processes.

Conclusions

Welding is a process that is amenable to numerous forms of feedback control. However, when one concentrates on the ***Process Control***, one must deal with the multivariable, non-linear, highly coupled nature of the process. Simple single variable control designs can perform well, especially when adaptive techniques are used to deal with the non-linear behavior, (e.g. Suzuki and Hardt, 1990). However, as detailed above, the coupling present in the existing processes, and in particular, GTAW and GMAW, precludes exploiting multivariable control to its fullest.

The origin of this problem is traced here as a one of an uncontrolled temperature distribution in the weldment. No manner of feedback can overcome this basic physical process; however, rather simple process modifications have been shown to greatly increase the process latitude. In their generalized form, they involve providing a controlled heat flux ***distribution*** into the weldment, and also perhaps, a controlled mass and heat flux from filler material.

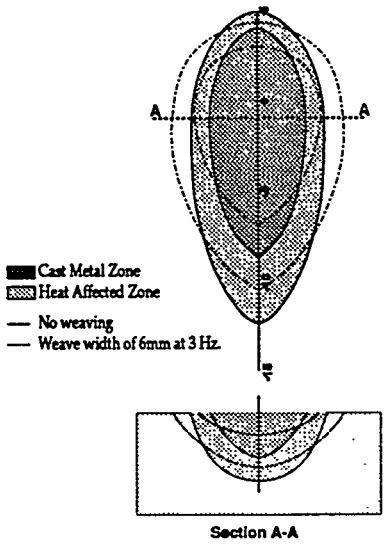


Fig. 11 Effect of High Frequency Weaving on Surface Isotherms, Predicted from a Transient Conduction Model of GTAW (From Masmoudi and Hardt, 1992)

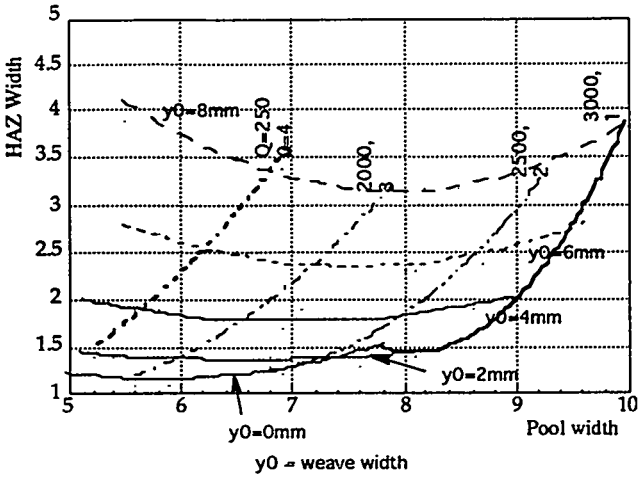


Fig. 12 Process Input Output Map For Various Weaving Parameters (from Masmoudi and Hardt(1992) Note that the original non-weaving process is indicated by a single line on this plot

References

Carslaw H.S, and Jaeger, J.C, *Conduction of Heat in Solids*, 2nd Edition, Oxford University Press, London, UK, 1959.

Doumanidis, C.C., and Hardt, D.E., "A Model for In-Process Control of Thermal Properties during Welding," *ASL*, Vol 113, 1991.

Doumanidis, C.C., and Hardt, D.E., "Multivariable Adaptive Control of Thermal Properties During Welding" *ASME Journal of Dynamic Systems Measurement and Control*, 1990

Doumanidis, C.C., and Hardt, D.E., "Simultaneous Control of Heat Affected Zone and Cooling Rate During Arc Welding," *Welding Journal*, May, 1990.

Hale, M. B., Multivariable Dynamics Modeling and Control of GMAW Weld Pool Geometry", Ph.D Thesis, MIT, Department of Mechanical Engineering, September, 1989.

Hale, M.B., and Hardt, D.E., "Multivariable Geometry Control of Welding-Part I: Process Modeling," *Symposium on Manufacturing Process Modeling and Control*, ASME, November, 1990a.

Hale, M.B., and Hardt, D.E., "Multivariable Geometry Control of Welding-Part II: Process Control," *Symposium on Manufacturing Process Modeling and Control*, ASME, November, 1990b.

Hardt, D.E., and Katz, J.M., "Ultrasonic Measurement of Weld Penetration," *Welding Journal*, Vol. 63, No. 9, September, 1984.

Lott, L.A., "Ultrasonic Detection of Molten Solid Interface of Weld Pools," *Materials Evaluation*, 42, March, 1984.

Masmoudi, R., and Hardt, D.E., "High-Frequency Torch Weaving for Enhance Process Controllability: Effect on Coupling of Pool Width and the Heat-Affected Zone Width," *Proc. Third Int. Conf.on Welding Research*, Gatlinburg, June, 1992.

Nomura, H., Satoh, Y., Tohno, K., Satoh, Y., and Kurotori, M, "Arc Light Intensity Controls Current in SA Welding System," *Welding and Metal Fabrication*, September, 1980.

Nunes A.C., "An Extended Rosenthal Model," *Welding Journal*, Vol. 62, No. 6, June, 1983.

Richardson, R.D., Gutow, R.A., Anderson, R.A., and Farson, D.F., "Coaxial Arc Weld Pool Viewing for Process Monitoring and Control," *Welding Journal* , Vol. 63, No. 3, March 1984.

Rosenthal, D., "Mathematical Theory of Heat Distribution During Welding and Cutting," *Welding Journal*, Vol. 20, No. 5, 1941,

Singer, A.R.E., "Spray Welding of Metals," UK Patent Application, 21422858 A, January, 1985.

Song, J-B. and Hardt, D.E., "Development of a Heat-Transfer Based Depth Estimator for Real-Time Welding Control," *Symposium on Manufacturing Process Modeling and Control*, ASME, November, 1990.

Song, J-B., "Multivariable Adaptive Control for GMA Welding using a Thermally Based Depth Estimator," Ph.D. Thesis, MIT, Department of Mechanical Engineering, January, 1992.

Suzuki, A., Hardt, D.E., and Valavani, L. "Application of Adaptive Control Theory to On line GTA Weld Geometry Regulation," *ASME J. of Dynamic Systems Measurement and Controls*, Vol. 113, March, 1990.

Tsai, N., "Heat Distribution and Weld Bead Geometry in Arc Welding," Ph.D. Thesis, MIT, Department of Materials Science and Engineering, April, 1983.

Vroman A.R., and Brandt, H., "Feedback Control of GTA Welding Using Puddle Width Measurement," *Welding Journal*, September, 1978, pp. 742-746.

Final List of Participants

ELEVENTH SYMPOSIUM ON ENERGY ENGINEERING SCIENCES

May 3-5, 1993

Argonne National Laboratory
Argonne, Illinois

Jan D. Achenbach
Center for Quality Engineering & Failure Prevention
Northwestern University
2137 Sheridan Road
Evanston, IL 60208-3020

Andreas Acrivos
The Levich Institute
City College of CUNY
Steinman Hall - T 1M
138th Street at Convent Avenue
New York, NY 10031

David M. Barnett
Department of Materials Science & Engineering
Stanford University
Building 550, Rm. 550 K
Stanford, CA 94305-2205

Howard Brenner
Department of Chemical Engineering
Massachusetts Institute of Technology
Cambridge, MA 02139

Shoei-Sheng Chen
Energy Technology Division
Argonne National Laboratory
Building 335
9700 South Cass Avenue
Argonne, IL 60439

James S. Coleman, ER-15
Office of Energy Research
U.S. Department of Energy
Washington, DC 20585

Subhendu Datta, ER-15
Division of Engineering & Geosciences
U.S. Department of Energy
Office of Basic Energy Sciences
Washington, DC 20585

Monte K. Drost
Energy Sciences Department
Battelle Pacific Northwest Laboratories
P.O. Box 999, K5-20
Richland, WA 99352

Daniel Frederick
Department of Engineering Science & Mechanics
Virginia Polytechnic Inst. & State University
Blacksburg, VA 24061

Huajian Gao
Department of Mechanical Engineering
Stanford University
267 Durand
Stanford, CA 94305-4040

Fred M. Glaser
Office of Fossil Energy
U.S. Department of Energy
Washington, DC 20585

Alan L. Graham
Department MEE-9
Los Alamos National Laboratory
MS 6789
Los Alamos, NM 87545

David E. Hardt
Laboratory for Manufacturing & Productivity
Department of Mechanical Engineering
Massachusetts Institute of Technology
77 Massachusetts Avenue, Bldg. 35-234
Cambridge, MA 02139

Alan R. Kerstein
Department 8351
Sandia National Laboratories
P.O. Box 969
Livermore, CA 94551-0969

Yukikazu Iwasa
Francis Bitter National Magnet Laboratory
Massachusetts Institute of Technology
Room NW14-3101
Cambridge, MA 02139-4307

Jin O. Kim
Ctr. for Quality Engineering and Failure Prevention
Northwestern University
2137 North Sheridan Road
Evanston, IL 60208-3020

Kailasam R. Iyer
Engineering & Environmental Science Division
U.S. Army Research Office
P.O. Box 12211
Research Triangle Park, NC 27709-2211

Gordon S. Kino
Ginzton Laboratory 4085
Stanford University
Stanford, CA 94305

Eric H. Jordan
Department of Mechanical Engineering
The University of Connecticut
191 Auditorium Road
Storrs, CT 06269-3139

Dusan Krajcinovic
Dept. of Mechanical and Aerospace Engineering
Arizona State University
Tempe, AZ 85287-6106

Mark Kachanov
Department of Mechanical Engineering
Tufts University
Medford, MA 02155

Erhard Krempl
Dept. of Mech. Eng., Aeronautical Eng. & Mechanics
Rensselaer Polytechnic Institute
Troy, NY 12180-3590

Leon M. Keer
Department of Civil Engineering
Northwestern University
2145 Sheridan Road
Evanston, IL 60208

Harri K. Kytömaa
Department of Mechanical Engineering
Massachusetts Institute of Technology
77 Massachusetts Avenue
Cambridge, MA 02139

Wing K. Liu
Department of Mechanical Engineering
Northwestern University
2145 Sheridan Road
Evanston, IL 60208

Philip C. Michael
Francis Bitter National Magnet Laboratory
Massachusetts Institute of Technology
NW14-2506
77 Massachusetts Avenue
Cambridge, MA 02139

Oscar Manley, ER-15
Division of Engineering and Geosciences
U.S. Department of Energy
Office of Basic Energy Sciences
Washington, DC 20585

Michael J. Miksis
Dept. of Engineering Sciences & Applied Math.
Northwestern University
2145 Sheridan Road
Evanston, IL 60208

Isaac D. Mayergoyz
Electrical Engineering Department
University of Maryland
College Park, MD 20742

Michael R. Moldover
NIST
Building 221, Rm. A105
Gaithersburg, MD 20899

Frank A. McClintock
Department of Mechanical Engineering
Massachusetts Institute of Technology
77 Massachusetts Avenue, Rm. 1-304
Cambridge, MA 02139

Joseph O'Gallagher
Department of Physics
The University of Chicago
5720 South Ellis Avenue
Chicago, IL 60637

Mark J. McCready
Department of Mechanical Engineering
University of Notre Dame
171 Fitzpatrick Hall
Notre Dame, IN 46556

Andrea Prosperetti
Department of Mechanical Engineering
Johns Hopkins University
121 Latrobe Hall
Baltimore, MD 21218

Mark O. McLinden
Thermophysics Division
NIST
MS 838.02
325 Broadway
Boulder, CO 80303

Walter G. Reuter
EG&G Idaho, Inc.
Idaho National Engineering Laboratory
P.O. Box 1625
Idaho Falls, ID 83415-2218

Horst J. Richter
Thayer School of Engineering
Dartmouth College
Hanover, NH 03755-8000

Stuart C. Snyder
EG&G Idaho, Inc.
Idaho National Engineering Laboratory
P.O. Box 1625
Idaho Falls, ID 83415

Irven H. Rinard
Department of Chemical Engineering
City College of CUNY
138th Street at Convent Avenue
New York, NY 10031

Pol D. Spanos
Department of Mechanical Engineering
Rice University
6100 South Main Street
Houston, TX 77005

Stanislav I. Rokhlin
Department of Welding Engineering
The Ohio State University
190 West 19th Avenue
Columbus, OH 43210

Clifford M. Surko
Physics Department 0319
University of California-San Diego
LaJolla, CA 92093

Martin J. Sablik
Dept. of Instrumentation and Space Research
Southwest Research Institute
P.O. Drawer 28510
San Antonio, TX 78228-0510

Ken L. Telschow
EG&G Idaho, Inc.
Idaho National Engineering Laboratory
P.O. Box 1625
Idaho Falls, ID 83415-2209

Reuel Shinnar
Department of Chemical Engineering
City College of New York
110 Ash Drive
Great Neck, NY 11021

Timothy R. Troutt
Department of Mechanical Engineering
Washington State University
Pullman, WA 99164-2920

Herschel B. Smartt
Idaho National Engineering Laboratory
P.O. Box 1625
Idaho Falls, ID 83415-2210

Kevin P. Walker
Engineering Science Software, Inc.
315 Log Road
Smithfield, RI 02917

Zellman Warhaft
Department of Mechanical & Aerospace Engineering
Cornell University
244 Upson Hall
Ithaca, NY 14853-7501

Carl Youngdahl
Reactor Engineering Division
Argonne National Laboratory
Building 331
9700 South Cass Avenue
Argonne, IL 60439

Peter C. Wayner, Jr.
Department of Chemical Engineering
Rensselaer Polytechnic Institute
Troy, NY 12180-3590

Herbert M.C. Yuen
Department of Mechanical Engineering
Northwestern University
The Technological Institute
2145 Sheridan Road
Evanston, IL 60208

James R. Welty
Department of Mechanical Engineering
Oregon State University
Corvallis, OR 97331

Gerald Wilemski
PSI Technology
Physical Sciences Inc.
20 New England Business Center
Andover, MA 01810

Roland Winston
Department of Physics
The University of Chicago
5640 South Ellis Avenue
Chicago, IL 60637

Barbara Wyslouzil
PSI Technology
Physical Sciences Inc.
20 New England Business Center
Andover, MA 01810