Research Directions in Computational Mechanics

U.S. National Committee on Theoretical and Applied Mechanics, Commission on Engineering and Technical Systems, National Research Council

ISBN: 0-309-57251-7, 144 pages, 6 x 9, (1991)

RESEARCH DIRECTIONS IN COMPUTATIONAL MECHANICS

U.S. National Committee on Theoretical and Applied Mechanics
Manufacturing Studies Board
Commission on Engineering and Technical Systems
National Research Council

NATIONAL ACADEMY PRESS
Washington, D.C. 1991

NOTICE: The project that is the subject of this report was approved by the Governing Board of the National Research Council, whose members are drawn from the councils of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine. The members of the board responsible for this report were chosen for their special competences and with regard for appropriate balance.

This report has been reviewed by a group other than the authors according to procedures approved by a Report Review Committee consisting of members of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine.

The National Academy of Sciences is a private, nonprofit, self-perpetuating society of distinguished scholars engaged in scientific and engineering research, dedicated to the furtherance of science and technology and to their use for the general welfare. Upon the authority of the charter granted to it by the Congress in 1863, the Academy has a mandate that requires it to advise the federal government on scientific and technical matters. Dr. Frank Press is president of the National Academy of Sciences.

The National Academy of Engineering was established in 1964, under the charge of the National Academy of Sciences, as a parallel organization of outstanding engineers. It is autonomous in its administration and in the selection of its members, sharing with the National Academy of Sciences the responsibility for advising the federal government. The National Academy of Engineering also sponsors engineering programs aimed at meeting national needs, encourages education and research, and recognizes the superior achievements of engineers. Dr. Robert M. White is president of the National Academy of Engineering.

The Institute of Medicine was established in 1970 by the National Academy of Sciences to secure the services of eminent members of the appropriate professions in the estimation of policy matters pertaining to the health of the public. The Institute acts under the responsibility given to the National Academy of Sciences by its congressional charter to be an adviser to the federal government and, upon its own initiative, to identify issues of medical care, research, and education. Stuart Bondurant, M.D., is acting president of the Institute of Medicine.

The National Research Council was organized by the National Academy of Sciences in 1916 to associate the broad community of science and technology with the Academy's purposes of furthering knowledge and advising the federal government. Functioning in accordance with general policies determined by the Academy, the Council has become the principal operating agency of both the National Academy of Sciences and the National Academy of Engineering in providing services to the government, the public, and the scientific and engineering communities. The Council is administered jointly by both Academies and the Institute of Medicine. Dr. Frank Press and Dr. Robert M. White are chairman and vice chairman, respectively, of the National Research Council.

This project was sponsored by the National Science Foundation, contract number INT-9016207-C.

Library of Congress Catalog Card Number 91-67466

International Standard Book Number 0-309-04648-3

Limited report copies available from:
National Research Council
The U.S. Nat'l Committee for Theoretical and Applied Mechanics - HA270
2101 Constitution Avenue, N.W.
Washington, DC 20418
202/334-2570

Additional copies for sale from: The National Academy Press 2101 Constitution Avenue, N.W. Washington, DC 20418 202/334-3313 1-800-624-6242

Printed in the United States of America

S-490

U.S. NATIONAL COMMITTEE ON THEORETICAL AND APPLIED MECHANICS (USNC/TAM)

MEMBERS OF THE INTERNATIONAL UNION OF THEORETICAL AND APPLIED MECHANICS (USNC/TAM EX-OFFICIO MEMBERS)

BRUNO A. BOLEY (*Personal Member*), Columbia University, New York, NY
DANIEL C. DRUCKER (*Bureau Member*), University of Florida, Gainesville, FL
JAN ACHENBACH (*Congress Committee Member*), Northwestern University, Evanston, IL
ANDREAS ACRIVOS (*General Assembly Member*), City College of City University of New York, New York, NY
STEPHEN H. CRANDALL (*General Assembly Member*), Massachusetts Institute of Technology, Cambridge, MA

Staff

THOMAS C. MAHONEY, Director, Manufacturing Studies Board
DANA G. CAINES, Staff Associate

MANUFACTURING STUDIES BOARD

Staff

VERNA J. BOWEN, Staff Assistant
DANA G. CAINES, Staff Associate
LUCY V. FUSCO, Staff Assistant
THEODORE W. JONES, Research Associate
THOMAS C. MAHONEY, Director
KERSTIN B. POLLACK, Deputy Director, and Director of New Program Development
MICHAEL L. WITMORE, Research Assistant

ACKNOWLEDGMENTS

The U.S. National Committee on Theoretical and Applied Mechanics thanks the following members of The Subcommittee on Research Directions in Mechanics for their individual commitment to and extra effort in the production of this report:

J. Tinsley Oden (*Chair*)
Andreas Acrivos
Ted Belytschko
Thomas Geers
Sidney Liebovich
Marshall Tulin
Milton Van Dyke

The committee also wishes to recognize those who authored the appendices and thank them for their hard work and invaluable participation.

D. Anderson, University of Texas at Arlington
H. Armen, Grumman Aircraft Systems
I. Babuska, University of Maryland
R.E. Ewing, University of Wyoming
S.J. Fenves, Carnegie Mellon University
J.E. Flaherty, Rensselaer Polytechnic Institute
A.F. Ghoneim, Massachusetts Institute of Technology
R. Glowinski, University of Houston
P. Haff, Duke University
A. Needleman, Brown University
A.K. Noor, University of Texas
S. Orszag, Princeton University
P.J. Rabier, University of Pittsburgh
W.C. Rheinboldt, University of Pittsburgh
O. Richmond, Alcoa Aluminum Company of America
M. Shinozuka, Princeton University
G. Turkiyyah, Carnegie Mellon University
O. Walton, Lawrence Livermore Laboratories

In addition, the committee gratefully acknowledges the contributions of many more individuals than can be conveniently listed here. Many colleagues gave helpful advice and suggested improvements to the report. Their participation in reviewing drafts of the report and offering constructive criticism is greatly appreciated.

FOREWORD

This document is the first in a planned series of five reports on research directions and trends in mechanics. Besides computational mechanics, the series will cover research directions in fluid mechanics, solid mechanics, experimental mechanics, and dynamics and control. The committee will assess the state of the art for each field and bring attention to developments that will allow the United States to maintain its leadership status and technical preeminence in an increasingly fiercely competitive and politically uncertain world.

This document describes current trends and future research directions in computational mechanics, as well as the status of computational mechanics in the United States. It suggests directions for further research, particularly in areas where gaps exist in current knowledge and major advances are crucial to continued technological developments in the United States. It consists of an Executive Summary and a collection of fourteen appendices. The Executive Summary is provided for readers not familiar with the technical aspects of the subject matter. The definition of computational mechanics, its importance in developing technology, and a summary of recommendations on principal areas of research requiring attention and support over the next decade are presented in the Executive Summary. More detailed technical discussions of specific research areas are given in the appendices.

If measured in terms of impact on advances in modern technology, the subjects of mechanics and its recent offspring, computational mechanics, are among the most important areas of research and development in the industrialized world. The subject is, arguably, the key to many further developments in manufacturing, robotics, defense systems, and other fields. Computational mechanics originated in the United States, grew out of the development of electronic computation in this country, and is basic to future growth in important segments of the economy. It is a field whose welfare is often taken for granted by major federal agencies that, some critics say, are more content with promoting the development of computing machines than with the development of conceptual tools to apply the machines to meaningful problems. In the meantime, Japan and Western Europe are investing heavily in computational mechanics and are gaining prominence both in expertise and in the world market share of computational methods and software.

CONTENTS

EXECUTIVE SUMMARY

COMPUTATIONAL MECHANICS: THE SUBJECT

What is it that engineers and scientists do that accounts for the remarkable advances in transportation, communication, manufacturing, and technology since the beginning of the industrial revolution? Without doubt, one of the most important things they do is to model natural phenomena. They develop conceptual and mathematical abstractions to simulate physical events. According to the late John Von Neumann,

> "The sciences do not try to explain, they hardly even try to interpret, they mainly make models. By a model is meant a mathematical construct which, with the addition of certain verbal interpretations, describes observed phenomena. The justification of such a mathematical construct is solely and precisely that it is expected to work."

Many of these mathematical models are based on fundamental scientific laws, axioms of physics exacted from centuries of research on the behavior of mechanical systems under the action of natural forces. Today this subject is referred to simply as mechanics—a term that encompasses broad fields of science concerned with the behavior of fluids, solids, and complex materials. Mechanics is vitally important to virtually every area of technology and remains an intellectually rich subject taught in all major universities. It is the focus of research in schools of mechanical, civil, and aerospace engineering, and engineering science and mechanics, as well as petroleum and chemical engineering and applied mathematics and physics. Time and again, investments in mechanics research have produced enormous dividends to the industrialized world.

All mechanics models are characterized mathematically by very complex expressions, many of which, until the advent of electronic computation, stood outside the reach of the scientific and engineering communities. Over the last three decades, however, the computer has made it possible to solve many problems of mechanics, dramatically expanding the

capabilities for mathematical modeling. It has made possible the detailed analysis and design of a multitude of new products and an understanding of many natural phenomena elevated the standard of living throughout much of the world. There now exists a new and growing body of knowledge connected with the use of computational methods and devices to analyze the mechanics models—computational mechanics.

MAJOR INDUSTRIES AND TECHNOLOGICAL AREAS AFFECTED BY COMPUTATIONAL MECHANICS

Several major technological areas are heavily dependent on the continued growth of computational mechanics. These include manufacturing process and product simulation, space exploration and defense, environmental phenomena, and fundamental scientific research. Further comments on each of these areas follows.

Manufacturing: Computer Simulation of Materials, Processes, and Products

Many of the great industrial achievements made in the United States in recent years have been possible because of the appearance of new and innovative manufacturing processes that take advantage of new materials. The design of most of these processes would not have been possible without detailed computer simulated material behavior under anticipated lifetime conditions and the products' desirable features. Of the myriad examples that could be cited, the following are some of the manufacturing applications of computational modeling:

1. *Plastics and injection molding*. Computer simulations of molten plastics and the formation of plastic parts represent a tool used with increasing frequency in a growing number of manufacturing applications.
2. *Computer chips and components*. The design and analysis of computer chips, computer components, microchip interconnections, and the overall integrity of computer boards subject to mechanical and thermal effects require computer modeling of the anticipated stress, deformation, and temperature in the design of stable and robust computer parts.
3. *Biomedical devices*. Prostheses, artificial organs, hospital equipment, and other devices that must function reliably in

life-threatening circumstances are manufactured under conditions simulated by computer modeling techniques.

4. *Sheet metal forming*. Simulations of residual stresses, cracks, fatigue life, and other properties of metal products produced in sheet metal formation are widely used, particularly in the automobile and aircraft industries, for the design of parts fabricated from sheet metal.
5. *Crashworthiness of automobiles*. Recent advances in simulation of automobile crashes have made it possible to test automotive designs at a fraction of the cost and time required to do full-scale vehicle testing.
6. *Conventional aircraft and aerospace vehicles*. Perhaps no industry is more dependent on computational modeling than the modern aerospace industry. It relics on computer simulations of high-speed flow around aircraft and space vehicles as a key design tool. It is virtually impossible to check the performance of hypersonic aircraft and spacecraft by full-scale testing, and the only means to determine the merits of specific design iterations is through computer simulation.
7. *Combustion phenomena*. The design of internal combustion engines, solid propellant rocket motors, and thermal energy conversion processes is now being done through numerical modeling of the underlying fluid mechanics and chemical reaction phenomena.

Space Exploration and Military Defense

The use of computational mechanics in contemporary aerospace industries is essential in the design and control of space stations and space vehicles. Detailed computer simulations play a crucial role in the behavior assessment of these space structures in orbit. Similar types of computer simulations are under way to study the design and performance issues in orbital mechanics, submarine technology, and the function of other defense systems in simulated real-life environments. For instance, advances have been made in computer simulation of the mechanics of nuclear and nonnuclear explosions, and the impact of projectiles (e.g., birds and meteors) on military and nonmilitary targets. It is hoped that computer modeling can displace the expensive testing of such events.

Environmental Phenomena

Computational mechanics is increasingly important in simulating both large-and small-scale phenomena affecting the environment, for instance, pollution control. The dispersion of contaminants in the atmosphere, food, agricultural products, aquifers, and oceans and estuaries is governed by physical processes that can be modeled. Computational mechanics will play a fundamental role in developing the mechanisms to control, contain, and abate environmental pollution. Computational mechanics is also being applied to fossil energy recovery and conversion. Computer simulation of oil and gas fields and implementation of advanced oil recovery methods could help increase current oil supplies. Simulation of chemical processes, fluidized beds, energy structure behaviors, and related technologies could significantly enhance energy conservation efficiencies.

Fundamental Scientific Research

There is a growing use of computational mechanics techniques as a fundamental research tool. This ranges from the use of computer modeling for design of study experiments for new material to the materials performance response to stress, damage, and deformation under various loading conditions. Events that occur on submicrosecond time scales are difficult to study in depth in a laboratory but can be modeled computationally and studied on whatever time scale the investigator chooses. On the other hand, events that take place over centuries cannot be studied in practical laboratory experiments but can be modeled on a computer. This includes the study of environmental phenomena such as planetary atmospheres, geological processes, earthquakes, and ice flows, as well as stars, galaxies, and various astronomical events. These examples emphasize the importance of modeling physical events.

SIGNIFICANCE OF COMPUTATIONAL MECHANICS TO NATIONAL INTERESTS

Computational mechanics is an integral and major component in many fields of engineering design and manufacturing. Major established industries such as the automobile, aerospace, chemical, pharmaceutical, petroleum, electronics, and communications, as well as emerging industries such as biotechnology, rely on computational mechanics-based

capabilities to simulate and model complex systems for the analysis, engineering, design, and manufacture of high-technology products. Detailed discussion of specific areas are given in the Appendices.

Recently, an Office of Technology Assessment (OTA) paper "Seeking Solutions: High-Performance Computing for Science" (April 1991), was requested by the Senate Committee on Commerce and Transportation and the House Committee on Science, Space, and Technology to examine high-performance computing as part of the infrastructure proposed by the White House Office of Science and Technology Policy's (OSTP) National High-Performance Computing and Communications initiative. In the OSTA paper, high-performance computing is portrayed as playing an increasingly significant role in all sectors of our economy, including manufacturing, service, agriculture, government, national security and defense. The paper states that:

> "....concern for the Japanese challenge in high-performance computing goes beyond the competitiveness of the U.S. supercomputing industry. Computational simulation in engineering design and manufacturing is becoming a major factor in maintaining a competitive posture in high-technology industries....It is in the availability and application of high-performance computing to increase productivity and improve quality where the greatest future economic benefits may lie."

The OTA paper has been reinforced by a recent report from the National Research Council's Board on Mathematical Sciences titled, "Mathematical Sciences, Technology, and Economic Competitiveness" (1991). Advances and capabilities in computational and mathematical modeling (computational mechanics) have been determined to have a direct connection to economic competitiveness. These advances and capabilities represent the tools of international competition and are important for maintaining America's supremacy in science and engineering. This is particularly significant at this time in our history, when our economy has been labeled as "an economy going nowhere," and our ability to commercialize our technological capabilities in biotechnology industries is questionable. The trends for our high technology industries have been unfavorable, as illustrated in Figure 1. These trends, identified in a 1987 report from the Council on Competitiveness, titled *Picking Up the Pace*, for several key industries, clearly show

Figure 1
Erosion of U.S. Share of Technology Markets

that as a nation we have been doing a poor job in commercializing our science and technology base.

Of particular concern to this report are trends in the computer industry. Computational mechanics play a large role in the software used by this industry. For example, a CRAY survey of supercomputer use in engineering showed that 10 percent is used in crashworthiness simulation (described in Appendix 2), and 30 percent is used in computational fluid dynamics (Appendix 12).

Software development in these fields is a native technology whose further development and application to industrial problems is inextricably and symbiotically linked to the health of the computer industry, for leadership in software is a vital element in the export of computers. Figure 2 shows that despite the U.S. computer industry's continuing heavy investment in R&D, the U.S. share of the world's computer systems market fell from about 60 percent to 40 percent between 1983 and 1989, while Japan's share rose from 8 percent to 22 percent, and Europe's share grew from 10 percent to 15 percent. Furthermore, Figure 3 illustrates that in 1990 the computer systems industry's trade balance was zero for the first time.

In March 1991 the National Critical Technologies Panel, organized under the direction of OSTP, issued a report that identified 22 technologies that, according to the panel's chairman, Dr. William Phillips, "are considered essential for the U.S. to develop in the interests of the nation's long-term security and economic prosperity."

The critical technologies identified in this panel's report are listed in Table 1 in six broad categories. It is important to note that one category, Information and Communication, is directly related to many of the R&D activities in computational mechanics, while the remaining five (and most of their individual technologies) are linked to developments and applications from the field of computational mechanics.

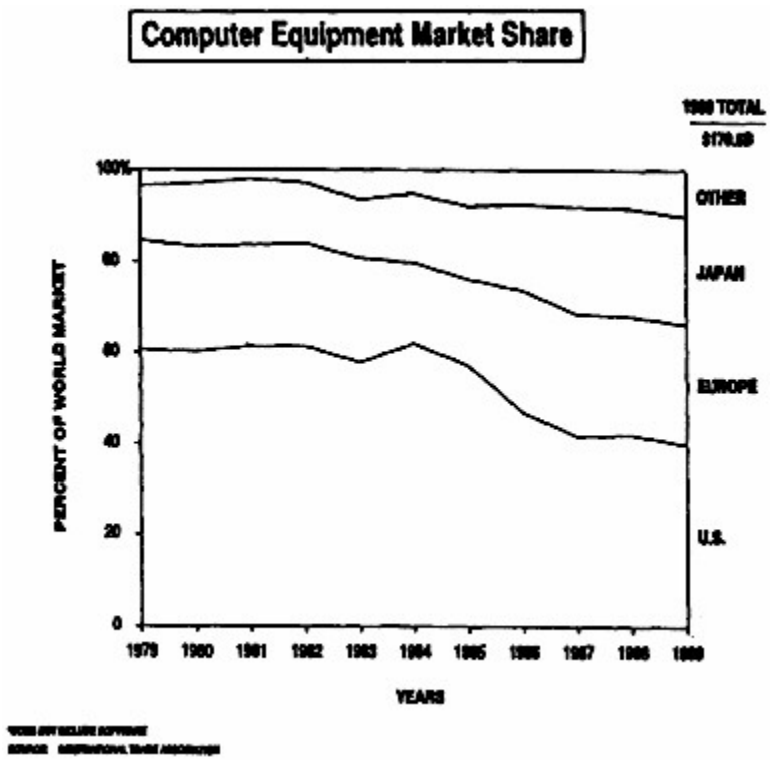

Figure 2
Computer Equipment Market Share (from a report prepared by the Computer Systems Policy Project (CSPP), "Perspective on U.S. Technology Policy, Part 1: The Federal R&D Investment.")

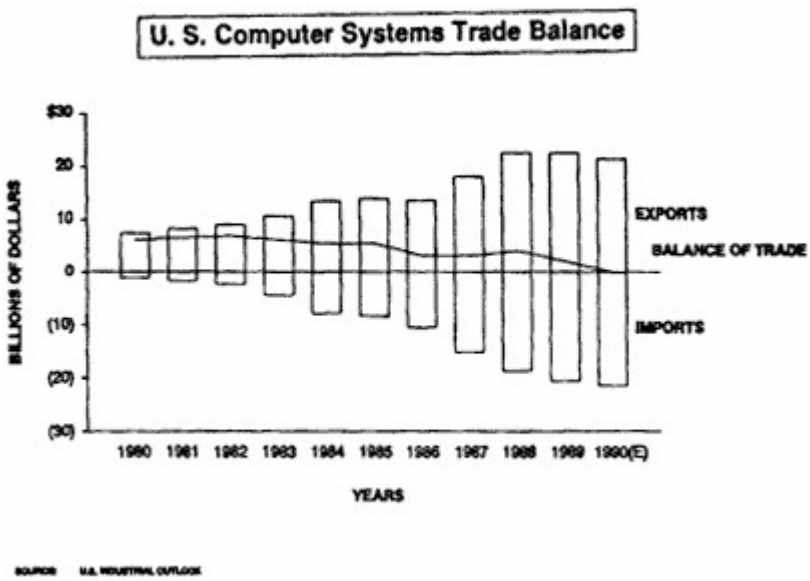

Figure 3
U.S. Computer Systems Trade Balance (from a report prepared by the Computer Systems Policy Project (CSPP), Perspective on U.S. Technology Policy, Part 1: The Federal R&D Investment)

NATIONAL CRITICAL TECHNOLOGIES

MATERIALS

- Materials synthesis and processing
- Electronic and photonic materials
- Ceramics
- Composites
- High-performance metals and alloys

MANUFACTURING

- Flexible computer integrated manufacturing
- Intelligent processing equipment
- Micro-and nanofabrication
- Systems management technologies

INFORMATION AND COMMUNICATIONS

- Software
- Microelectronics and optoelectronics
- High-performance computing and networking
- High-definition imaging and displays
- Sensors and signal processing
- Data storage and peripherals
- Computer simulation and modeling

BIOTECHNOLOGY AND LIFE SCIENCES

- Applied molecular biology
- Medical technology

AERONAUTICS AND SURFACE TRANSPORTATION

- Aeronautics
- Surface transportation technologies

ENERGY AND ENVIRONMENT

- Energy technologies
- Pollution minimization, remediation, and waste management

The High-Performance Computing and Communications Initiative

The High-Performance Computing and Communications (HPCC) initiative represents a major strategic investment for the nation that is expected to yield both economic and social returns. This initiative can have substantial impact on the field of computational mechanics, if a concurrent investment is made in computational mechanics that can lead to significant progress in the efforts critical to this nation's needs.

The HPCC initiative represents an ambitious and welcome opportunity to reverse the trends in our high-technology industries and, in particular, in our computer systems industry. This 5-year, multibillion-dollar program offers great opportunities for development and application of the technology associated with computational mechanics. Indeed, many of the goals, strategies, and proposed programs, embodied in "Grand Challenges," will depend on the successful implementation of the technological capabilities afforded by computational mechanics.

The goals, strategy, and program components of the HPCC program are presented in Table 2.

Among the various components of this program, the one most pertinent to computational mechanics falls within the largest—Advanced Software Technology and Algorithms (ASTA), which represents 41 percent of the proposed funding. One of the subcomponents of ASTA—Computational Techniques—is described as follows:

Computational techniques: The focus of the HPCC program on scalable parallel computing systems dictates that significant advances in computational techniques will be needed. The design and theory of algorithms are as important as hardware or networking improvements in reaching teraflop computational performance. Research in computational techniques will include parallel algorithms, numerical and mathematical analyses, parallel languages, computational models, and program refinement techniques.

The various activities of the HPCC initiative will contribute to a number of "grand challenge" problem areas. The grand challenges are defined in the 1987 OSTP report, titled *A Research and Development Strategy for High Performance Computing*, as fundamental problems in science or engineering with broad implications, whose solution would be facilitated by the application of high performance computing resources.

Table 3 indicates how many research directions in computational mechanics contribute to resolving the grand chal

Goals: Strategic Priorities

Extend U.S. technological leadership in high performance computing and computer communications.

Provide wide dissemination and application of the technologies both to speed the pace of innovation and to serve the national economy, national security, education, and the global environment.

Spur gains in U.S. productivity and industrial competitiveness by making high performance computing and networking technologies an integral part of the design and production process.

Strategy: Integrating Priorities

Support solutions to important scientific and technical challenges through a vigorous R&D effort.

Reduce the uncertainties to industry for R&D and use of this technology through increased cooperation between government, industry and universities and by the continued use of government and government-funded facilities as a prototype user for early commercial HPCC products.

Support the underlying research, network, and computational infrastructures on which U.S. high performance computing technology is based.

Support the U.S. human resource base to meet the needs of industry, universities, and government.

Program Components

High Performance Computing Systems

Research for Future Generations of Computing Systems
System Design Tools
Advanced Prototype Systems
Evaluation of Early Systems

Advanced Software Technology and Algorithms

Software Support for Grand Challenges
Software Components and Tools
Computational Techniques
High Performance Computing Research Centers

National Research and Education Network

Interagency Interim NREN
Gigabits Research and Development

Basic Research and Human Resources

Basic Research
Research Participation and Training
Infrastructure
Education, Training, and Curriculum

lenges of the HPCC initiative. It is clear that without significant advances in computational mechanics, successful resolutions of these challenges will be impossible.

Computational mechanics is closely related to several research initiatives of title U.S. government and the national needs identified in these initiatives. Appendix 6 describes how computational mechanics can be used to design new materials and to predict the life and failure of existing materials. However, significant progress in modeling is needed before this potential can come to fruition.

Following this section, a summary of research directions in computational mechanics is presented, in which the basic research areas that feed these applications are discussed. If the field is to progress over the next decade, these basic research areas require sustained attention and support.

CONCLUDING REMARKS

In its assessment of computational mechanics in the United States, the committee has identified the following broad areas that require focused research and financial support over the next decade:

1. *Adaptive methods for nonlinear problems*. Adaptive methods that allocate computational resources to the portions of a simulation where it is most needed are crucial to the types of computation on the leading edge of mechanics. Simulations of manufacturing processes, crashworthiness, and materials failure all require that adaptivity be feasible. Research is required in the fundamentals of error analysis and control for nonlinear problems, strategies for implementing adaptivity in mechanics problems, and treatment of various classes of mechanics problems by adaptive methods. Particular difficulties are foreseen in the implementation of adaptive strategies on massively parallel computers. Promising opportunities are also seen in the exploitation of artificial intelligence and knowledge-based systems in controlling adaptivity and other aspects of algorithm performance.
2. *Optimal control of computational modeling*. This is a new area of research that draws on the use of artificial intelligence and knowledge-based systems, methods of *a posteriori* error estimation, adaptive methods, and parallel computing to optimize and control the computational process, making it as efficient and reliable as possible. These component subjects are expected to be important areas of research for at least the next 10 years.

Table 3 Relevance of Computational Mechanics Major Technical Areas to the Grand Challenges of HPCC

	GRAND CHALLENGE PROBLEMS												
	WEATHER	MATERIALS	SEMICONDUCTORS	SUPERCONDUCTORS	VEHICLE SIGNATURE	TURBULENCE	VEHICLE DYNAMICS	NUCLEAR FUSION	COMBUSTION	ENHANCED OIL & GAS RECOVERY	COMPUTATIONAL OCEAN SCIENCES	VISION	ANTI-SUB. WARFARE
ADAPTIVE METHODS AND ERROR ESTIMATION	*	*	*	*	*	*	*	*	*	*	*		*
PROTOTYPE TEST SIMULATION	*	*	*	*	*	*	*	*	*	*	*	*	*
PARALLEL COMPUTATION	*	*	*	*	*	*		*	*	*	*		*
AI IN MATH MODELING	*	*	*	*	*	*	*	*	*	*	*	*	
ENIVRONMENTAL CONTAMINATION AND GFD SIM										*	*		
MATERIALS RESEARCH		*	*	*									*
STRUCTURES AND STRUCTURAL DYNAMICS			*										*
NONLINEAR EQUATIONS AND BIFURCATION	*	*	*	*	*	*	*	*	*	*	*		
UNCERTAINTY AND STOCHASTIC PROCESSES	*	*			*	*					*		*
CHEMICALLY REACTING FLOW AND COMBUSTION					*	*			*				
NUMERICAL MODELING OF TURBULENCE	*					*			*		*		*
GENERAL CFD	*		*			*	*		*	*	*		*
NONLINEAR DYNAMICS OF MULTIBODY SYSTEMS		*							*	*	*		
MANUFACTURING PROCESSES			*		*							*	*

3. *Structures and materials*. The computational modeling of new materials as well as the behavior and control of structures composed of such materials constitute one of the most important and exciting areas of research for the next decade. New advances in computer modeling techniques and a better understanding of the micromechanics of structural materials bring within the realm of possibility the actual design of new structures that exhibit special strength, fatigue life, stiffness, and fracture resistance. The role of computational mechanics modeling in these areas is an essential one; however, many basic problems must be resolved to achieve these goals.
4. *Computational fluid dynamics (CFD)*. This field, which includes traditional computational fluid dynamics, combustion, hypersonics, chemically reacting flows, environmental pollution and contamination modeling, weather prediction, direct numerical simulation of turbulence, and turbulence modeling, has suffered from a flawed philosophy advocated in the past by some agencies that the size and speed of available computers will drive the advances in CFD, making research on methodology of secondary importance. Recent developments prove that a contrary philosophy is needed for the healthy growth of CFD over the next decade. True advances in these subjects will result from interdisciplinary research in fluid dynamics, computer science and engineering, and numerical analysis in which all aspects of CFD modeling are addressed. In particular, many traditional approaches to CFD modeling are now obsolete. There is a need for considerable research that will develop effective techniques to overcome the extensive list of long-standing open problems.
5. *Reliability of computational modeling*. The issue of computational mechanics reliability will be of special concern during the next decade. A variety of research areas will impact this issue, including the development of theories for probabilistic methods, uncertainty and stochastic processes in mechanics, *a posteriori* error estimation, postprocessing, and knowledge-based systems. The use of hierarchical modeling concepts also emerges as an area that requires research to provide a basis for testing the models' reliability.
6. *Interdisciplinary parallel computing*. To achieve higher levels of computational power, such as computer processing speeds in the teraflop range, hardware designers are increasingly turning to massively parallel computers. The efficient utilization of these computers will require fundamentally new concepts in computational mechanics algorithms.

Research in parallel computing techniques in computational mechanics must be done by interdisciplinary teams with expertise in computational mechanics, numerical analysis, and

computer science. Interaction among computer scientists, engineers, numerical analysts, and hardware developers is essential to achieving advances. Parallel programming constructs targeted for computational mechanics must be developed, and new programming languages and software tools must be standardized. New languages must have capabilities rich enough to address readily the complexities commonly found in computational mechanics. The design and development of special hardware to optimize performance for particular mechanics computations should also be undertaken.

1

ADAPTIVE METHODS AND ERROR ESTIMATION

As a first step, virtually all modeling procedures in use in computational mechanics involve a process of discretization in which continuum models of nature are transformed into discrete or digital forms manageable by digital computers. This discretization process is critical to the overall computer simulation and governs its accuracy, efficiency, and general effectiveness.

Most discretization procedures involve the construction of a mesh or grid overlaid on the volume of matter to be studied; quantities of interest within gridcells or quantities defined at gridpoints are evaluated computationally. Other techniques model the so-called spectral content of continua with finite but high-order spectral representations. Collectively, these grid parameters characterize the discretization. Indeed, the accuracy with which the discrete model can represent the continuum usually depends on parameters such as the grid space (mesh size), density of gridpoints, and order of the representation.

It is clear that in this necessary first step of the computational process, the discretization, an error is always made, because a discrete model cannot possibly capture all of the information embodied in continuum models of gaseous, fluid, or solid materials. This inherent error has been a subject of concern and a topic of research for many decades and remains a source of many open questions—how can the error be measured, controlled, and effectively minimized? These issues will be among the most important topics of research in computational mechanics for the next decade and are at the heart of the reliability of computer simulations of nature. If the mathematical models of mechanics were perfect representations of mechanical events (which, of course, they are not), their utility in simulating events would be solely dependent on the discretization process used in computations and the errors it produces.

A POSTERIORI ERROR ESTIMATION

The subject concerned with measuring or estimating the discretization error that prevails in a computed simulation is called *a posteriori* error estimation. If it were possible to measure the discretization error with some level of mathematical precision, the reliability of the computed simulation could, at least partially, be assessed. Moreover, the entire discretization process could conceivably be modified or controlled to minimize and control the error.

The subject of *a posteriori* error estimation is in the early stages of development, although significant progress has been made in recent years for linear elliptic problems. Typically, finite element methods, finite difference methods, or boundary element methods are used in the discretization process, and the error is measured by error indicators that express the error in a single element or gridcell in an appropriate norm. Numerical experiments are generally performed on benchmark problems for which the exact error can be calculated, and the effectiveness of the error estimate can be assessed by computing effectively indices defined by

$$\xi = \frac{\text{estimated error}}{\text{true error}}$$

Today, most of the *a posteriori* error estimation theory pertains to linear elliptic problems and employs energy-norm estimates in some cases. Estimates for which 0.8 1.2 can be achieved. However for more general problems, particularly time-dependent or nonlinear problems, the theory is much less developed. There is also an urgent need to develop methods of error estimation in other norms since different classes of problems need different measures of error.

Nonlinear problems present special challenges in error estimation. The discretization process can introduce spurious solutions that have little bearing on the true behavior of the system under study. In addition, the theory of *a posteriori* error estimation for bifurcation problems, hyperbolic conservation laws, problems with multiple scales, and problems with resonance is in its infancy and deserves detailed study in the future. Work on methods of error estimation and control for the Navier-Stokes equations of fluid dynamics has mainly been a subject of ad hoc experimentation; no mathematical analysis of methods for this important class of problems exists.

Undoubtedly, this subject will be the focus of much research during the next decade.

Most existing methods of error estimation pertain to finite element methods of discretization. Additional work is needed on error estimation procedures for boundary element methods, finite difference methods, spectral and spectral element schemes, and other related techniques. For the short term, research will continue to focus on models characterized by linear elliptic equations, but research on the important areas of time-dependent and nonlinear problems will be vital to future developments in adaptive methods throughout the next decade.

Adaptive Methods

Again, if one can estimate even roughly the error induced in the discretization process, it is then possible to adjust the discretization parameters (e.g., mesh size, order of the approximation, density of gridpoints, even the solution algorithm) to control the computation (the error, stability, overall performance of methods used in the analysis of the model). Methods of discretization and numerical analysis that automatically adjust these parameters are called adaptive methods. Their goal is the optimal control of the computational process to produce the best results for the least effort. Differences in various adaptive schemes hinge on how this optimum is defined, how control parameters are selected and measured, and how the effort is defined. The control parameter is usually the discretization error in each gridcell. Some strategies exist in which numerical stability is also a factor, so that the time step, for example, is also controlled to maintain stability at minimum computational cost.

The best adaptive procedures function independently of the user, who merely prescribes a level of error that he/she can tolerate or a dollar value (cost) that he/she is willing to endure to complete a simulation. Thereafter, the adaptive code makes the decisions necessary to produce solutions within the user-specified limits. Once the control parameters (the gridcell, element errors, or some reasonable approximation of them) are available, the adaptive code attempts to adjust them to meet control objectives—to minimize the error. Typically, error control is achieved by refining the mesh in areas of the solution domain where errors are too large and coarsening the mesh (using larger elements) where the error is small, or relocating nodes to increase nodal densities near regions of high error. Also, one could increase the order of

approximation and expect the accuracy of the approximation to be increased. Thus, in adaptive finite element methods, several broad types of adaptivity can be used in the control process:

1. *h-methods*: The mesh size h is used to control error. Error is reduced by refining the mesh or regenerating a new finer mesh. To obtain an optimal h-mesh (one with the least error possible for a fixed number of refinements), provisions for also coarsening a mesh must be included in the adaptive strategy.
2. *r-methods*: A fixed number of elements of a given order are used in a mesh, but nodes are relocated to reduce error in certain regions. The *r*-methods are thus moving grid methods in which the number of degrees of freedom is fixed, but the node locations are adapted to control error.
3. *p-methods*: The spectral order p of the approximation is raised or lowered to control error. In finite element methods or boundary element methods, the order p corresponds to the degree of the polynomial shape function used over an element.
4. *Combined methods*: Combinations of *h*-, *r*-, and/or *p*-adaptivity are used.

Figures 1.1a and 1.1b show a mesh modeling supersonic flow around a space shuttle in which *h*-method adaptivity has been employed to optimize the mesh structure to produce accurate simulation of flow features important in assessing the performance of the design—such as the profiles of pressure distribution shown.

The potential advantages of adaptive methods over conventional methods of computational mechanics are enormous. For example, for many typical problems in two dimensions uniform mesh and degree distribution would lead to 100,000 or more equations, whereas adaptive approaches involving only 700 to 1,000 equations give results with comparable accuracy. Theoretically, in many cases adaptive procedures can lead to exponential convergence rates compared to algebraic convergence of classical low-order nonadaptive approaches. The potential impact on computational mechanics of exponentially convergent computational schemes may be one of the most important factors affecting research in this area for the next decade. Figure 1.2 illustrates the performance of exponentially convergent finite difference, finite

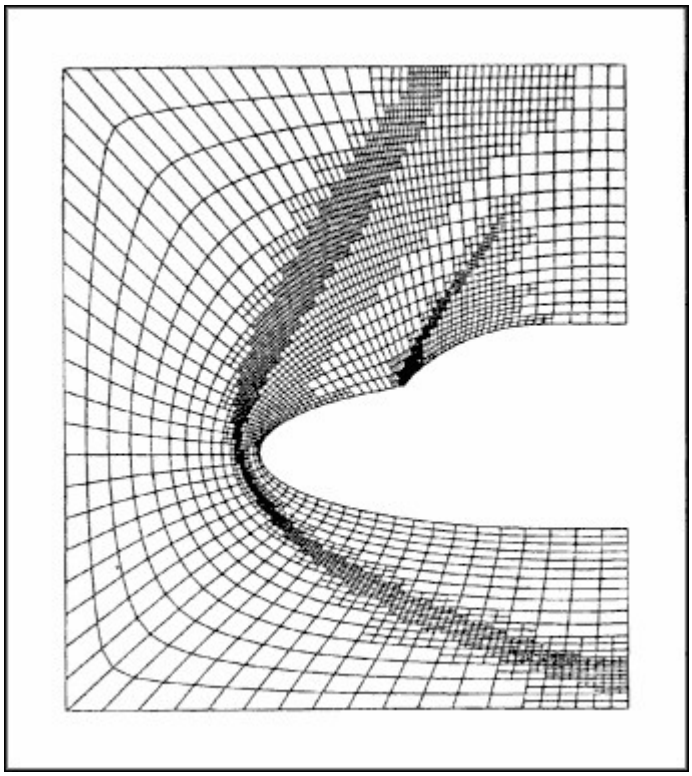

Figure 1.1a
An h-adapted finite element mesh about a shuttle-like body. Note that the adaptive algorithm has automatically refined the mesh to capture shocks in the flow field.

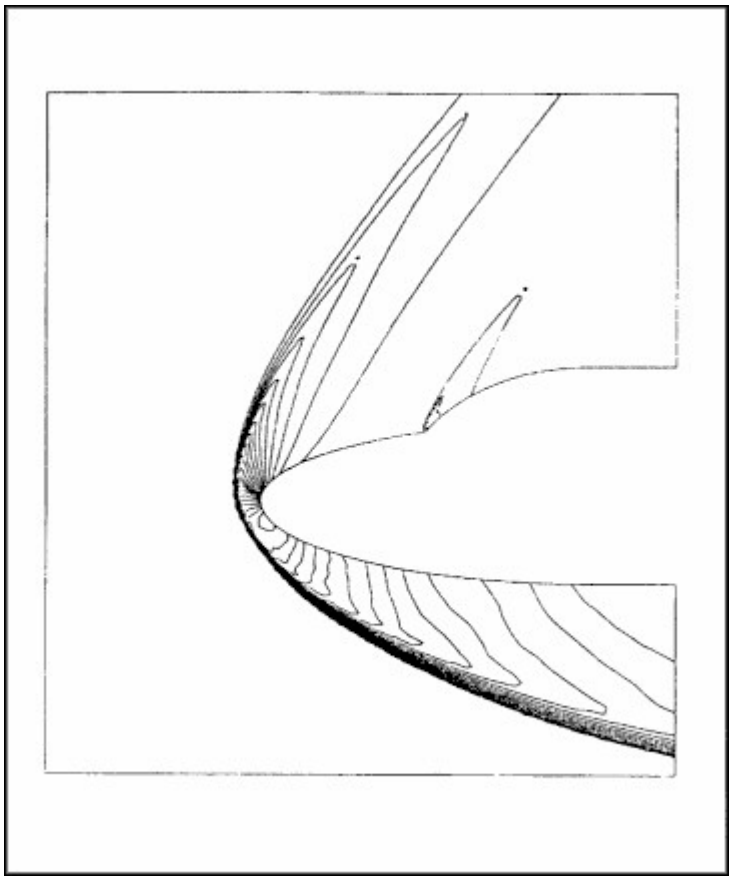

Figure 1.1b
Computed pressure contours on the optimally refined mesh.

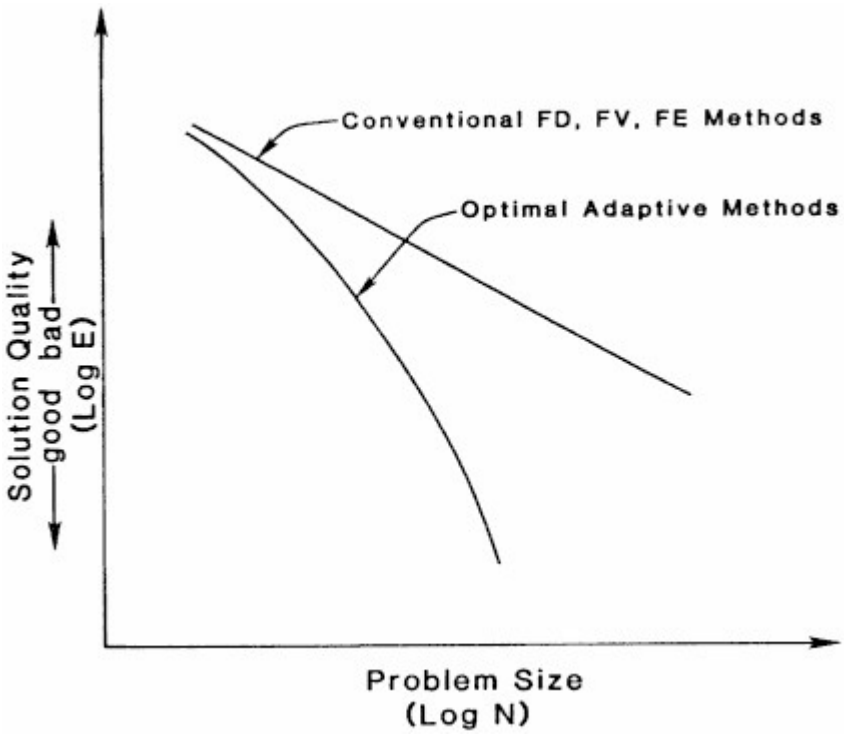

Figure 1.2
Numerical performance of various classes of numerical methods: logarithm of error versus logarithm of the number N of unknowns.

volume, or finite element schemes. The significance of the curves in this figure is that no matter what the capacity of the computer on hand to solve large-scale problems in computational mechanics, a level of accuracy can always be specified that cannot be attained by conventional methodologies. Such specified accuracies may be attained using superalgebraic convergent schemes (e.g., adaptive methods) since these can deliver results of a specified accuracy with many orders of magnitude fewer unknowns than conventional methods.

To date, certain *p*-version adaptive finite element methods, combined *hp* adaptive finite element methods, and spectral, pseudospectral, and spectral element methods have attained exponential convergence rates. Further research on adaptive strategies that produce optimal distributions of discretization parameters and exponential rates of convergence is urgently needed and should be a principal topic of research during the next decade.

Several other research issues on adaptive methods require additional study. Adaptive procedures should reflect various special aims of the computation, such as stress intensity, velocity, contact pressure, and vorticity. There is need for flexible adaptive principles that will focus on these specific quantities. Very little has been done in this direction. In time-dependent problems, successful approaches based on *r*-methods (relocating mesh points in time) combined with *h*-methods (refinement and coarsening) have been developed recently. Further research on these types of discretization is needed. In finite difference methods, some success has been achieved with moving overlapping meshes for hyperbolic problems and other types of time-dependent problems. In nonlinear solid mechanics, much remains to be done in the development of effective adaptive schemes. be enriched by the appearance of adaptive methods.

The successful implementation of adaptive methods often leads to nonsparse systems of equations that are not readily handled by conventional linear equation solvers. Research on the development of special techniques for solving the algebraic equations arising from adaptive methods is needed. These may include domain decomposition methods for multiprocessor computation, multigrid techniques, preconditioned iterative techniques, and related methods.

Postprocessing. A topic somewhat related to error estimation is referred to as (mathematical) postprocessing of numerical solutions to produce results for enhanced accuracy and utility. Postprocessing techniques extract extra information present in computed results to obtain even better simulations. For example, these schemes may exploit supercon

vergence properties of finite element methods or employ extrapolation techniques that use data from sequences of meshes obtained in adaptive processes, or they may use so-called extraction schemes that use Green's formulas to "extract" superaccurate properties of solutions at points in the domain where high accuracy is sought. The basic goal is to extract maximum information and precision from computed simulations. To date, the subject has been developed only for model linear elliptic problems. Much research is needed to extend these postprocessing ideas to more general classes of problems.

In summary, the major research directions in the general areas of adaptive methods and error estimation are to:

- develop reliable *a posteriori* error estimation of computed data of interest with the effectivity index of order $0.9 \leq \quad \leq 1.1$.
- develop *a posteriori* error estimation techniques for nonlinear problems in solid mechanics;
- develop *a posteriori* error estimation techniques in fluid mechanics for both compressible and incompressible flow;
- explore and develop *a posteriori* estimates and adaptive methods for space-time approximations; determine optimal techniques for *h*-, *p*-, and *r*-adaptive schemes;
- develop a modeling reliability theory, including possible expert system development and bracketing theory;
- devise postprocessing techniques for enhancement of solution accuracy; and
- develop adaptive modeling in which criteria for changes of models are incorporated in the adaptive process.

2

PROTOTYPE TEST SIMULATION

Prototype test simulation by computers is rapidly becoming an important part of the design process in many industries. For example, in the automotive industry crashworthiness design is increasingly performed by computer simulation rather than physical prototypes. Computer simulation of automotive crashes offers the promise of both reduced cost (a crash test of a prototype costs $300,000 to $700,000 versus about $100,000 for a simulation) and reduced design time. Similar trends are also evident in other industries, for example, the manufacture of glass, tires, and jet engines (to prevent catastrophic failure during bird impact).

Simulation also is being used in structure designs where, historically, the ultimate capacity has been "designed in" through empirical safety factors. For example, in the seismic design of dams, liquid-storage tanks, and nuclear power plants, nonlinear simulation is increasingly used to examine the ultimate capacity and the consequences of partial failures.

In addition, the entire area of weapons effects in the defense community relies significantly on nonlinear simulation. For example, determination of the ultimate capacity of protective structures is often performed by a combination of scale-model tests and computer simulation. Computer simulation has also had a dramatic impact on the design of armor and armor-piercing weapons.

FINITE ELEMENT METHODOLOGY

The major catalyst for this rapid evolution has been the development of efficient finite element methods for nonlinear mechanics. The past two decades have seen the emergence of powerful nonlinear finite element algorithms and rapid advances in the fundamental understanding of nonlinear mechanics and material behavior. In the past the use of fully integrated finite element methods proved to be too expensive for practical computation. On the other hand, underintegrated elements were unstable and produced meaningless results. Further research, however, led to theories of numerical control

for stabilization methods and made possible significant improvements in the efficiency of nonlinear computations.

These new methods have been critical in the adoption of nonlinear techniques in crashworthiness analysis. The models required are very large and complex and involve many difficult phenomena such as contact-impact friction, local buckling, and strain-rate effects. A typical model is shown in Figure 2.1. Such a simulation still requires 20 to 80 hours of supercomputer time; without these techniques, the computer time required for a simulation would be prohibitive.

These advances in nonlinear techniques have been critically dependent on the advances in continuum mechanics made in the 1950s and 1960s and their adaptation to a form that facilitates rapid computation without violation of basic physical principles. Today, nonlinear computer programs increasingly make use of frame-invariant formulations, which are applicable to arbitrarily large strains, and advances in the theory of plasticity to describe material behavior in the nonlinear range. However, these procedures have been streamlined by algorithmic developments such as the so-called radial return form for plasticity and domain integration for fracture.

Despite these important advances, in many respects the field is in its infancy and the many simulations that engineers want to perform routinely cannot be handled reliably today. Among the major handicaps is the inability to treat strength degradation that occurs near a structure's failure point.

In this regime the governing equations often change type from hyperbolic to elliptic and vice versa. Without scientifically sound methodologies for treating such behavior, reliable calculations are impossible for many materials due to the long computer times required for many simulations to achieve even a modicum of accuracy. It is evident that bringing the simulation times to the level needed to make them truly useful in the engineering design process requires adaptive methods and implementations on parallel computers.

NEW TECHNOLOGIES

Adaptive methods are essential for streamlining the modeling process. Today, separate finite element models are made for both linear analysis used for normal performance and nonlinear simulation. This duplication often requires 3 to 6 man-months of effort and takes months in the design process. Adaptive methods will enable the same models to be used. Then adaptivity will provide the mechanism for allocating elements where they are needed in a particular nonlinear simulation.

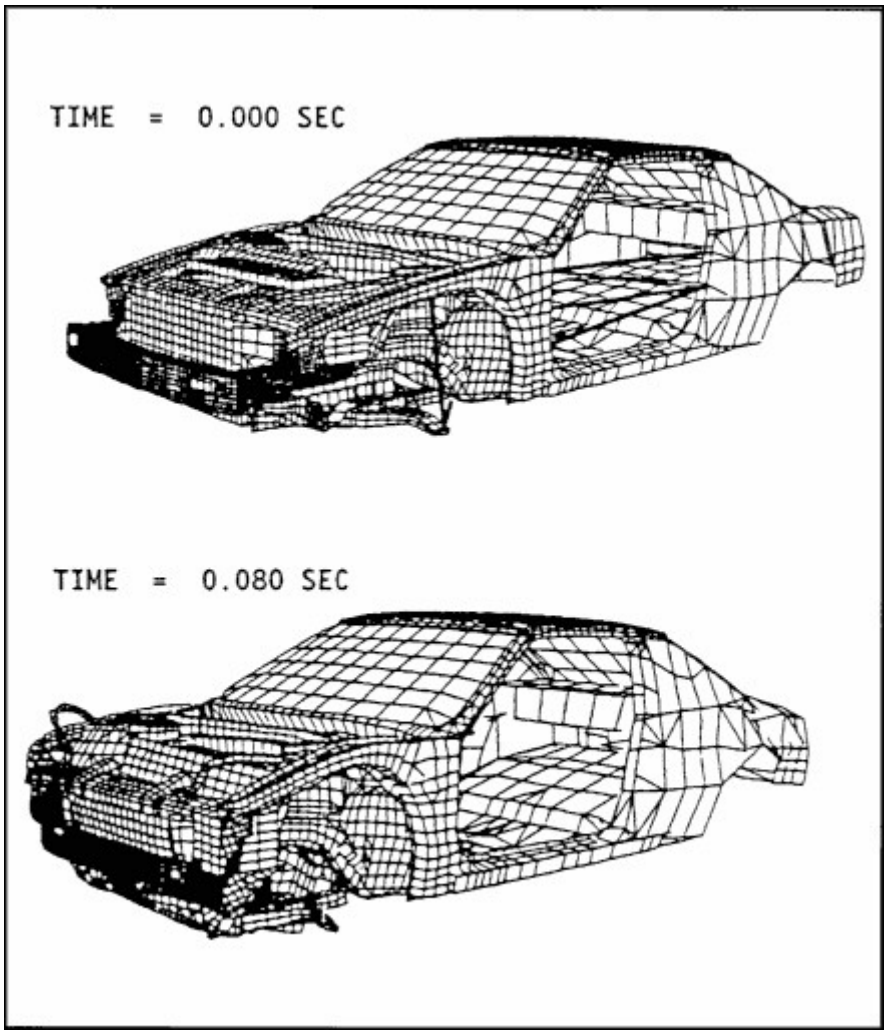

Figure 2.1
A finite clement model for crashworthiness Source: Johnson, 1989, in *Crashworthiness and Occupant Protection in Transportation Systems*, *AMD-Vol. 16, ASME, ed. by T.B. Khalil and A.I. King, pp. 27–37.*

The treatment of failure will require methods that can effectively handle multiple scales and coupled processes. There are 3 important classes of failure: (1) buckling, (2) fracture, and (3) failure due to strain localization. The last two, in particular, usually involve scales far smaller than that of the test objects. Although adaptive methods offer substantial promise, current adaptive methods are not sufficiently capable of treating these problems. Even in buckling, the latter stages often exhibit severe localization of deformation due to plastic hinge-line formation.

These processes are often coupled to heat transfer and even the kinetics of phase change. This coupling adds an additional dimension of complexity, for it often involves different time scales. All of these difficulties must be resolved effectively to make computer prototype testing a practical tool.

The efficient implementation of nonlinear mechanics algorithms to new computer architectures will require careful coordination between the mechanics modeling processes and the computer architecture. Past experience has shown that compilers simply cannot extract even a small percentage of the computer's capability with an algorithm designed for a von Neumann computer. A careful redesign of the algorithm based on an analysis of how it can be rearranged cognizant of the mechanics to be simulated is crucial to the development of effective algorithms and software.

The payoffs in achieving these goals will be substantial. If computer simulation of prototype testing can be integrated into the design process, lead times for new designs can be reduced, and more effective and economical designs can be developed. For example, in automobiles, actuator sensors must be placed so that they can detect impact with an object, such as from a full-scale frontal crash. This is a trial-and-error process that benefits from feedback. If feedback is provided by prototype tests, the process is slow and new ideas require considerable confirmation time. Development of a faster simulation method can dramatically reduce design time.

Similarly, in weapons-effects work the development of methods that can simulate near-failure regimes will provide more reliable estimates of weapons and targets in such regimes. Many products that are now routinely tested for ultimate capacity and redesigned several times before introduction to the market can rely on prototype test simulation for much of the design process. The insight provided by computer simulation, which can provide a much more complete picture of the product's behavior, and the reduced design times will contribute immensely to U.S. competitiveness in many industries.

3

PARALLEL COMPUTATION

COMPUTATION AND SUPERCOMPUTERS

Computation has played a central and critical role in mechanics for more than 30 years. Current supercomputers are being used to address challenging problems throughout the field. The Raveche Report[1] covers a broad range of disciplines. Nevertheless, it cites specific accomplishments where supercomputers have had profound impact in mechanics, such as the two-orders-of-magnitude reduction in the number of wind tunnel models tested, design of artificial heart valves with minimal degradation, and design and manufacture of efficient mechanical devices with respect to heat and mass transfer. Other problems described as being "ripe for impact" include full aircraft flow analysis, turbulent flow simulation, composite materials design, understanding phase transitions, polymeric materials design, and analysis of multiphase and particulate flows. For instance, solution of the Euler equations of compressible flow about a complete aircraft would greatly enhance vehicle design capabilities and dramatically reduce the time from design conception to fabrication and testing. Dependence on complex, time-consuming, and expensive wind tunnel testing would be greatly reduced. Turbulent flow simulations would lead to a fundamental physical understanding of the mechanisms involved in turbulence. This, in turn, would lead to more realistic turbulence models that could enhance vehicle stability, control, and performance.

Solutions to these problems could be obtained in approximately 1 hour on a computer having 10G Flop (1010 floating point operations per second) performance and 100 million

[1] H.J. Raveche, D.H. Lawrie, and A.M. Despain, 1987, *A National Computing Initiative: The Agenda for Leadership*, Report on the Panel on Research Issues in Large-Scale Computational Science and Engineering, Society for Industrial and Applied Mathematics, Philadelphia.

words of memory, which places it barely out of the range of current supercomputer capabilities. Algorithmic advances and hardware innovation notwithstanding, a greater reliance on parallelism will be needed to solve many of the complex problems addressed in this report.

Parallel computation involves a program's division into tasks that can be executed concurrently by several processing units. A task can be a single instruction, a group of instructions such as a loop, or a procedure. While theoretical models of parallel computation have focused on a hypothetical computer known as a parallel random access machine (PRAM), available hardware is quite varied. For present purposes, hardware will be classified as either a shared- or a distributed-memory computer; however, systems are becoming increasingly complex and such a simple division is not appropriate in many situations.

Shared-memory systems involve several processors connected to a common memory by a high-speed bus, for example. In principle, each processor has equal access to any data element; however, with current technology, communication delays develop when too many processors are involved. Hierarchical memory systems where each processor has a local memory cache have improved performances somewhat, but it is believed that architectures of this type are limited to tens or hundreds, rather than thousands, of processors.

Current supercomputers employ a limited form of parallelism based on a shared memory. Processors generally have vector capabilities, thus utilizing a mixture of coarse-and fine-grained parallelism. So-called graphics supercomputers that offer computational performance with superior visual display capabilities in a workstation are also of shared-memory design.

Systems having a more massive level of parallelism have had memory distributed among each of several processors. Each processor and its local memory are connected through a network to a limited number of other processing elements. For present purposes, a *processing element* shall refer to a processor with its local memory and input/output connections to the network. Various network topologies are available; the most popular include two-dimensional grids and hypercubes. Processing elements in a grid are located at vertices and have network connections to their nearest north, east, west, and south neighbors. With a hypercube-connected network, 2^d processing element located at the vertices of a d-dimensional cube are connected along the d edges emanating from each vertex. Thinking Machines Corporation's Connection Machine, for example, has a 512-megabyte memory distributed uni

formly among 65,536 one-bit processors. Each 16 processors form the vertices of a 12-dimensional hypercube, and pairs of 16-processor clusters share a floating-point unit.

Processors on the Connection Machine operate synchronously in lock step or single instruction, multiple data (SIMD) fashion with a separate control unit dictating the instruction to be executed simultaneously on all processors. MasPar Corporation's recently introduced MP-1 is another SIMD machine having 16,384 four-bit processing elements, each having network connections to its nearest eight neighbors. Hypercube-connected asynchronous or multiple instruction, multiple data (MIMD) systems with as many as 10 dimensions (1,024 processors) have been available for some time. With each processor executing its own program, balancing computation and communication is critical. In fact, load balancing is the major programming problem on MIMD computers.

Past and likely future advances in processor speed and memory are traced in Figure 3.1. This forecast, which appears to be reasonable given straight forward developments of semiconductor devices, suggests that Teraflop (10^{12} floating point operations per second) performance will be available by the mid-1990s.

Additional architectural advances will provide platforms for greater experimentation and vastly increased performance. Massively parallel MIMD systems have been described recently in the literature. Systems placing a greater reliance on hierarchies of parallelism will emerge. For example, a likely compromise between the flexibility of an MIMD system and HP LaserJet Series IIHPLASEII.PRSy of processors connected to a global memory with each processor being a medium-grained network of fine-grained vector processors. Software management of such complex systems will be a major challenge of the 1990s.

Originally poor on high-performance systems, software environments have been improving steadily. Parallel versions of Lisp, C++, and Fortran-77, for example, are available on the Connection Machine. Most systems offer at least one high-level language having some parallel programming enhancements and some parallel debugging tools. Research on new parallel computing languages is continuing; however, software is still the greatest bottleneck in parallel computation. A lack of standards, robust and rich programming structures, software development tools, and debugging aids has discouraged widespread use. Left alone, this situation will become even more critical as increasingly complex hardware systems emerge.

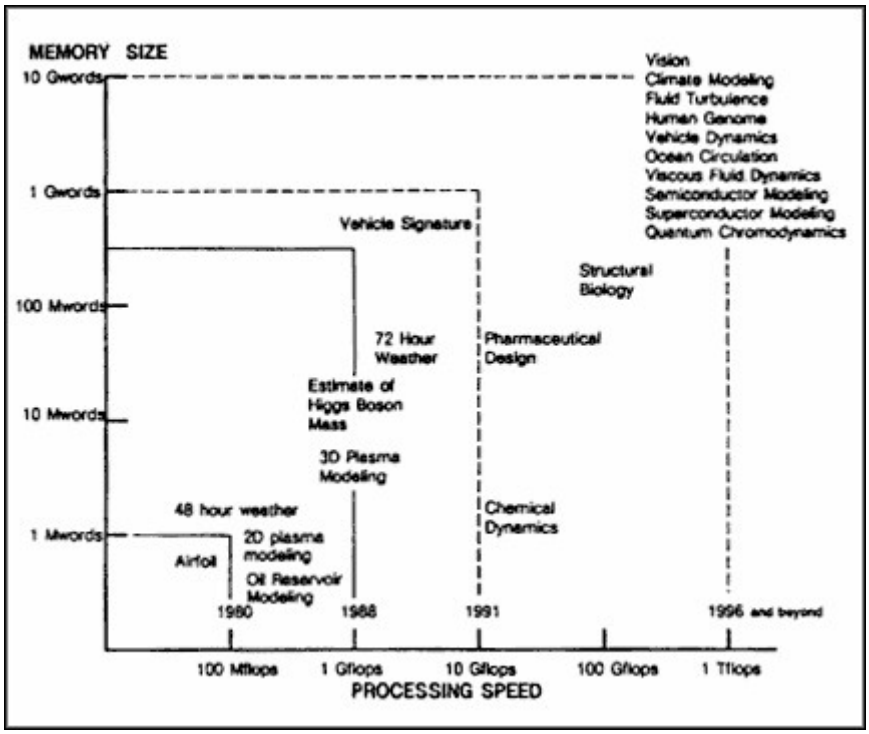

Figure 3.1
Some grand challenges of high-performance computing and their projected computational requirements (from the Federal High-Performance Computing Program).[2]

[2] T.M. Walker, ed., (1989).

The greatest advances in computing performance have, and will continue to come from, algorithmic innovation. The fast Fourier transform, adaptive *hp*-refinement, and fast multipole method are but three examples. Systems and pressor software must improve so scientists and engineers do not have to divert a great deal of time from their primary disciplinary goals in order to understand the nuances of parallel programming. With this accomplished, algorithmic breakthroughs will occur more frequently.

THE FEDERAL HIGH-PERFORMANCE COMPUTING PROGRAM

Several studies conducted during the 1980s, terminating with the Raveche Report and the FCCSET (Federal Coordinating Council for Science, Engineering, and Technology) Report,[3] addressed the need for the United States to maintain superiority in high-performance computing. As a result, the Bush administration introduced the Federal High-Performance Computing (HPC) Program in 1989.[4] The HPC Program calls for increasing the current $500 million funding level for high-performance computing by $150 million during the first year to $600 million by year five. Intimately connected with parallel computation, the HPC Program clearly addresses several issues central to computational mechanics. Listed among their "grand challenges" are climatology, turbulence, combustion, vehicle design, oil recovery, oceanography, and viscous flow. Progress in these areas, made possible by advanced parallel computing systems and software, will result in enhanced productivity, greater efficiency of engineering systems, and increased scientific knowledge.

Research Opportunities

The greatest success in algorithm development for parallel computers has been the solution of linear algebraic systems. Techniques and software exist for solving dense and

[3] P.G. Huray, ed., 1987, A Research and Development Strategy for High Performance Computing, Executive Office of the President, Office of Science and Technology Policy.

[4] T.M. Walker, 1990, "High-Performance Computing Update," Computing Research News, Vol. 2, No. 2.

sparse systems on a variety of architectures. Relative to this, parallelization of other portions of computational mechanics problems is in its infancy.

Experimentation with procedures for solving solid and fluid mechanics problems on massively parallel SIMD arrays has had some degree of success. For problems of moderate size, one could associate an element, computational cell, or node of the mesh with an individual processor of the parallel system. With computing systems having the concept of virtual processors, this strategy would be viable for problems of arbitrary size. However, as previously noted, the network connections of the computer may not be rich enough to support a convenient mapping of elements or nodes onto processors. Problems using uniform meshes and finite difference approximations are, naturally, simpler to map to SIMD or MIMD networks than are formulations using unstructured grids having general connections. One technique for solving unstructured-grid finite element problems on a SIMD array uses a two-step procedure with an initial mapping of elements of the mesh onto processors followed by a second mapping of the nodes onto processors for the solution of the linear system. Another approach utilizes an underlying uniform tree structure associated with the grid to create a mapping. Nevertheless, there is a great deal more work to be done, and mapping the discrete problem to the computational topology continues as a major unsolved problem. This situation will be further exacerbated by the emergence of new architectures with different network connections.

Adaptive finite difference and finite element methods employing automatic mesh (*h*-type) refinement, order (*p*-type) enrichment, and/or mesh redistribution (*r*-type refinement) are among the most efficient and robust procedures for solving mechanics problems on serial computers. Optimal adaptive enrichment strategies can even produce exponential convergence rates as problem sizes increase (cf. Figure 1.2). Very little has been done, however, to develop parallel adaptive procedures. Factors limiting progress include reliance on nonuniform unstructured meshes, variable-order methods, and dynamic data structures such as trees. The efficiency afforded by adaptivity will always be important at some level and would permit solving problems that could not be solved with uniform structures. Furthermore, complex three-dimensional geometries will place a greater emphasis on adaptive mesh refinement and order enrichment rather than on *a priori* mesh generation and order specification.

Adaptive procedures typically use hierarchical data structures that contain suitable parallel constructs. Procedures

for mapping such trees to some parallel architectures are known, for example, and used in computer vision and computer-aided design. Use of these procedures with adaptive techniques has begun and, while much more remains to be done, there is some cause for optimism. Rapidly convergent *hp*-refinement procedures use different combinations of mesh refinement and order enrichment locally and seem more suited to MIMD computation than to execution on a computer operating in lock-step fashion. Many mesh redistribution techniques often include global influences and also seem to be more appropriate for MIMD computation. Envisioned MIMD computer systems containing clusters of SIMD arrays may overcome some of these difficulties, but efficient processor utilization and load balancing remain challenging issues to the successful parallel implementation of adaptive *hp*- and *r*-refinement procedures.

A 10-dimensional hypercube-connected MIMD computer has approximately the same bit-processing capability as the largest SIMD arrays. Software environments on MIMD computers, however, are less developed than those on SIMD computers due to the greater complexity involved in distributing control structures. Model mechanics problems involving wave propagation, compressible flow, and elastic deformation have been solved with great success; but realistic problems with embedded heterogeneities such as heterogeneous materials, phase transitions, and elastic-plastic responses present difficulties similar to those described for parallel adaptive procedures. Computation on SIMD systems will be inefficient due to large variations in local response, and computation on MIMD systems will require complicated load balancing.

Solution-based reliability measures provided by *a posteriori* error estimation could potentially serve as metrics for balancing processor loading. Physically motivated domain decomposition, whereby asymptotic or other approximate analyses are used to estimate the computational effort needed in various portions of the problem domain, are another possible means of load balancing that should be explored further.

Education, Experimentation, and Access

Researchers in computational mechanics must be educated and have access to promising advanced systems as soon as they become available. Undue reliance should not be placed on existing systems, and conclusions regarding the superiority of one architectural design relative to another should not be

drawn too quickly. Software developed with a lack of standardization and a lack of input from the engineering community can only hinder progress and further discourage the use of parallel systems. Interdisciplinary cooperation is clearly needed. Parallel programming constructs and languages must strive for standardization and portability. New computing languages must have capabilities rich enough to address complexities commonly found in computational mechanics.

Access to supercomputers has been improved considerably by the construction of national and regional networks and the funding of computational research sponsored by the National Science Foundation supercomputing centers. However, relatively few engineers have access to more massive and experimental parallel computing systems. Current massively parallel systems are the prototypes of future computers, and their use is now vital toward achieving the algorithmic breakthroughs that will provide necessary performance improvements. Although less expensive than production supercomputers, massively parallel systems are still costly and, as noted, have a more complex programming environment. Use should increase with the availability of smaller hardware systems at reduced cost and with expected improvements in software systems and tools. Scaled-down SIMD systems still provide hundreds to thousands of processors and thus have a reasonable degree of parallelism. Cost is declining to the point where individual research groups can afford to purchase systems.

Computation is peripherally involved in most mechanics curricula and is a central part of others; however, parallel programming in education is rare. Courses emphasizing parallel programming design and methodology must be developed now in order to meet the engineering challenges of the future. Programs should emphasize the interdisciplinary nature of the subject and include studies in mechanics, computer science, and numerical analysis. Appropriate computer science courses should give students a knowledge of parallel algorithm design, programming languages, data bases, operating systems, architectures, and algorithm analysis. Computational experimentation is essential. Hardware must either be available locally or conveniently accessible via networks. Software must be developed to illustrate the fundamental concepts and encourage algorithmic experimentation without undue reliance on frustrating semantic details or architectural idiosyncrasies. Programs will undoubtedly begin at the graduate level; however, undergraduate exposure to parallel computation should not be delayed.

Much of the future of high-performance computing lies in massive parallelism and, although present parallel systems

may bear little resemblance to their successors, experience gained now is vital to developing new algorithms and solution strategies, building a knowledge base, and developing and designing new architectures. Steps must be taken to provide increased access to parallel systems of moderate and massive proportions. Moderately sized systems having modest cost may be distributed widely. They are important for education, experimentation, software development, and establishment of a local cadre of expertise. They will accelerate learning of the more complicated programming methodology and greatly stimulate interdisciplinary interaction. Large systems at remote centers may be used for production and testing of scalability or parallel algorithms.

4

ARTIFICIAL INTELLIGENCE IN MATHEMATICAL MODELING

ARTIFICIAL INTELLIGENCE AND KNOWLEDGE-BASED SYSTEMS

Artificial intelligence (AI) is as much a branch of computer science as are its other branches, which include numerical methods, language theory, programming systems, and hardware systems. While computational mechanics has benefited from, and closely interacted with, the latter branches of computer science, the interaction between computational mechanics and AI is still in its infancy. Artificial intelligence encompasses several distinct areas of research each with its own specific interests, research techniques, and terminology. These sub-areas include search technologies, knowledge representation, vision, natural language processing, robotics, machine learning, and others.

A host of ideas and techniques from AI have the potential to impact the practice of mathematical modeling. In particular, knowledge-based systems and environments can provide representations and associated problem-solving methods that can be used to encode domain knowledge and domain-specific strategies for a variety of ill-structured problems in model generation and result interpretation. Advanced AI programming languages and methodologies can provide high-level mechanisms for implementing numerical models and solutions, resulting in cleaner, easier to write, and more adaptable computational mechanics codes. A variety of algorithms for heuristic search, planning, and geometric reasoning can provide effective and rigorous mechanisms for addressing problems such as shape description and transformation, and constraint-based model representation. Before discussing the applications of AI in mathematical modeling, we briefly review knowledge-based expert systems and problem-solving techniques.

Knowledge-Based Expert Systems

A good standard definition of knowledge-based expert systems (KBES) is the following: knowledge-based expert systems are interactive computer programs incorporating judgment, experience, rules of thumb, intuition, and other expertise to provide knowledgeable advice about a variety of tasks.

Many computer-aided engineering professionals initially react to this definition with boredom and impatience. After all, conventional computer programs for engineering applications have become increasingly interactive. They have always incorporated expertise in the form of limitations, assumptions, and approximations, as discussed above, and their output has long ago been accepted as advice, not as "the answer" to a problem.

There is a need, therefore, to add an operational definition to distinguish the new wave of KBES from conventional algorithmic programs that incorporate substantial amounts of heuristics about a particular application area. The distinction should not be based on implementation languages or on the absolute separation between domain-dependent knowledge and generic inference engine. The principal distinction lies in the use of knowledge. A traditional algorithmic application is organized into two parts: data and program. An expert system separates the program into an explicit knowledge base describing the problem-solving knowledge and a control program or inference engine that manipulates the knowledge base. The data portion or context describes the problem and the current state of the solution process. Such an approach is denoted as knowledge based.

Knowledge-based systems, as a distinctly separate AI research area, are about a decade old. This decade of research has seen many changes in the importance placed on various elements of the methodology. The most characteristic change is methodological; the focus has shifted from application areas and implementation tools to architectures and unifying principles underlying a variety of problem-solving tasks.

In the early days of knowledge-based systems, the presentation and analysis were at two levels: 1) the primitive representation mechanisms (rules, frames, etc.) and their associated primitive inferencing mechanisms (forward and backward chaining, inheritance, demon firing, etc.), and 2) the problem description. Unfortunately, it turned out that the former descriptions are too low level and do not describe the kind of problem that is being solved, while the latter descriptions are necessarily domain specific and often incomprehen

sible and uninteresting for people outside the specific area of expertise.

A description level is needed that adequately describes what heuristic programs do and know—a computational characterization of their competence that is independent of both task domain and programming language implementation. Several characterizations of generic tasks that arise in a multitude of domains have been presented recently in the literature. Generic tasks are described by the kind of knowledge on which they rely and their control of problem solving. Generic tasks constitute higher-level building blocks for expert systems design. Their characterizations form the basis for analyzing the contents of a knowledge base (for completeness, consistency, etc.) to describe the operation and limitations of systems and to build specialized knowledge acquisition tools.

Problem Solving

Many problem-solving tasks can be formulated as a search in a state space. A state space consists of all the states of the domain and a set of operators that change one state into another. The states can best be thought of as nodes in a connected graph and the operators as edges. Certain nodes are designated as goal nodes, and a problem is said to be solved when a path from an initial state to a goal state has been found. State spaces can get very large, and various search methods to control the search efficiency are appropriate.

1. Search reduction. This technique involves showing that problem's answer cannot depend on searching a certain node. There are several reasons this could be true: (a) No solution can be in the subtree of this node. This technique has been called "constraint satisfaction" and involves noting that the conditions that can be attained in the subtree below a node are insufficient to produce some minimum requirement for a solution. (b) A solution in another path is superior to any possible solution in the subtree below this node. (c) The node has already been examined elsewhere in the search. This is the familiar dynamic programming technique in operations research.
2. Problem reduction. This technique involves transforming the problem space to make searching easier. Examples of problem reduction include: (a) planning with macro operators in an abstract space before getting down to the details of actual operators; (b) means-end analysis, which attempts to reason backward from a known goal; and (c) sub-goaling,

which breaks down difficult goals into simpler ones until easily solved ones are reached.

3. Adaptive search techniques. These techniques use evaluation functions to expand the "next best" node. Some algorithms (A*) will expand the node most likely to contain the optimal solution. Others (B*) will expand the node that is most likely to contribute the most information to the solution process.
4. Using domain knowledge. One way to control the search is to add additional information to nongoal nodes. This information could take the form of a distance from a hypothetical goal, operators that may be usefully applied to it, possible backtracking locations, similarity to other nodes that could be used to prune the search, or some general goodness in formation.

APPLICATIONS IN MATHEMATICAL MODELING

Mathematical modeling is the activity devoted to the study of the simulation of physical phenomena by computational processes. The goal of the simulation is to predict the behavior of some artifact within its environment. Mathematical modeling subsumes a number of activities, as illustrated by Figure 4.1.

The following sections discuss the applications and potential impacts of AI technology on various mathematical modeling activities. The mathematical modeling activities presented include model generation, interpretation of numerical results, and development and control of numerical algorithms. Note that these activities are not independent, and this organization is used primarily to assist in the exposition of ideas.

Model Generation

The term model generation is used to encompass all activities that result in the generation of models of physical systems suitable as input for a computational mechanics program. The generation of mathematical models from physical descriptions of systems is a problem of great practical importance. In all disciplines that use computational mechanics—aerospace, nuclear, marine, civil, and mechanics—there is a need to model an increasingly wider range of phenomena in all stages of system design from the earliest conceptual studies to the most detailed component performance evaluation. In addition, there is an urgent need for much

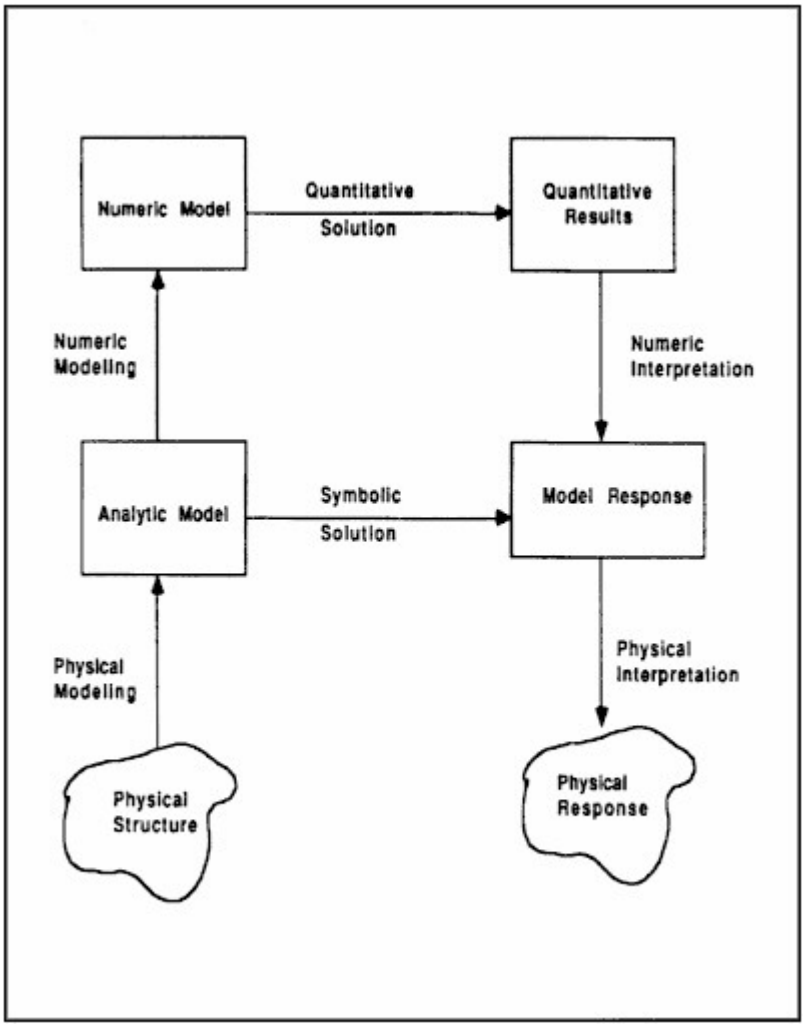

Figure 4.1
Mathematical modeling process.

closer integration of computational mechanics evaluations into computer-aided design (CAD) and to extend analyses to computer-aided manufacturing where there is a great interest in analyzing not just the finished components, structures, or systems but also the manufacturing processes themselves, such as casting, forging, or extrusion.

With the availability of literally hundreds of computational mechanics codes, including a large number of general-purpose finite element programs with a broad range of capabilities, model generation has become the primary activity of the analyst. However, in the current state of the art, the preparation of input data is a tedious, error-prone, and time-consuming process. Analysts are forced to interact with programs at a level much lower than the conceptual level at which they make decisions. Hence, there is a need for higher-level program interfaces that will free analysts from details, allowing them to generate models in terms of high-level behavioral descriptions, thereby increasing their productivity and improving the quality and reliability of their analyses.

Moreover, because of the very small number of experienced modelers who can confidently and reliably model physical problems and the increasing need for modeling expertise, it has also become increasingly important to capture and organize the modeling knowledge of experienced analysts and transfer it to the less experienced novice analysts. The methodology of AI and knowledge-based systems promises to provide an opportunity to respond to the needs identified above.

Model generation tools can be discussed at three levels of increasing abstraction.

1. Intelligent help systems. Intelligent help systems address the issue of providing consulting advice to nonexpert engineers. The subject of help could be either how to use a particular analysis program or what model parameters and procedures are appropriate for particular physical systems. Help systems are not connected to analysis programs and are not meant to provide complete solutions to modeling problems. They simply guide the user—typically the novice user—in conducting some modeling tasks.

   Typically, help systems act as interactive passive consultants. They query the user on some key aspects of the problem and, based on the key problem features, inform the user on the appropriate sequence of commands, program options to select, analysis strategies to invoke, numerical parameters to assign, etc. The interaction is often through a question-and-answer session and custom menus. These help systems can be readily built using simple shells that provide

forward and/or backward chaining capabilities. With the advent of powerful personal computer software that can write and organize knowledge and communicate with the user through standard interfaces (HyperCard, HyperX), systems can be properly integrated in a variety of analysis and design environments.

2. Customized preprocessors. Customized preprocessors are knowledge-based programs that are integrated into the environment in which they operate. Customized preprocessors extract relevant features from a data base that describes the physical object to be modeled (often a simple geometric model). These features play the role of higher-level, symbolic descriptions that provide semantics to geometric entities. Features are used to classify various components and match them to corresponding analysis methods and parameters (e.g., finite element mesh density). These parameters are then used to derive special-purpose interfaces to produce input files for the appropriate analysis programs of the environment.

   The advantage of customized systems is that the user intervention in the modeling process is minimal. Essentially, the user is only required to enter some description of the physical object to be modeled. The preprocessors rely on the fact that the object can be adequately described in terms of a predetermined set of features encoded in the knowledge base and that a set of rules for modeling, analyzing, and evaluating these features exists. The structure of this class of knowledge-based systems can be analyzed in terms of three primitive tasks: (1) data abstraction (definitional, qualitative, or generalization abstractions); (2) heuristic associations between the abstracted features that characterize the object and the categories of known models; and (3) refinement that specializes a chosen model category to the particular problem at hand.

   Unfortunately, customized preprocessors are typically limited to narrow domains. This is due to the fact that they rely on structuring the objects of the domain in very specific ways: to fit the templates of a set of *a priori* chosen features. The ways in which models can be used must be anticipated and fixed when the system is designed. As the domain expands, significant knowledge engineering effort is required to find, organize, and encode the myriad pieces of knowledge needed to extract all the relevant features and analyze their potential interactions. The combinatorial explosion of rules needed to cover very large domains can become prohibitive.

3. High-level model generation tools. High-level model generation tools incorporate techniques that are more flexible than the heuristic classification approaches used by the systems discussed above. In particular, the goal of these tools

> is to put at the analyst's disposal a set of powerful representations and reasoning mechanisms that can be invoked as needed and that serve as a means of high-level specification of modeling operations.

A representation of modeling knowledge that can provide effective modeling assistance is an explicit representation of the assumptions that can generate various model fragments. Assumptions are the basic entities from which models are constructed, evaluated, and judged as adequate. Analysts often refer to and distinguish from various models by the assumptions they incorporate. The vocabulary of assumptions corresponds more closely to how analysts describe and assess the limitations of proposed models. Hence, the explicit representation and use of modeling assumptions in a modeling assistant can make the modeling operations correspond more closely to analyst's methods and could make it easier to organize and build a knowledge base. Assumptions encode a larger chunk of knowledge than rules and hence can provide a conceptual structure that is clearer and easier to manage than the typical knowledge-based or rule-based systems.

Turkiyyah and Fenves provide an example of how assumptions can be represented in a modeling assistant for their Generation and Interpretation of Finite Element Models in a Knowledge Based Environment[5]. In this representation, assumptions are modular units that, besides a prescription of how they affect the model, incorporate declarative knowledge about the scope of their applicability and their relevance for various modeling contexts, as well as their (heuristic) *a priori* and (definitional) *a posteriori* validity conditions. Assumptions can be used either directly by the analyst or indirectly through analysis objectives. When an analysis objective is posted, a planning algorithm selects an appropriate set of assumptions that can satisfy the modeling objective. These assumptions can then be applied automatically to generate a model that can be input into a finite element program.

Geometric reasoning techniques also provide the basis for high-level model generation tools. For example, geometric abstractions such as the skeleton can capture significant spatial object characteristics. Effectively, the skeleton is a symbolic representation of shape that corresponds to the way in which analysts visualize and describe shape and shape information, namely in terms of axis and width for elongated subdomains, center and radii for rounded subdomains, bisec

[5] R-90-188, Civil Engineering Department, Carnegie Mellon University, 1990.

tor and angle for pointed subdomains, etc. Because such abstractions are domain independent and hence general purpose, they can be used to suggest simplifications to the model (e.g., replace certain elongated two-dimensional regions by one-dimensional beam elements). They can also be used to subdivide a spatial domain into subregions of simple structure that can then be meshed directly by mapping or other mesh generation techniques.

MODEL INTERPRETATION

This section describes the AI potential to assist in postanalysis operations. Postanalysis operations are generically referred to as interpretation, although they involve distinctly different types of processes, including model validation, response abstraction, response evaluation, and redesign suggestions.

MODEL VALIDATION

Model validation determines if the mathematical model's numerical results accurately reflect the modeled system's real behavior. Knowledge-based techniques provide practical mechanisms to represent and characterize one important class of possible errors—idealization errors. Recently, a framework to validate idealized models has been proposed such that if the system model is methodically generated through the application of a set of assumptions any idealization error can be traced to one or more of those generative assumptions. Furthermore, each assumption encodes the conditions under which it is valid; hence, model validation involves checking the validity conditions of individual assumptions. There are many ways of verifying assumptions ranging in complexity from the evaluation of simple algebraic expressions to analysis of a model that is more detailed than the original one.

ABSTRACTION OF NUMERICAL RESULTS

Response abstraction is the task of generating some abstract description of the raw numerical results obtained from the analysis program. This description is presented in terms of high-level aggregate quantities representing key response parameters or behavior patterns.

Response abstractions can be classified in two types. The first, functional response abstractions, depends on the role that various subsystems or components play outside the model proper, that is, the meaningful aggregate quantities that are generated depend on knowledge of characteristics beyond the geometric and material properties of the system. The ability to generate the function-dependent response abstraction depends on the ability to represent the functional information that underlies the object modeled.

The second type of response abstraction is function independent. One seeks to recognize patterns, regularities, and interesting aspects and generate qualitative descriptions (e.g., stress paths) of the numerical results, independent of the functional nature of the object being modeled. Techniques from computer vision and range data analysis can be used to generate these interpretations. Well-developed vision techniques such as aggregation, segmentation, clustering classification, and recognition can be applied to the task. One interesting use of response abstractions assists the user in checking the "physical reasonableness" of numerical results by comparing the response abstractions between more refined models and simpler models.

CONFORMANCE EVALUATION

Conformance evaluation is the task of verifying that the computed results satisfy design specifications and functional criteria such as stress levels, ductility requirements, or deflection limitations. Conformance evaluation is largely a diagnostic problem, and the well-developed techniques for diagnosis can be applied to the task. Conformance evaluation requires heuristics on what are the (1) possible failure modes, (2) applicable code and standard provisions, (3) response quantities (stresses, deflections, etc.) affected by the provisions, and 4) approximations of the provisions and responses necessary.

A major issue in developing expert systems for conformance evaluation is the representation of code and standard provisions in a form suitable for evaluation yet amenable to modification when standards change or an organization wishes to use its own interpretations. One technique suitable for this purpose is to represent standards provisions by networks of decision tables.[6]

[6] The use of this representation in an expert system environment is demonstrated in J.H. Garrett and S.J. Fenves, 1989, "Knowledge Based Standard-Independent Member

INTEGRATION IN DESIGN

Analysis is rarely, if ever, an end in itself. The overwhelming use of analysis is to guide and confirm design decisions. One important application of AI techniques is to provide redesign recommendations when the analyzed system's response is not satisfactory. One problem that has to be addressed is the nature of the knowledge base that can generate redesign recommendations. Should it be separate and independent, or should it use the same modeling knowledge responsible for generating and interpreting models? The second approach formulates redesign as a goal-oriented problem: Given some deficiencies uncovered by an analysis, what modifications to the design object are required so that a model whose response satisfies the design specifications can be generated?

Another problem that has to be addressed if computational methods are to be adequately incorporated in CAD is the general capability to provide analysis interpretations and design evaluations compatible with the progress of the design process from the initial conceptual sketch to a fully detailed description. Evaluations should occur at increasingly higher degrees of refinement throughout the design process. Early simple models can provide early feedback to the designer (or design system) on the global adequacy of the design, while evolved models, paralleling the evolving design, help to guide the designer in the detailed design stages. An important issue in developing general mechanisms for hierarchical modeling is how to generate and represent various kinds of geometric abstractions.

NUMERICAL MODEL FORMULATION

The term formulation denotes the process of producing a computational mechanics capability—a set of numerical routines—from a representation of a physical phenomenon. It is well known that the development of a computational mechanics program is time consuming, expensive, and error prone. Processes that can help in the quick development of reliable numerical software can be of great practical benefit. Ideas from AI can contribute significantly to various aspects of the formulation process: performing symbolic computations, expressing subroutines in a form that makes them reus

Design," Journal of Structural Engineering, Vol. 115, No. 6, June.

able, designing large systems with appropriate data abstractions, assisting in the synthesis of computational mechanics programs, and integrating heuristics and knowledge-based methods into numerical solutions. These aspects are examined in turn.

SYMBOLIC PROCESSING

One aspect of program development that is particularly time consuming and error prone is the transition from a continuum model involving operations of differentiation and integration to a computational model involving algebraic and matrix operations. A branch of AI deals with symbolic computations, culminating in symbolic computation programs such as MACSYMA and Mathematica. Programs in this class operate on symbolic expressions, including operations of differentiation and integration, producing resulting expressions in symbolic form. A particularly attractive practical feature of these programs is that the output expressions can be displayed in Fortran source code format.

The potential role of symbolic processing has been investigated by several researchers. Studies indicate that symbolic processing can significantly assist the generation of computational model components to be incorporated in the source code. Symbolic generation of source code for clement stiffness and load matrices can eliminate the tedious and error-prone operations involved in going from a differential algebraic representation to a discrete procedural representation. Additional run-time efficiency improvements are possible through functionally optimized code and the use of exact integrations. Finally, conceptual improvements are possible, such as symbolically condensing out energy terms that contribute to shear and membrane locking.

REUSABLE SUBROUTINES

Numerical subroutines that perform function evaluations, domain and boundary integrations, linear and nonlinear equation solving, etc., abound in computational mechanics codes. However, the typical implementation of these subroutines bears little resemblance to our mathematical knowledge of the operations they perform. They are written as a sequence of concrete arithmetic operations that include many mysterious numerical constants and are tailored to specific machines. Because these routines do not exhibit the structure

of the ideas from which they are formed, their structure is monolithic, handcrafted for the particular application, rather than constructed from a set of interchangeable parts that represent the decompositions corresponding to the elemental concepts that underly the routine. Often numerical routines are difficult to write and even more difficult to read.

The idea of expressing mathematical routines constructively is widely applicable. Even the simplest routines that are of ten thought of as "atomic"—such as $\sin(x)$—can be constructed from their primitive constituent mathematical operations, that is, periodicity and symmetry of the sine function and a truncated Taylor expansion. Abelson et al.[7] shows how Romberg's quadrature can be built by combining a primitive trapezoidal integrator with an accelerator that speeds the convergence of a sequence by Richardson extrapolation. The idea is that instead of writing a subroutine that computes the value of a function, one writes code to construct the subroutine that computes a value.

Such a formulation separates the ideas into several independent pieces that can be used interchangeably to facilitate attacking new problems. The advantages are obvious. First, clever ideas need to be coded once in a context independent of the particular application, thus enhancing the reliability of the software. Second, the code is closer to the mathematical basis of the function and is expressed in terms of the vocabulary of numerical analysis. Third, the code is adaptable to various usages and precisions because the routine's accuracy is an integral part of the code rather than a comment that the programmer adds; just changing the number that specifies the accuracy will generate the single, double, and quadruple precision versions of a subroutine.

Writing subroutines in this style requires the support of a programming language that provides higher-order procedures, streams, and other powerful abstraction mechanisms available in functional languages. The run-time efficiency does not necessarily suffer. The extra work of manipulating the function's construction need be done only once. The actual function calls are not encumbered. Moreover, because functional programs have no side effects, they have no required order of execution. This makes it exceptionally easy to execute them in parallel.

[7] H. Abelson et al., 1989, "Intelligence in Scientific Computing," Communications of the ACM, Vol. 32, No. 5, pp. 546–562, May.

PROGRAMMING WITH DATA ABSTRACTIONS

The current generation of computational mechanics software is based on programming concepts and languages that are two or three decades old. As attention turns to the development of the next generation of software, it is important that the new tools, concepts, and languages that have emerged in the interim be properly evaluated and that the software be built using the best appropriate tools.

Designs for finite element systems based on object-oriented concepts have begun to emerge in the literature. As these designs show, object-oriented programming, an offshoot of AI research, can have a major impact on computational mechanics software development. It is possible to raise the level of abstraction present in large-scale scientific programs (i.e., allowing finite element programmers to deal directly with concepts such as elements and nodes) by identifying and separating levels of concern. Programs designed in this manner allow developers to reason about program fragments in terms of abstract behavior instead of implementation. These program fragments are referred to as objects or data abstractions, their abstract quality being derived from precise specifications of their behavior that are separate and independent of implementation and internal representation details.

MODEL SYNTHESIS ASSISTANCE

While the bulk of today's computational mechanics production work is done by means of large comprehensive programs, there is a great deal of exploratory work requiring the development of "one-shot" ad hoc custom-built programs. Developers of such ad hoc programs may have access to subroutine libraries for common modules or "building blocks" but not much else. These developers frequently have to reimplement major segments of complete programs in order to "exercise" the few custom components of their intended program.

One potential application of AI methodology is an expert system to assist in synthesizing computational programs tailored to particular problems, on the fly. The system would require some specifications of the program goal; the constraints (e.g., language, hardware environment, performance); and the description of custom components (e.g., a new equation solver, a new element, a new constitutive equation for a standard element) as input. The system's knowledge base would contain descriptions of program components with their attributes (lan

guage, environment, limitations, interface descriptions, etc.) and knowledge about combining program components, which might include the knowledge to write interface programs between incompatible program segments. The expert system would have to use both backward and forward chaining components—the former to break down the goal into the program structure and the latter to select program components to "drive" the low-level custom components.

MONITORING NUMERICAL SOLUTIONS

Combining numerical techniques with ideas from symbolic computation and with methods incorporating knowledge of the underlying physical phenomena can lead to a new category of intelligent computational tools for use in analysis. Systems that have knowledge of the numerical processes embedded within them and that can reason about the application of these processes can control the invocation and evolution of numerical solutions. They can "see what not to compute" (Abelson 1989) and take advantage of known characteristics of the problem and structure of the solution to suggest data representations and appropriate solution algorithms.

The coupling of symbolic (knowledge-based) and numerical computing is particularly appropriate in situations where the application of pure numerical approaches does not provide the capabilities needed for a particular application. For example, numerical function minimization methods can be coupled with constraint-based reasoning methods from AI technology to successfully attack large nonlinear problem spaces where numerical optimization methods are too weak to find global minima. To derive a solution to the problem, domain-specific knowledge about problem solving in terms of symbolic constraints guides the application of techniques such as problem decomposition, constraint propagation, relaxation, and refinement.

COMPREHENSIVE MODELING ENVIRONMENTS

As higher level modeling tools are built and larger modeling knowledge bases are constructed, issues such as integration, coordination, cooperative development, and customization become critical. A framework for a general finite element modeling assistant has also been proposed recently in the literature. The framework is intended to permit a

cooperative development effort involving many organizations. The key feature of the framework is that the system consists of a set of core knowledge sources for the various aspects of modeling and model interpretation that use stored resources for the problem-dependent aspects of the task. In this fashion, new problem types, as well as individual organizations' approaches to modeling, only involve expansion of the resources without affecting the knowledge sources.

In the comprehensive framework envisaged, the core knowledge sources would perform the function of model generation and interpretation and program selection (with possible customization and synthesis) and invocation. The three major resources used by these knowledge sources are as follows:

1. Physical class taxonomies. These represent an extended taxonomy or semantic network of the various classes of physical systems amenable to finite element modeling and the assumptions appropriate for each class. Their purpose is to provide pattern matching capabilities to the knowledge sources so that the definition of problem class and key problem parameters can be used by the knowledge sources in their tasks at each level of abstraction. The major design objective in developing these taxonomies will not be to avoid exhaustive enumeration of individual problems to be encountered, but rather to build a multilevel classification of problem types based on their functionality, applicable assumptions, behavior, failure modes, analysis strategies, and spatial decompositions. It is also expected that a large part of knowledge acquisition can be isolated into modifying these taxonomies either by specialization (customization to individual organization) or generalization (merging or pooling knowledge of separate organizations).
2. Program capability taxonomies. In a manner similar to the above, these taxonomies represent the capabilities, advantages, and limitations of analysis programs. The taxonomy must be rich enough so that the knowledge source invoking the programs can make recommendations on the appropriate program(s) to use based on the high-level abstractions generated by the other knowledge sources or, if a particular program is not available in the integrated system, to make recommendations on alternate modeling strategies so that the available program(s) can be effectively and efficiently used. As with the previous taxonomy, the program capability taxonomy needs to be designed so that knowledge acquisition about additional programs can be largely isolated to the expansion of the taxonomy data base.

3. Analysis programs. The programs, including translators to and from neutral files as needed, are isolated in the design to serve only as resources to solve the model. The issues in this interconnection are largely ones of implementation in coupling numerical and knowledge-based programs. Modern computing environments make such coupling relatively seamless.

Research Issues

This section discusses briefly two important problems that must be addressed before reliable modeling environments, such as the one discussed above, can be built. The first problem is the need to provide more flexibility to knowledge-based systems, and the second is the need to compile a core of modeling assumptions.

FLEXIBLE KNOWLEDGE-BASED SYSTEMS

The present generation of knowledge-based systems has been justly criticized on three grounds: they are brittle, idiosyncratic, and static.

Present knowledge-based systems are brittle—in the sense used in computer science as a contrast to "rugged" systems—in that they work in a very limited domain and fail to recognize, much less solve, problems falling outside of their knowledge base. In other words, these systems do not have an explicit representation of the boundaries of their expertise. Therefore, there is no way for them to recognize a problem for which their knowledge base is insufficient or inappropriate. Rather than exhibiting "common sense reasoning" or "graceful degradation," the systems will blindly attempt to "solve" the problem with their current knowledge, producing predictably erroneous results. Current research on reasoning from first principles will help overcome this problem. Combining first principles with specialized rules will allow a system to resort to sound reasoning when few or no specialized items in its knowledge base cover a situation. First principles can also be used to check the plausibility of conclusions reached by using specialized knowledge.

A KBES developed using the present methodology is idiosyncratic in the sense that its knowledge base represents the expertise of a single human domain expert or, at best, that of a small group of domain experts. The system thus reproduces only the heuristics, assumptions, and even style of

problem solving of the expert or experts consulted. The nature of expertise and heuristics is such that another, equally competent expert in the domain may have different, or even conflicting, expertise. However, it is worth pointing out that a KBES is useful to an organization only if it reliably reproduces the expertise of that organization. At present, there appear to be no usable formal methods for resolving the idiosyncratic nature of KBESs. There are some techniques for checking the consistency of knowledge bases, but these techniques are largely syntactic. One practical approach is to build a domain-specific metashell that contains a common knowledge base of the domain and excellent knowledge acquisition facilities for expansion and customization by a wide range of practitioners.

Present KBESs are static in two senses. First, the KBES reasons on the basis of the current contents of its knowledge base; a separate component, the knowledge acquisition facility, is used to add to or modify the knowledge base. Second, at the end of the consultation session with a KBES, the context is cleared so that there is no provision for retaining the "memory" of the session (e.g., the assumptions and recommendations made). Research on machine learning is maturing to the point where knowledge-based systems will be able to learn by analyzing their failed or successful performance—an approach called explanation-based learning.

COMPILATION OF MODELING KNOWLEDGE

One task of great practical payoff is the development of a knowledge base of modeling assumptions that contains what is believed to be the shared knowledge of analysts. Such a core knowledge base will be beneficial in two important ways. First, it could be used as a starting point to build a variety of related expert systems, hence making the development cycle shorter. Second, such a knowledge base could become the "corporate memory" of the discipline and, hence, could give insights into the nature of the various aspects of modeling knowledge. One starting point to build such a knowledge base is to "reverse engineer" existing models to recognize and extract their assumptions.

Two useful precedents from other domains offer guidance. Cyc is a large-scale knowledge base intended to encode knowledge spanning human consensus reality down to some reasonable level of depth—knowledge that is assumed to be shared between people communicating in everyday situations. Cyc is a 10-year effort that started in 1984.

Progress to date indicates that the already very-large knowledge base (millions of assertions) is not diverging in its semantics and already can operate in some common situations.

Knowledge-based Emacs (KBEmacs) is a programmer's apprentice. KBEmacs extends the well-known text editor Emacs with facilities to support programming activities interactively. The knowledge base of KBEmacs consists of a number of abstract programs (cliches) ranging from very simple abstract data types such as lists to abstract notions such as synchronization and complex subsystems such as peripheral device drivers. The fundamental idea is that the knowledge base of cliches encode the knowledge that is believed to be shared by programmers. KBEmacs has been used successfully to build medium-sized programs in Lisp and Ada.

SUMMARY

The objective of this appendix was to present some of the concepts and methodologies of AI and examine some of their potential applications in various aspects of computational mechanics. The methodologies sketched herein are maturing rapidly, and many new applications in computational mechanics are likely to be found. Undoubtedly, AI methodologies will eventually become a natural and integral component of the set of computer-based engineering tools to the same extent as present-day "traditional" algorithmic tools. These tools will then significantly elevate the role of computers in engineering from the present-day emphasis on calculation to the much broader area of reasoning.

5

ENVIRONMENTAL CONTAMINATION AND GEOMECHANICAL SIMULATION

One of the most important problems facing society today is the environment's protection from various sources of contamination. In order to prevent or remove contamination from the air, earth, and water, large-scale computational models are needed that can accurately predict and help control the transport of contaminants. Contamination is found at all length scales. The strongly nonlinear properties of transport models make the scale-up necessary for global simulation extremely difficult.

Another problem that poses a critical threat to the country's economic stability is our rapidly growing dependence on foreign energy sources. The United States produces less than one-half of its petroleum and does not have the overall supplies or increased production capabilities needed to fully support its energy requirements for any extended period of time. Substantial additions must be made to reserves through the location of new sources via exploration or the development of ways to enhance the production of known resources with better technology. The Bush administration has recently acknowledged these needs by recommending significant increases in the Department of Energy's budget aimed at improving enhanced oil recovery (EOR) technology.

These broad topics, although quite different in societal motivation and concerns, actually possess many common technical issues from a computational mechanics standpoint. The research areas of subsurface contaminant transport, geographical exploration, and enhanced production of fossil fuels exhibit a common thread for the need to understand and predict the movement of fluids or mechanical disturbances through heterogeneous porous media on many length scales. They all require large-scale simulations on enormous computational domains. Each of these problems has been mentioned in the context of grand challenges in computational science. They each hold the potential for significant impact through major advances in computational mechanics.

STATUS OF GEOMECHANICAL SIMULATION

Although the increase in computing capabilities and advances in computational mechanics has allowed significant progress in these area's simulation capabilities, existing models are not adequate to address many of the important aspects of these problems. Basic capabilities have been established to predict certain aspects of contaminant transport, petroleum recovery, and seismic exploration. However, models involving less restrictive assumptions are needed to understand and predict the physical processes and to address societal problems effectively.

Currently, basic trends and large-scale movements of contaminants through fairly homogeneous media can be predicted. However, the multiphase flow properties of transport in unsaturated zones are not well understood, nor is the transition from unsaturated to saturated flow with contaminants that form separate phases. The mass transport between phases needed to model migration and vaporization of volatile organic contaminants in the unsaturated zone is not currently modeled sufficiently well to develop efficient monitoring techniques for gasoline spills or leaks. Tools are still needed to model varying soil properties well enough to accurately predict flow in heterogeneous media. Although restoration techniques are beginning to be addressed by both microbial and chemical flooding techniques, the growth/decay, absorption/desorption, or heterogeneous flow parameters are not well enough known to develop effective remediation strategies.

In primary or secondary petroleum recovery where multiphase flow occurs to natural formation pressure or to artificial water flooding, the basic mechanisms are understood, and massive field-scale simulations can aid greatly in developing field production strategies. However, an average of 60 to 70 percent of the original oil is left in the reservoir after this production; this residual oil is the target for EOR techniques. It involves injecting fluids or chemicals into the pores of the reservoir rock, which reduces the surface tension or viscosity that traps the oil and allows the previously trapped hydrocarbon to flow toward production wells. Modeling of the complex fluid/fluid or fluid/rock interactions involved in these physical and chemical processes leads to large coupled systems of nonlinear equations. Although significant progress has been made in these applications, they still require major advances in computational mechanics before the effect of a field-scale EOR process in a heterogeneous reservoir can be routinely predicted.

Although billions of dollars are spent each year on geophysical prospecting techniques, this application still remains one of the least understood problems in computational mechanics due to its lack of uniqueness and the difficulty in solving the so-called inverse problem. The probability of locating a reserve with seismic exploration can clearly be increased, but the ability to predict the earth's internal structure on any fine scale is still well beyond contemporary modeling techniques. Seismic techniques can currently identify the major structural traps that could potentially contain hydrocarbons. However, refined computational mechanics techniques could greatly aid in the detection of more subtle stratigraphic traps or pressure cells and significantly increase petroleum reserves.

PAST COMPUTATIONAL ADVANCES

Current modeling capabilities in these important applications have been made possible through a variety of advances in computational mechanics. Numerical solution of the convection-dominated models used for both contaminant transport and petroleum recovery is quite difficult. The complicated fluid/fluid interactions that govern the multiphase or multicomponent flow processes occur in narrow moving regions of interaction, which often involve small viscous or diffusive phenomena. They are strongly nonlinear and strongly coupled systems that must be solved simultaneously. The enormous computational domains and complex inverse procedures to obtain the parameters for flow have posed, in general, significant challenges to the computational mechanics community.

Large upstream-weighted cell-centered finite difference codes have been developed and optimized in the petroleum industry. Efficient solution of the large coupled systems of nonlinear partial differential equations has required advances in linearization and quasi-linearization techniques, block eliminations, and iterative solution methods, as well as in data storage and retrieval.

The need to treat complex boundaries and flow regimes in contaminant transport has led to the use of finite element methods for these applications. Advances in the use of finite elements in fluids have appeared in recent years. Since governing fluid/fluid interactions occur in moving zones, techniques such as streamline diffusion methods are becoming more popular and hold great promise. Efficient linear solution techniques such as multigrid methods have been used recently

to solve the nonstructured matrices arising from triangular and irregular grids.

Petrov-Galerkin methods, with the choice of test function based on stabilizing the transport terms, have been used with some success. Similarly, Eulerian-Lagrangian techniques and the modified method of characteristics have proven quite useful in treating transport in porous media. These methods require accurate fluid velocities. Mixed finite element methods have been used to yield accurate approximations of the latter, even under heterogeneous and unstable flow conditions.

Many of the phenomena governing transport in porous media are highly localized in time and space. Although local grid refinement is common in finite element codes, its use in large structured finite difference codes has been greatly restricted due to the increased complexity of the resulting matrices. Recently, the combination of domain decomposition techniques and local refinement in time and space has been applied efficiently in large industrial cell-centered codes using finite difference discretization.

Effective solution of the inverse seismic problem requires accurate and efficient solution of the "forward" problem where the response to input sources is calculated on a specified set of geological structures. The forward problem must be solved repeatedly with different structural locations in the course of solving the inverse problem associated with determining the true subsurface configuration from measured surface data. Given the vastness of the earth's interior, the choice of a reasonable computational domain requires the introduction of artificial computational boundaries. Until recent advances were made in the development of absorbing boundary conditions, computed reflections from such computational boundaries destroyed information about subsurface structure. Improvements have also been made in the modeling of reflections from interior stratigraphic discontinuities. Such advances in computational mechanics for the forward problem now allow the inverse problem to be addressed meaningfully.

The emergence of vector-based supercomputers has revolutionized the solution of each problem mentioned above. Different applications can take advantage of differences in hardware architectures, thus requiring distinct algorithm developments. For example, the CYBER 205 architecture performs quite well for seismic problems where extremely long vectors can be manipulated effectively. Although the CRAY computers cannot take advantage of very long vectors as well as the CYBER 205, the cycle time is significantly better and

the reservoir simulation codes perform extremely well on the CRAY. Highly vectorized codes have been developed for each of the major vector computers by taking advantage of their particular architectural features. Although some codes have been ported to the emerging parallel architecture computers, the machines' parallelizing compilers are not well developed. The choice of the granularity level for code development on parallel algorithms for complex problems is not clear, but the potential for massively parallel machines is enormous.[8]

NEEDED ADVANCES

Much research is needed in the development of more accurate and useful models in the environmental and energy areas. A major goal is a better understanding of the model properties, especially in the nonlinear regime and how they scale through various representative lengths. There is the need to fully develop the ability to use the computer as a laboratory to test model concepts, parameters, sensitivities, etc., in order to obtain better knowledge of the properties of the models and the processes they represent. If these large-scale simulators are to be used as a tool to build intuition and understanding, overall computational efficiency is essential. Visualization techniques that allow the scientist or engineer to view the results of his/her calculations in real time are also critical in this computational laboratory concept.

One of the major difficulties in accurate earth transport modeling is its heterogeneous nature at many different length scales. It is very difficult to obtain effective global parameters from those that vary rapidly on smaller-length scales, even for linear problems. The nonlinear properties of models greatly complicate their development. Volume averaging, homogenization, geostatistical averaging, and the current use of micromodels are to develop effective parameters. Extensive comparison between fine-grid and coarse-grid simulations using effective parameters will be necessary for a major breakthrough. In the scale-up process the basic constitutive laws must also be described differently to relate the different effective flow properties at various length scales. The homogenization from Navier-Stokes models at a pore scale to Darcy's law at field scale is an important example of this idea. The effects of heterogeneity and viscous fingering, however,

[8] See Appendix 3, Parallel Computation.

may require qualitatively different diffusion terms on various scales.

Although the importance of reliability is clearly acknowledged in many fields such as structural design, it is not emphasized enough in the environmental and energy areas. To evaluate and control environmental damage, there is a distinct need to develop bounding calculations for the transport of contaminants. Designing waste disposal strategies requires the ability to certify the reliability of the computational models. Also, if the local moving interfacial phenomena in EOR methods are not resolved, there will be no guarantee that the production strategy will be successful, much less optimal.

As noted earlier, adaptive strategies have received considerable attention in many areas of computational mechanics. The use of *a priori* and *a posteriori* error estimators has proven to be quite effective for many application areas. These concepts must be developed more fully for nonlinear dynamic applications and must be utilized in energy and environmental problems.

A promising area for major advances in adaptive strategies for large-scale codes is the use of domain decomposition techniques. Adaptive local grid refinement methods are being developed for accurate treatment of localized phenomena. Combinations of these with domain decomposition techniques can achieve the benefit of adaptivity without the loss of efficiency. Global decomposition of the computational domain into pieces for parallel solution or decomposition of the problem on coarser-grained physical criteria each allows for more efficient solution algorithms. Also, different constitutive equations can be used on different domains to address scale dependencies. Major advances in these various forms of domain decomposition can help to exploit the enormous potential of the parallel architecture supercomputers.

Although massively parallel computing can potentially break the speed barriers of serial computers, significant advances in utilizing parallelism will require extensive research. The compilers for these parallel architecture computers are far from being able to determine automatically the optimum granularity for parallelism. Similarly, even if a granularity is given, there is a great need for algorithm development for efficient utilization of the computers' parallel capabilities.

Finally, major advances in computational mechanics will be considerably more effective if they are achieved in a multidisciplinary environment. Significant progress can be made by interaction of scientists or engineers who understand the aspects of the physical problem, mathematicians and

numerical analysts who know the mathematical and numerical properties of the mathematical model, and computer scientists who can help take full advantage of the architecture of the specific computer to be used.

Significant advances in computational mechanics as described above could have enormous payoffs in the areas of energy and environment. The ability to model and understand the physics of flow on increasing length scales in heterogeneous media could allow major breakthroughs in the ability to control the expanding contamination of our environment and to develop effective remediation strategies. Although we are far from these capabilities now, they are possible through scientific advances and are critical to maintaining environmental integrity. The rapid progress being made in computer hardware computational capabilities finally allows enough physics to be incorporated into the complex EOR models to permit accurate flow descriptions at various scales. The algorithm development to utilize this hardware growth is lagging severely. The enormous potential of enhanced modeling to increase significantly our accessible petroleum reserves can be achieved only through major advances in computational mechanics.

6

MATERIALS RESEARCH: NUMERICAL MODELING OF NEW MATERIALS

Computational mechanics is an essential tool in the quest to engineer materials for specific purposes. Many of the manufacturing and performance-limiting issues of these materials are appropriately addressed within a continuum mechanics framework. The numerical solution of continuum mechanics problems is used to (1) analyze microscale deformation modes in order to design materials and processes, (2) develop macro-scale constitutive and damage laws, and (3) design machines and structures that fully utilize the capabilities of these new materials.

The performance of a material in service is controlled by a large set of properties. One purpose of computational mechanics studies is to link properties as measured in standard tests to performance under service conditions. For example, standard test procedures exist to measure a material's resistance to crack growth. Computational mechanics analyses are used to correlate performance in such a test with crack growth behavior under service conditions and relate the property measured in this test to material microstructure. Another purpose is to relate properties to measurable and controllable features of the material's microstructure such as the orientation of crystal structure, the configuration of phases, and the morphology and distribution of second-phase particles.

Much of the computational mechanics of materials is finite element based and therefore, to a large extent, is an unanticipated outgrowth of development in computational structural mechanics. Formulations and algorithms developed in the structural mechanics context for problems involving both geometrical and material nonlinearities have provided the foundation to develop a powerful capability to analyze and predict material behavior. Although the foundations have been laid, much remains to be done to make this capability an engineering design tool in the same sense as computational structural mechanics. The potential payoff is enormous.

One can envision real-time calculations being used to control processes in order to maximize throughput and ensure

desirable product properties, even tailoring product properties in real time for different applications. For example, the mechanical properties and formability of metal sheet depend on the crystallographic texture (the orientation of the crystal axes of the microscopic grains that make up a polycrystalline structural metal). In turn, the texture of the metal sheet coming out of a rolling mill depends on the mill settings. At present, calculation of texture under complex loading histories is itself a subject of much research. The ability to compute texture development quickly and accurately would permit cheap and efficient tailoring of sheet products for specific applications. This capability will undoubtedly be developed. The only question is when and by whom.

Off-line analyses permit "what if" studies of the implications of alternative manufacturing options and alternative material microstructures. The design of composite materials involves considering a very large number of possibilities, even presuming that the basic matrix and reinforcing material are given. What are the optimal volume fraction and second-phase size and shape distributions necessary to achieve specified performance goals at reasonable cost? If the shape of the reinforcement phase can be changed, does improvement in performance justify the extra effort and cost? How strong should the interface bond be? Too strong a bond can reduce energy dissipation in certain failure modes, which makes the composite brittle and unsafe in certain applications. Too weak a bond can reduce stiff ness and degrade other mechanical properties to the point where the composite is unsuitable for the intended application.

The following examples illustrate some current computational capabilities on issues related to new material design and performance assessment.

Figure 6.1 shows some results for an aluminum-lithium alloy that fails by grain boundary cavitation at room temperature. A complete compact tension specimen was analyzed, but Figure 6.1 shows the mode of crack propagation in the region just ahead of the crack. The material is described by a porous plastic constitutive relation that allows for creation of a new free surface, and the elements in which there has been a complete loss of stress-carrying capacity are deleted from the mesh plot. The model is a highly idealized one; nevertheless, there are no adjustable parameters. All material parameters input into the analysis were independently measured or inferred. The quantitative agreement between the model predictions and the experimentally measured values was quite good. The predicted critical stress intensity values were within a factor of two of the experimental values, while the

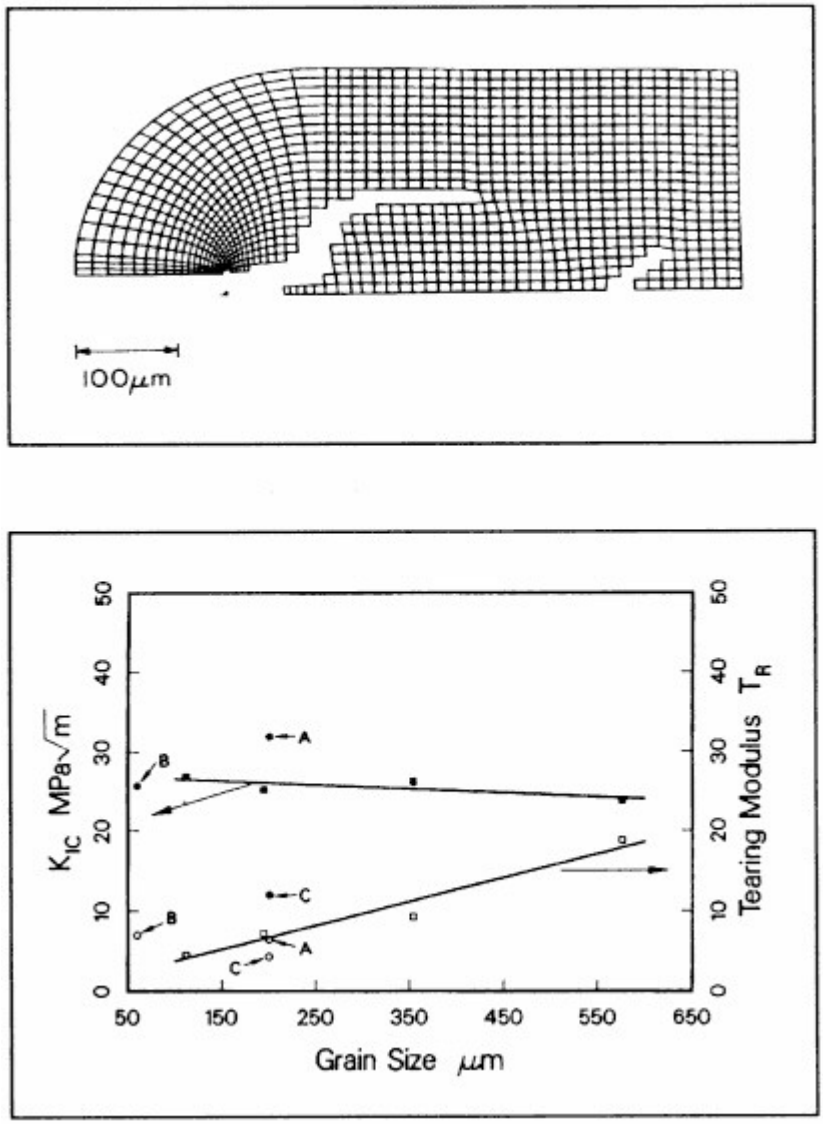

Figure 6.1
Mode of crack propagation and comparison of predicted and experimental toughness and tearing modulus. Source: R. Becker, A. Needleman, S. Suresh, V. Tvergaard, and A.K. Vasudevan, 1989, *Acta Metallurgica*, Vol. 37, pp. 99–120. Reprinted with permission, Pergamon Press, New York, 1988.

predicted and observed values of the tearing modulus (a measure of material resistance to crack growth) were typically within 50 percent of each other.

A possible method for improving the toughness of ceramics is to use materials, such as zirconia partially stabilized with magnesia, that undergo a stress-induced phase transformation. Constitutive models have been developed that represent the shear strains and volume change accompanying the phase transformation. Finite element analyses based on these models have successfully predicted toughness enhancements and the mode of fatigue crack growth at notches. Figure 6.2 shows that the direction of crack growth follows the predicted shape of the residual tensile stress contours near a notch tip growth from a notch tip and predicted contours of constant normal stress at the notch tip. Another key issue is to understand the role of microcracks on toughness to improve the performance of ceramic materials. Figure 6.3 shows a picture of a random array of microcracks in a model of a polycrystalline ceramic and computed crack growth resistance curves. The resistance to the extension of a macroscopic crack builds up as a result of the wake of microcracked material left behind by the advancing large crack. This gives an increase in toughness of up to 40 percent for mode I conditions (tensile crack loading). However, subsequent work showed that there is no appreciable gain in toughness due to microcracking when the crack loading contains a significant shear component. This has important implications for the design of ceramics tough enough for a broad range of structural applications.

Structural steel is a complex multiphase material. The development of high-strength steels, with properties optimized for specific applications, poses a considerable challenge. In high-purity, high-strength steels, for an important range of loading conditions, the ductility and toughness-limiting event is void nucleation by debonding of grain-refining particles. Figure 6.4 shows the predictions of a continuum model for the debonding strain as a function of grain-refining particle size, with the volume fraction for each size adjusted to attain a constant grain size. The prediction is that substantial improvements in toughness can be obtained by decreasing the grain-refining dispersion particle size. Whether or not such performance actually occurs remains to be seen.

Convergence of calculation procedures is an important issue. The phenomena computed must be reliable characteristics of the problem formulation and not of the spatial or temporal discretization. For material modeling problems that may involve large deformations, path-dependent material response, instabilities, and the creation of new free surface, essentially

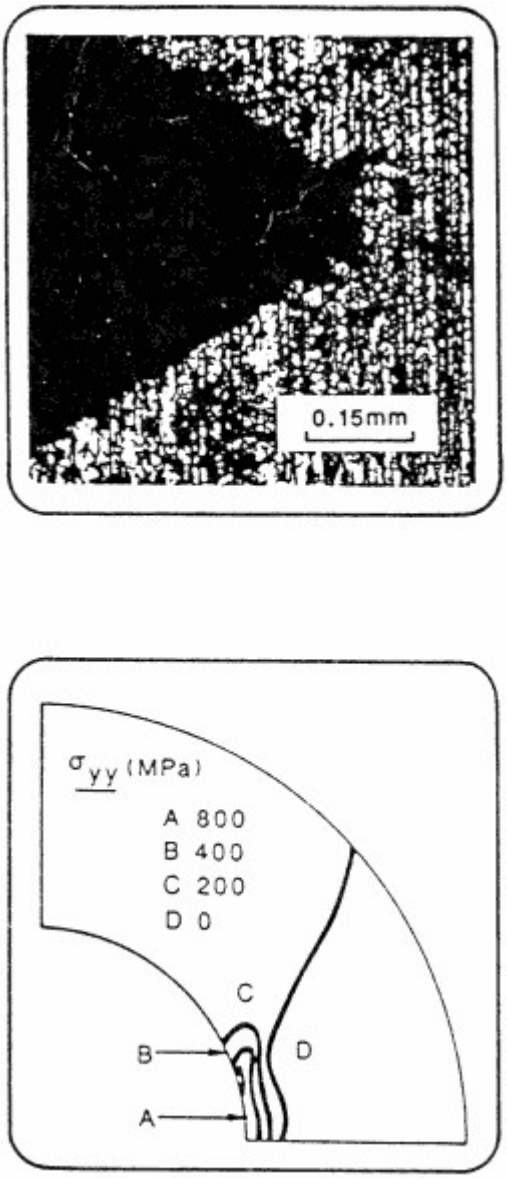

Figure 6.2
Crack growth from a notch tip and predicted contours of constant normal stress at the notch tip. Source: S. Suresh and J.R. Brockenbrough, 1988, *Acta Metallurgica*, Vol. 36, pp. 1455–1470. Reprinted with permission, Pergamon Press, New York, 1988.

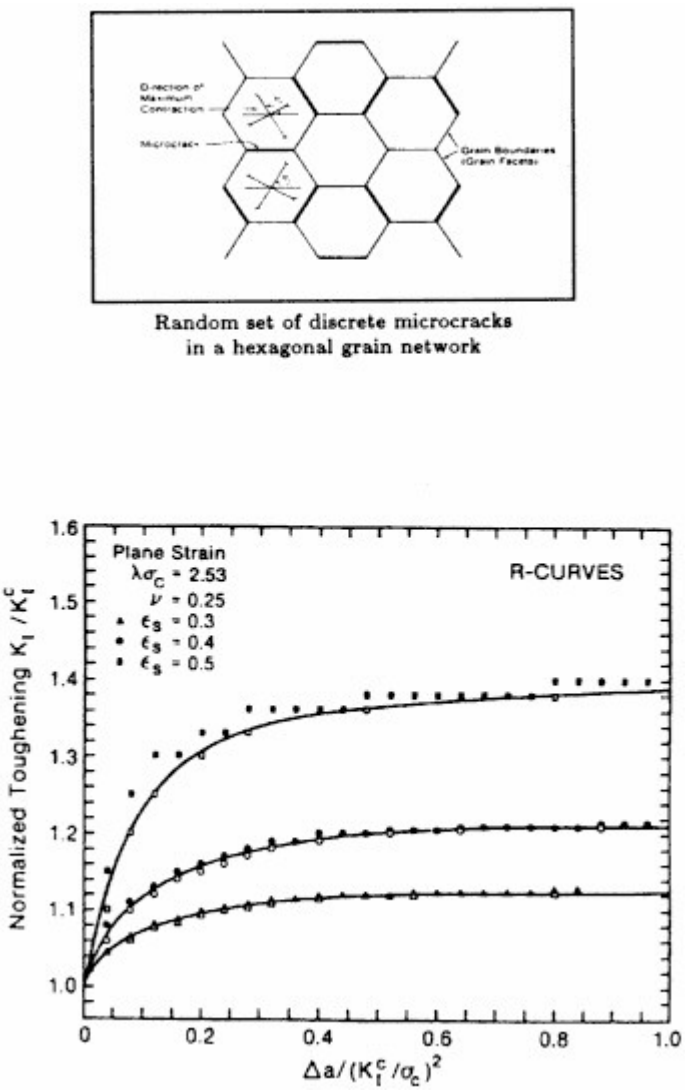

Figure 6.3
Sketch of facet microcracks and predicted crack growth resistance. Source: P.G. Charalambides and R.M. McMeeking, 1987, *Mechanics Materials*, Vol. 6, pp. 71–87. Reprinted with permission, Elsevier-*Science Publishers, Amsterdam, The Netherlands.*

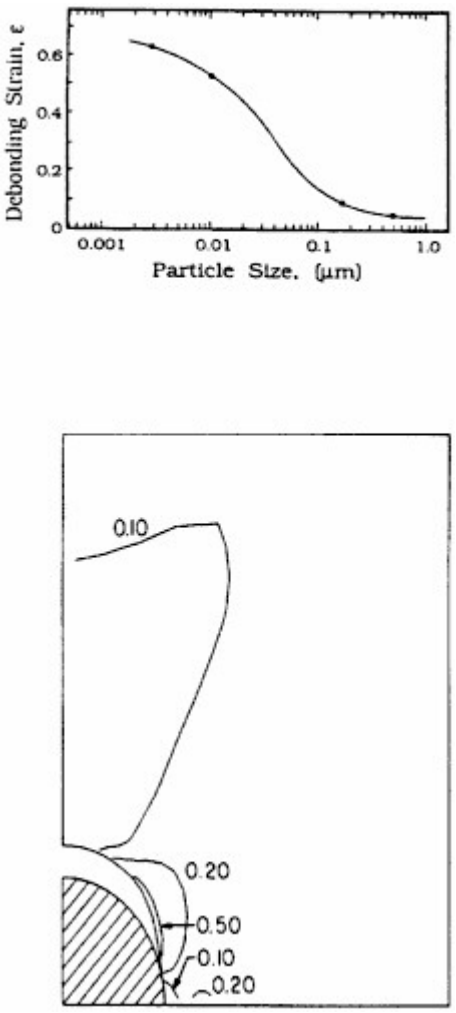

Figure 6.4
Debonding process and predicted particle debonding strain versus particle diameter. Source: G.B. Olson, "Overview: Science of Steel," in *Innovations in Ultrahigh-Strength Steel Technology*, Proceedings of the 34th Sagamore Army Materials Research Conference, ed. by G.B. Olson, M. Azrin and E.S. Wright, pp. 3–66. Reprinted with permission. Creep crack growth versus time for two meshes. Source: F.Z. Li, A. Needleman, and C.F. Shih, 1988, *International Journal of Fractures*, Vol. 36, pp. 163–186.

nothing is known on a fundamental mathematical basis about the convergence behavior of the algorithms used. Numerical experimentation has given some feeling for which features are mesh sensitive and which are not and how this sensitivity depends on apparently subtle features of the problem formulation. Figure 6.5 shows results from a finite element calculation of creep crack growth using a micromechanically based constitutive relation that allows for the loss of stress-carrying capacity due to grain boundary cavitation. The finer mesh is obtained by halving the grid spacing in each spatial direction. There is clearly very little mesh dependence for the crack growth rate. The features of the problem formulation and the numerical procedure that permit such mesh-independent results to be obtained need to be understood better. (The highly viscous nature of creep response is one such feature.)

It is worth noting that these computational models have two spatial dimensions while the actual phenomena are three-dimensional. The reason is simply that the computer time and storage required for fully three-dimensional analyses would preclude, or at least severely limit, the parameter studies that the investigations require. Clearly, such studies would benefit from improved algorithms for path-dependent material response that are flexible enough to accommodate physically based constitutive descriptions, instabilities, and progressive creation of new free surface. In order to transform these types of computations from research projects into engineering tools, significant advances in computational capability and understanding are required.

Brief descriptions are given below of areas where improved computational mechanics capabilities appear likely to have a real impact on the design, processing, and performance evaluation of new materials. The focus is on issues that are especially material related, and in this area computational and modeling issues often are not readily decoupled; an accurate predictive capability depends on an accurate physical model. Although purely modeling issues will not be discussed, a key need for further progress is the development of accurate constitutive models for large strain inelastic deformation that are suitable for use in computations. Also, general computational issues, while important, are not highlighted.

THE LOCALIZATION OF DEFORMATION

When ductile solids are deformed sufficiently far into the plastic range, it is frequently observed that a smoothly

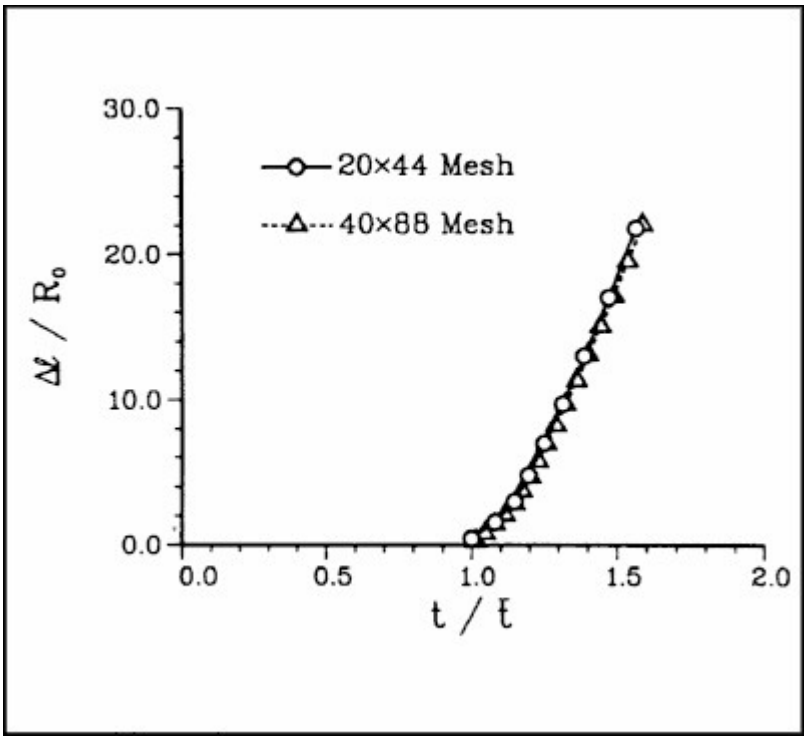

Figure 6.5
Creep crack growth versus time for two meshes. Source: F.Z. Li, A. Needleman, and C.F. Shih, 1988, *International Journal of Fractures* , Vol. 36, pp. 163–186. Reprinted with permission, Kluwer Academic Publishers, Amsterdam, The Netherlands.

varying deformation pattern gives way to one involving highly localized bands of deformation. This phenomenon of flow localization occurs in a wide variety of solids: ductile single crystals, polycrystalline structural metals, polymers, granular materials, and concrete. Such localizations are often termed shear bands because the deformation mode in the localized band is primarily one of shear. Shear bands have a dual significance: as a precursor to fracture and as a mechanism for further deformation. The ability to predict their onset, evolution, and subsequent failure is essential for efficient control of a variety of manufacturing processes and design of material microstructures.

Problems involving shear localization present special complications for numerical methods in general and finite element methods in particular. For rate-independent solids, a shear band is associated with a change in the character of the governing equations. Under quasi-static loading conditions, the equations governing incremental equilibrium lose ellipticity, while under dynamic loading conditions, wave speeds become imaginary. Since the governing equations have no natural length scale, the width of the shear band is arbitrary. Hence, numerical solutions to localization problems can exhibit an inherent mesh dependence because in a grid-based numerical solution the minimum width of the band of localized deformation is set by the mesh spacing. Furthermore, global quantities such as the overall stiffness characteristics of the body depend on the mesh size used to resolve the band of localized deformations.

Pathological mesh dependence can be eliminated by a "regularization" of the governing equations. This regularization can take a variety of forms, including accounting for material rate sensitivity or introducing a nonclassical strain gradient modification to the constitutive relation. The regularization needs to be understood and justified on physical grounds. Different regularizations present different numerical challenges, and the implications of these for numerical accuracy, efficiency, and stability need to be understood. There are also important mathematical questions concerning the existence, regularity, and stability of solutions to localization problems. A more complete understanding of physically appropriate and computationally convenient regularizations is needed.

However, regardless of the formulation, the mesh sets the minimum band width at one grid spacing. Conventional displacement finite element methods can only resolve bands of concentrated deformation when the band interfaces follow element boundaries. Unless special care is used to align

element boundaries along band directions, the mesh introduces a shear band broadening that can control the course of strain localization and mask the effects of physical response mechanisms on localization. At present, the onset of localization and the orientation of localization bands can be predicted accurately if an adequate constitutive model is available for the material. In certain circumstances, including a restriction to one or two spatial dimensions, strain localization development and failure modes can be computed that are in qualitative and, in some cases, quantitative agreement with the experiment. Current computational capabilities need to be extended in order to predict reliably general localization and failure phenomena. Recently, much progress has been made on developing enhanced and enriched element formulations that permit resolution of sharp gradients oriented arbitrarily with respect to the mesh. These methods open up the possibility of analyzing shear band development in general three-dimensional problems.

Since shear band widths are generally much smaller than the overall body size and develop at locations that are not always known *a priori* , localization problems are ideally suited to benefit from adaptive mesh methods and, perhaps, combined adaptive-enrichment strategies. This has not happened yet, but the potential benefits are great.

PROGRESSIVE CREATION OF NEW FREE SURFACE

Fracture involves the creation of new free surface. In traditional approaches to fracture analysis, the presence of a single dominant flaw is presumed, and the key issue is whether that flaw will grow under the given loading conditions. This question can be addressed by solving a boundary value problem for the flawed body and then imposing an appropriate failure criterion. However, in order to investigate the relation of ductility and strength to material microstructure or to analyze fracture in circumstances where there is no single dominant flaw, an approach is needed where the material's constitutive relation embodies a physically appropriate model of the failure process. In such boundary value problem formulations, failure—the creation of new free surface—is an outcome of the solution procedure. The creation of new free surface is also a desired outcome in many shaping and forming operations. The aim in this type of application is to develop the capability to predict and control how the free surface evolves as process parameters are varied. It is important to note that because of the progressive nature of the separation process the formula

tion must be fully coupled, that is, it must contain a description of the failure mechanism and predict its evolution. An uncoupled calculation where a stress analysis is carried out and then a failure criterion imposed is inadequate for the task.

Several methods can be used to analyze problems involving the creation of new free surface—for example, an "element vanish" technique and a critical energy release rate fracture criterion have been used for ductile and brittle solids, respectively. The critical energy release rate formulation has enabled fundamental questions concerned with the micro-mechanical origin of toughness in ceramics to be addressed by computational modeling. The element vanish technique is applicable when the material constitutive relation allows for a complete loss of stress-carrying capacity and has been used successfully in a wide variety of circumstances where the failure mechanism is ductile void growth. However, other ductile separation mechanisms occur; for example, in machining operations, shear band-type localizations can occur and, the large strains within a band of localized deformation subsequently lead to the creation of new free surface. Quantifying this process is important to understanding when localized deformation leads to separation.

The development of accurate models for the creation of new free surface by a variety of physical mechanisms, and of algorithms for the efficient and reliable implementation of these models is an important need for safety assessment and process modeling for broad classes of materials.

SURFACES, INTERFACES, AND MATERIAL-DEPENDENT LENGTH SCALES

Identification of appropriate material length scales is an issue that arises in a variety of guises. Predictions of fracture necessarily, from dimensional considerations, involve a material length scale. Predicting shear band widths also involves a material length scale. Whenever phenomena are governed by both surface and volume effects, the ratio of material volume to surface area introduces a length scale. In coupled field problems, phenomena such as heat conduction and diffusion introduce length scales. In multiphase materials, it is well known that the absolute size, as well as the volume fraction of the second-phase constituent, plays an important role in determining performance.

There are many important situations where phenomena at surfaces and interfaces play a dominant role. For example, the ductility and toughness of composite materials are often

governed by the bonding between the constituents. Surface quality in machine products and wear under service conditions depend on surface rather than bulk properties. The incorporation of surface characterizations into the formulation often introduces a material length scale as well as allowing for surface-specific phenomena (e.g., slip-stick instabilities).

However, most of what is known about continuum formulations and their numerical implementation pertains to circumstances independent of size scale. An improved understanding is needed of the implications of material length scales for the convergence and efficiency of numerical algorithms.

COUPLED FIELD PROBLEMS

Computational approaches are extremely important in nontraditional areas of application where coupled processes occur between mechanical, chemical, thermal, and electronic phenomena. Examples include hydrogen embrittlement and stress corrosion resistance of materials. Creep at high temperatures typically involves mass diffusion as well as dislocation creep. Thermal variations encountered in material processing can significantly affect properties and performance. In this regard it is worth emphasizing that an understanding of the effects of thermomechanical processing on materials is important even when the ultimate application is not mechanical. Mechanical effects on nonmechanical behavior will undoubtedly become an increasing concern. For example, an important problem in microelectronics, where significant advances have been made by finite element analyses, concerns the prediction of piezoelectrically induced threshold voltage shifts in gallium arsenide metal semiconductor field effect transistors. The development of efficient and robust computational procedures for complex coupled field problems is an area of potentially great impact.

PHYSICALLY BASED INELASTIC CONSTITUTIVE LAWS

Computational mechanics has an important role to play in the development of constitutive laws for new materials through an analysis of the deformation behavior of material microstructures. Modern developments in averaging methods for polycrystalline metals are giving predictions of deformation-induced texture development that are accurate enough to have an impact on industrial processes. Unit cell model studies of composites and solids with various forms of

microdamage (i.e., microvoids and microcracks) are elucidating key features of material deformation and failure modes.

Available phenomenological inelastic constitutive models for composite materials are almost all based on a characterization of the distribution in terms of a single parameter, the volume fraction. Computational studies of irregular distributions are being used to assess the extent to which the overall response of such materials does admit a one-parameter characterization. Recent studies have shown important circumstances where the details of the distribution have a profound effect on overall response. The appropriate characterization of distributions that go beyond the one-parameter volume fraction description for the nonlinear behavior of hetereogeneous materials is an open issue. It seems clear that computational studies will play a major role in identifying appropriate characterization parameters and will be a key ingredient in a yet-to-be-developed extension of current approaches to homogenization for inelastic solids.

7

STRUCTURES AND STRUCTURAL DYNAMICS

Computational structural mechanics and dynamics (CSMD) covers a broad spectrum of activities, including computational methods for predicting the response, failure, and life of structures and structural components and automated methods of structural synthesis and optimization. Also, CSMD is an important component in the multidisciplinary analysis and design of many engineering systems.

RECENT ADVANCES

Major advances in CSMD continue to take place on a broad front. The new advances are manifested by the development of sophisticated computational models that simulate mechanical, thermal, and electromagnetic responses of structures; efficient discretization techniques (viz., improved finite element, boundary element, and hybrid/analytical numerical methods); stochastic-based modeling; and computational strategies and numerical algorithms for nonlinear static, dynamic, and fracture problems. Large software systems now exist for detailed modeling and analysis of structures. These systems are currently being applied to some rather complicated problems, and this is affecting the design cycle of engineering systems and their components.

Life prediction methods are required to perform structural integrity assessments of engineering systems. For metallic structures, empirical methods for life prediction are being replaced by fracture-mechanics-based computational methods. For composite structures, more work is needed to understand the mechanisms of failure initiation and propagation before computational methods can be widely used by industry. Because of the difficulty in understanding and modeling the failure phenomena, computational methods for strength and life predictions are lagging behind those of response predictions.

FUTURE DEVELOPMENTS

CSMD is likely to play a significant role in the future development of structures technology, as well as in the multi-disciplinary design of future engineering systems. For this to happen, major advances and computational tools are needed in a number of key CSMD areas. To this end, the research community must address a number of primary and secondary pacing items, many of which are discussed in this appendix. In identifying the pacing items, three factors are taken into account: characteristics of future engineering systems and their implications on CSMD, future computing environments, and recent and future developments in other fields of computational technology (notably fluid dynamics and computational mathematics) that can be adapted to CSMD.

Primary and Secondary Pacing Items

The primary pacing items in CSMD include detailed modeling of complex structures, prediction and analysis of failure of structural components made of new materials, effective computational strategies for large systems, computational methods for articulated dynamic systems, and quality assessment and control of numerical simulations. (The last pacing item is covered in Appendix 7.) The secondary pacing items include integration of analysis programs into computer-aided design/computer-aided engineering (CAD/CAE) systems and predata-postdata processing and the effective use of visualization technology. (The secondary pacing items are not discussed in this appendix.)

Detailed Modeling of Complex Structures

One of the most important steps for the accurate prediction of the response of a complex aerospace structure is the proper selection and sequencing of mathematical and discrete models with varying degrees of complexity. Hence, there is a need for development of automatic model generation facilities as well as smart interfaces to the analysis and design systems. The smart interfaces will be artificial intelligence (AI) based expert systems that run on workstations and can help the engineer in the initial selection of the model, its adaptive refinement, selection of the solution procedure, constraint representation, and interpretation of the results.

Work is currently being done by AI researchers on the development of intelligent computational tools that combine numerical techniques with traditional AI methods, particularly symbolic methods and computer vision. These tools are designed for automatic preparation, execution, and monitoring of numerical simulations and automatic interpretation of their results. These new tools have high application potential in many facets of CSMD, including preparation of a discrete model from high-level specification in symbolic form of structural characteristics and qualitative description of structural response characteristics using computer vision techniques to recognize them (see Appendix 4).

Predictions of Failure of Structural Components

Practical numerical techniques are needed for predicting the failure initiation and propagation in structural components made of new high-performance materials within terms of measurable and controllable parameters. Examples of these materials are high-temperature materials for hypersonic vehicles; piezoelectric composites; and electronic, optical, and smart material systems for space and other applications. For some of the materials, accurate constitutive descriptions, failure criteria, and damage theories are needed, along with more realistic characterization of interface phenomena (such as contact and friction). The constitutive descriptions may require investigations at the microstructure level or even the atomic level, as well as carefully designed and conducted experiments. (Numerical modeling of new materials is covered in Appendix 6.) Numerical simulation of failure is still a challenge and is feasible only under restricting assumptions. Considerable work is needed in this area.

Solution of Large-Scale Structural Problems

A number of large-scale problems exist for which solutions are not feasible, even on present-day large computers. Examples of these problems are dynamics of large flexible structures incorporating the effects of joint nonlinearities and hysteretic damping, interaction problems of large structures with harsh operational environments such as space structures for extraterrestrial bases, and large-scale multidisciplinary design problems. Solution of these problems requires the development of effective multilevel strategies and hierarchical modeling techniques. Promising multilevel strategies include

hybrid modeling/analysis techniques and partitioning methods. In hybrid modeling/analysis techniques, different analytical, numerical, and experimental techniques are combined to predict the response of the structure. Partitioning methods are based on the intuitively obvious and well-established practice of breaking a larger problem into smaller subproblems and generating the solution with information provided by the individual subproblems. Hierarchical modeling is a conceptual change from traditional finite element modeling. It is a strategy to handle structural response characteristics covering a wide variety of length scales (e.g., from global stress/displacement distribution in a composite component to stress/strain distribution in the fibers and the matrix). It can be performed by an initial selection of multiple mathematical models in different regions of the engineering system to take advantage of the simplifications available (through the known local nature of the response). The mathematical models used can range from simple one-dimensional continuum models to three-dimensional micromechanical models followed by adaptive refinement of the models for those regions in which the sensitivity of the response to the modeling details neglected exceeds a prescribed limit.

Computational Methods for Articulated Dynamical Systems

Broad classes of machines that function in space and on the earth's surface involve coupled effects of deformation, angular velocity, and acceleration relative to an inertial reference frame. Deformation may be thought of as relative to an accelerating frame that is associated with each moving component of the machine. Such applications include deployable space structures, orientation control of space systems, space and surface-bound robot and manipulator, ground vehicles such as cars and trucks, aircraft landing gear, and manufacturing equipment. Such machines are designed to control motion and transmit force in the presence of large relative motion between components.

Despite the breadth of these applications and their importance in many fields of engineering and applied science, dynamic simulation methods to predict motion and structural deformation of the components of such systems have lagged significantly behind the much better developed field of finite element structural analysis. While considerable progress has recently been made in computational methods for articulated structural dynamics (often called multibody dynamics or flexible multibody dynamics), much remains to be done in the

development of algorithms and computer software for automated formulation and solution of the differential algebraic equations of motion that describe the dynamics of articulated structures. Such simulation tools must be capable of accurately predicting coupling among bodies that make up the system, coupling between gross motion (angular velocity and acceleration) of individual bodies and their associated deformation fields, and prediction of constraint forces that act between components. Only then can effective performance and failure analysis be carried out prior to fabrication and hardware testing.

Another major objective of articulated structural dynamics is to provide accurate and effective simulation of dynamics and control of hydraulic subsystems that are used to impart desired motion to the system. Structural-control interaction is becoming increasingly important for lightweight extremely flexible articulated structures whose motion is controlled by feedback systems that can cause unwanted interaction and even instability. Finally, with the advent of space telerobotic systems, interaction among the human operator, the control system, and the flexible articulated structure must be accounted for. This will require real-time simulation of the articulated structure and its associated controls, in order to carry out experiments with the operator in the loop.

8

NONLINEAR EQUATIONS AND BIFURCATION

Finite-dimensional nonlinear systems of equations have been encountered in mechanics ever since the field's formalization in Newton's time. It is well known that, except for very simple cases, there are no general methods for obtaining exact solutions of such systems. This has been one of the reasons for the extensive use of linearized mechanical models throughout history. However, in recent years and with the development of electronic computers, there has been a growing trend toward extending the computational methodology to nonlinear problems.

Nonlinearities pervade nearly all parts of mechanics, and accordingly nonlinear systems of equations are found in connection with a multitude of problems. Some of the principal examples include systems arising (1) as discretizations of nonlinear equilibrium equations, (2) in connection with nonlinear optimization problems, (3) in implicit methods for the solution of dynamic processes, and (4) in the formulation of the nonlinear dynamics of constrained multibody systems.

The form and properties of these nonlinear systems are strongly affected by the problem area as well as by the different sources of nonlinearities encountered in most types of problems. The latter point can be illustrated by some typical examples in structural mechanics, including (1) geometrical nonlinearities due to nonlinear strain-displacement relations, (2) material nonlinearities due to nonlinear constitutive equations, (3) force nonlinearities due to nonlinear stress boundary conditions, and (4) kinematic constraint nonlinearities due to nonlinear displacement boundary conditions.

The choice and effectiveness of a solution technique for any nonlinear equation depend critically on factors such as (1) the form and properties of the equations, (2) the specific aims and accuracy requirements of the computation, (3) the amount of available information about the problem, and (4) the computing resources to be used.

The theoretical understanding of these and related factors has not reached, by far, a satisfactory stage. Many techniques and computer programs have been developed for

various types of problems, but their mathematical basis is often insufficiently understood, and there is rarely much known about the accuracy of the computed results.

Deficiencies in the numerical analysis of many problems are often directly attributable to an insufficient theoretical basis for the problems themselves. This is the case, for instance, in finite deformation plasticity where there are profound mathematical and computational difficulties in modeling phase changes, viscous effects, cracks, singularities, etc. At the other end of the spectrum, there are many novel questions relating to scientific computing, such as those concerning the effect of different types of computational resources and, especially, the development of new computer architectures on the selection, implementation, and performance of different numerical methods.

In line with the above remarks, it must be noted that the following discussion will touch only a few of the numerous open research topics in this area.

ITERATIVE TECHNIQUES

Mathematical theory shows that a system of n nonlinear equations in n unknowns

$$f_i(x_1, \ldots, x_n) = 0, \; i = 1, \ldots, n \quad (8.1)$$

may have many solutions. In fact, there exists a theorem to the effect that for certain systems any closed subset of $\mathbf{R}^n$ can occur as the solution set. This also implies that any numerical algorithm must be based on specific *a priori* knowledge about the existence and type of solution that is to be computed. It also means that there cannot be any best algorithm for general-purpose applications.

In practice, a basic task relating to a system (8.1) is the development of locally convergent iterative processes for the computation of an isolated solution x^*. These are processes for which there exists an open neighborhood U of x^* such that for any starting point(s) in U the computed iterates do not leave U and converge to x^*.

The literature on such locally convergent methods is certainly huge, but there are still numerous open problems, especially when it comes to practical applications. Such questions as storage management, data access, vectorization, and parallel processing have a profound influence on the overall procedure, especially when the dimension n is large.

For large n it becomes necessary to consider the use of secondary iterative techniques for solution of the linear systems at each step. Then effective control of the combined primary and secondary iterations assumes central importance. Some challenging questions of this type arise when, as a secondary process, a multigrid technique is applied that utilizes the fact that the nonlinear system is the discretization of some differential equation.

In connection with locally convergent iterative processes, it must be noted that the attraction region for a specific solution x^* is anything but a simply shaped set. In fact, while it is usually known that this region contains some ball around x^*, the space between the attraction balls belonging to different solutions of (8.1) generally has a fractal nature. Since the size of the attraction balls is rarely accessible, this makes early detection of suspected divergence of any iterative process for (8.1) a critical issue for the overall computational cost.

Some general problems concerning any such iterative process for (8.1) are the design of effective and reliable tests for termination of the process either when, as mentioned, divergence is suspected or when convergence within given error tolerances can be declared.

Various approaches have been proposed for control of round-off errors. In particular, interval arithmetic has been shown to be capable of providing excellent guarantees for computed solutions of nonlinear equations. But at present these methods can be expected, at best, to become applicable for the large problems arising in computational mechanics only when the needed operations are available in hardware.

PARAMETERIZED PROBLEMS

In general, nonlinear problems in computational mechanics involve a number of parameters that influence the behavior of the physical system and that have to be taken into account in the computation. For example, in structural mechanics these parameters may characterize intrinsic quantities such as material properties, geometric dimensions, or extrinsic influences such as load points, directions, and intensities. Thus, the system of nonlinear equations (8.1) has. the more general form

$$f_i(x_1, \dots, x_n, \lambda_1, \dots, \lambda_m) = 0,\ i = 1, \dots, n \quad (8.2)$$

involving a vector of parameters $\lambda = (\lambda_1, \dots, \lambda_m)^T \in \mathbf{R}^m$.

In general, it is rarely desired to compute only some specific state vector $x = (x_1, \ldots, x_n)$ which, together with a specific choice of the parameter vector , solves (8.2). Instead, interest centers on assessing the changes of the state vectors when the parameters are varied and, in particular, determining parameter values where a change in the physical solution behavior is expected, such as a loss of stability for certain equilibrium problems.

Under natural conditions, the set of all solutions

$$M = \{(x,), f_i(x,) = 0, i = 1, \ldots, n\} \quad (8.3)$$

forms a submanifold of dimension m in the product space $\mathbf{R}^n \times \mathbf{R}^m$ of the state and parameter variables. Then the question is to determine the shape and properties of this solution manifold.

The classical technique for analyzing the solution manifold (8.2) of (8.3) is the continuation method. It consists in the *a priori* selection of some path on the manifold M and in computing a sequence of points along that path starting from some known point. In practice, selection of the continuation path usually consists of specification of the parameters as some function i = i(μ), $i = 1, \ldots m$ of a single variable (μ). The solution step from a current point then consists of determining an appropriate predicted point along the path and application of a local iterative procedure to a suitable augmented nonlinear system $n + 1$ equations in as many unknowns. Various forms of continuation processes have been proposed that differ in choice of the predicted point, form of the augmented equations, and type of local iterative process.

Thus, all the errors for local iterative processes arise also in connection with continuation methods. In addition, it becomes important how to determine accurately the required "singular points" on the path where the solution behavior is expected to change. In this connection it is also critical to decide when the process leaves a particular component of the solution path. These questions of continuation method accuracy are as yet largely unresolved.

For a numerical analysis of a given solution manifold, it appears to be advantageous to consider continuation methods in a broader sense as a collection of numerical procedures for a variety of tasks, such as (1) following numerically any curve on the manifold and providing estimates of the accuracy of the computed points; (2) on any such curve determining the location of target points where a given variable has a specified

value and, if desired, at such a target point switching to the trace of a different *a priori* specified path; and (3) on any solution path determining the location and properties of the singular points where the solution behavior changes and, depending on the type of the detected point, taking appropriate action.

For specific applications, additional tasks may arise. For instance, if (8.3) represents the equilibrium equations of a system of differential equations, there may be interest in locating Hopf bifurcation points on a solution path where periodic solutions of the dynamic system branch off from the static equilibrium.

In recent years much research has been devoted to these various tasks, but numerous questions remain. Library-quality software is available only for some tasks. For specific applications, such as fluid dynamics problems or structural problems involving plasticity, creep, and viscoelastic effects, the situation is even more open.

The continuation processes, and more generally the above tasks, mean that a picture of an m-dimensional manifold in $n + m$ space is to be developed solely from computed information along *a priori* specified paths. But, in many problems in mechanics, it is essential to assess the sensitivity of a solution (x,) of (8.2) under potential changes of several parameters. In other words, the interest is to determine the curvature behavior of the manifold M at the particular point. For this purpose, in recent years, some methods have been proposed that allow for computation of triangulations of the manifold in the neighborhood of a given point and for subsequent determination of some measure of the curvature. These multiparameter methods were in their infancy, but are no longer so. In particular, they have not yet been implemented in production codes nor applied to large practical problems.

BIFURCATION WITH OR WITHOUT SYMMETRY

As noted earlier, for a deeper understanding of the behavior of a mechanical system, it is necessary to analyze the shape of the full solution manifold and, in particular, to determine the singular points where the solution behavior is expected to change. Here singularity may be characterized by the condition that the system's Jacobian matrix is not subjective with respect to the stated variable. Of course, in this context the choice of the parameters entering into the definitions of the equations is critically important.

Typical computational tasks arising in connection with bifurcation phenomena include accurate determination of the singular points along paths. Usually, this is accomplished via a characterization of the expected singular point as a solution of an augmented system. In their present form the available iterative methods for such augmented equations are local in nature and often require for their convergence the availability of exceptionally close initial guesses. Moreover, currently used augmentations generally assume knowledge of the dimension of the null space. It would be of practical importance to design improved algorithms where both these constraints are relaxed as much as possible.

Calculation of the Bifurcated Branches

Singular points along a path are the only ones where several solution branches may intersect and, hence, where multiple options arise for the path continuation. Different branches typically correspond to different physical behavior of the underlying mechanical system, although some branches may have little or no physical reality. However, it is generally necessary to calculate parts of a solution branch before it becomes possible to decide its physical relevance. While the cases of simple turning points and simple bifurcation points have been treated exhaustively in the theoretical and numerical literature, there is as yet little material available regarding calculation of all the branches emanating from multiple bifurcation points.

Calculation of Paths of Singular Points

Singularity theory has provided some rigorous basis for the intuitive idea that a hierarchy exists among singularities. This in turn has given some substance to the statement that a singular point is more or less singular than another one. Singularities of high rank in the hierarchy, especially organizing centers, are of practical interest because they admit a large number of qualitatively nonequivalent perturbations, and hence their analysis may reveal unsuspected phenomena. While both the dimension n of the state variable and the number m of the parameters are important factors for a given type of singularity to be generically present in an equation system, a singularity of high rank can hardly be detected by following a few randomly chosen paths. In fact, high-rank singularities are rather scarce and, hence, lie only on selected

and *a priori* unknown solution paths. In other words, detection of such singularities requires special strategies, such as following paths of singular points along which higher singularities may be found. In practice, paths of singular points may be obtained by solving the original system augmented with additional relations that ensure singularity. This technique has been implemented in some of the simpler cases, but a satisfactory general procedure is not available at this time.

It should be pointed out that one ingredient entering the solution of all three problems mentioned above is the availability of cost-effective and accurate techniques for the numerical evaluation of higher-order mapping derivatives. Some recent methods for calculating the curvature of the solution manifold have also provided interesting possibilities for the numerical calculation of second-order derivatives, but further contributions to this area are certainly needed.

Bifurcations are often closely connected with a loss or "breaking" of symmetry. Several problems, both theoretical and numerical, are worthy of examination. Certainly, it is important to be able to predict as accurately as possible what symmetries will be preserved along some bifurcated branches when a bifurcation point is encountered along a path of "full" symmetry. For this, bifurcation theory provides a certain number of criteria but by no means a comprehensive answer. On the positive side, it may be noted that the case of the dihedral or orthogonal groups and their subgroups governing symmetry in many problems of mechanics is now much better understood than the general case.

Among the many questions related to numerical aspects is, for instance, the preservation of symmetries of the original "continuous" (e.g. partial differential equation) version of the problem. Clearly, symmetry can be utilized in the computation only if the discretized equations have inherited at least part of the original problem's symmetry. This often requires specific constraints in the mesh design, etc., that can be met, in practice, only with some difficulties. Another important question concerns numerical identification of the invariant subspaces and, in finite element methods, of suitable bases of such subspaces. Here it may be noted that the required symmetry properties of the basis functions may significantly enlarge the size of their supports, which in turn may have a negative effect on the sparsity of the matrices involved in the calculations. It has not yet been examined how much this inconvenience balances the gain in the reduction of the dimension of the state space or whether the symmetries induce properties in

the matrices that may be exploited to make up for the deterioration of their sparsity.

Finally, it has to be emphasized that all computations can work only with singular points on the solution manifold of some finite dimensional nonlinear equations (8.2). Since most of these equations in computational mechanics represent discretizations of certain differential equations, the question arises whether any computed critical phenomena, such as a particular bifurcation point and its branching behavior, correspond to similar phenomena of the original problem and, if so, what errors are encountered. These questions have as yet only partial answers and represent considerable research challenges. In this connection it must be noted that for nonlinear problems the solution manifold of the discretized equations often shows different features from that of the original equations. In particular, there may be connected components of the approximate manifold that do not occur in the exact solution manifold. Such components have been called spurious and are being observed more and more often in various applications, but their study has only recently begun.

This entire range of questions about the discretization errors of multiparameter nonlinear problems is only in a very embryonic stage. The same is true for the design of efficient and reliable *a posteriori* error estimates, although results for certain specific classes of problems and discretizations have shown that estimators for the discretization errors at computed points on the solution manifold are computationally feasible. There is certainly a need for concentrated research in this area.

DIFFERENTIAL ALGEBRAIC EQUATIONS

Differential algebraic equations (DAEs) may be characterized as problems of the form

$$F(t,x,x') = 0, \quad (8.4)$$

where $F : \mathbf{R} \times \mathbf{R}^n \times \mathbf{R}^n$ ` $\mathbf{R}^n$ and the derivative D_pF with respect to the third variable has constant but not full rank $r < n$. Thus, in contrast to the case $r = n$, equation (8.4) cannot be transformed into an explicit ordinary differential equation (ODE) in $\mathbf{R}^n$ in the vicinity of any given (t_0, x_0, p_0) with $F(t_0, x_0, p_0) = 0$.

An important class of mechanical examples involving DAEs arises in the study of the dynamics of constrained multibody systems. Inherent here is the need to account for

algebraic kinematic constraints due to mechanical couplings that restrict relative and absolute motion of the components in a multibody system. Thus, if $x \quad \mathbf{R}_n$ represents a vector of generalized coordinates, the resulting DAE consists of the algebraic constraint equations

$$(x,t) = 0 \quad (8.5)$$

together with the constrained equations of motion

$$\mathrm{M}(x,t)x'' + \quad \mathrm{x}(x,t)\mathrm{T}z = Q(x,x',t) \quad (8.6)$$

where M denotes the mass matrix, z is the vector of Lagrange multipliers, and Q is the vector of generalized applied forces.

The formulation (8.5) – (8.6) has been known for centuries, but it was only recognized during the past decade that DAEs have different properties than ODEs and that their numerical solution involves particular difficulties. While there has been considerable progress in the development of specific computational methods for DAEs in general, at present there exist no broadly applicable and robust integrators suitable for automated application of multibody dynamic simulation methods in a computer-aided engineering environment. Recent developments of methods of generalized coordinate partitioning methods for local parametrization of the kinematic constraints appear to offer particular promise for the effective solution of (8.5) – (8.6), but much still remains to be done.

Other recent advances have shown that a satisfactory existence and uniqueness theory for initial value problems associated with DAEs can be worked out, which extends the standard ODE theory. By showing how DAEs locally reduce to ODEs on implicitly defined manifolds, these results suggest new techniques of solution and identify new challenges. For instance, it appears that the design of higher-order schemes for the treatment of ODEs on manifolds may have to include curvature considerations.

SINGULAR DIFFERENTIAL EQUATIONS

Singular differential equations differ from DAEs in that noninvertibility of the derivative D_pF in (8.4) occurs only at the points of a set with an empty interior. In other words, in a singular ODE, singularity may be termed an "accident." Such equations occur in a variety of physical problems, including, for instance, plasticity theory. They have been studied much less than DAEs despite their first appearance in the scalar case

$n = 1$ well over a hundred years ago. In fact, a rigorous analysis of singular ODEs in the general case has begun only recently but has already shown that singularities may be responsible for some quite peculiar dynamic behavior. Much theoretical and numerical work remains to be done in this little explored area.

9

UNCERTAINTY AND STOCHASTIC PROCESSES IN MECHANICS

Probabilistic computational mechanics is the methodology that forms the basis of the structure reliability and risk analysis of mechanical components and systems. Reliability and risk analyses are of critical importance both to public safety and the competitiveness of products manufactured in the United States.

Reliability analysis applications to enhance public safety include the performance of structures when subjected to seismic loads, determination of inspection cycles for aging aircraft, and evaluation of existing infrastructure such as bridges and lifelines. In the design of mechanical components and systems where safety is not a crucial issue, reliability engineering is also important because it can provide cost-effective methods for enhanced fabrication and inspection.

The problem common in reliability engineering is that certain features of the problem are uncertain or stochastic in character. Two of the most important sources of uncertainty in reliability engineering are unavoidable defects in the structures such as cracks and the environment, which includes factors such as load and temperature.

PROBABILISTIC FRACTURE MECHANICS

Cracks, whose behavior is described by the field of fracture mechanics, are one of the most pervasive causes of failure and therefore play a critical role in reliability engineering. Many of the problems of aging structures and aircraft, component life, and behavior under extreme loads are due to the growth of minor defects into major cracks. The growth of cracks is, however, an inherently stochastic process. Both the sizes and locations of the initial defects that lead to major cracks are random, and the growth of a crack under cyclic loading is stochastic in character. Generally, the growth of a crack under cyclic loading is modeled by the Paris law, in which the length of the crack a is governed by where n is the number of load cycles, and K is the range of the stress intensity factor in the load cycle; D and m are constants that are

fit to experimental data and exhibit significant scatter, or randomness.

$$\frac{da}{dn} = D(\Delta K)m \qquad (9.1)$$

In current engineering practice the reliability of a structure against excessive crack growth is usually ascertained by performing linear stress analysis and then using *S-n* charts, which provide the engineer with the probability of failure due to fatigue fracture of a component subjected to *n* cycles to a maximum stress *S*. These *S-n* charts are usually based on a simple rod specimen subjected to a uniaxial state of stress, which may be quite different from the complex stress pattern encountered in an actual component. Furthermore, the assumption of a perfect cyclic character with an amplitude that does not vary with time is usually quite unrealistic.

Computational mechanics is now reaching the stage where the actual growth of cracks in structures can be modeled along with the uncertainties in the crack growth law, initial flaw size, and randomness in loading. These methods can be based on Monte Carlo procedures; however, they are often expensive in terms of computer cost. Alternatively, the approximations of first-and second-order moment methods may be used. As described later, the latter may not be of sufficient accuracy in cases where the underlying problem is strongly nonlinear. To make these advances useful to engineers, better methods and an improved understanding of the limitations of available methods for these problems is needed.

The stochasticity in parameters *D* and *m* in the Paris law is probably due to randomness in the strength or toughness of the material and the randomness of the microstructure of the material. These ideas have been examined only very cursorily. A better understanding and methodologies for treating these problems are urgently needed for the following reasons: The development of the Paris law data involves many tests, which are often not feasible when advanced, high-cost materials are considered, and the Paris law is directly applicable only to mode I crack growth (crack growth under tension) and is not applicable to cracks that do not remain rectilinear, as in the presence of shear or in three-dimensional crack models.

By computational studies of the stochastic character of materials and their failure in conjunction with experiments, it may be possible to develop more generally applicable crack growth laws.

The implications of such improved computational mechanics methodologies are quite startling. It would be

possible to relate lifetimes of components to the size and distribution of defects that are introduced in the fabrication process and, thus, design fabrication processes and optimal cost effectiveness. Inspection cycle and nondestructive evaluation techniques for structures such as bridges, pipelines, and aircraft could be optimized for reliability and cost.

UNCERTAINTY AND RANDOMNESS IN LOADS

Loads are the second major source of uncertainty in reliability analysis. Loads, man-made or natural, acting on mechanical and structural systems are often difficult to predict in terms of their time of occurrence, duration, and intensity. The temporal and spatial load characteristics needed for detailed analysis are also subject to considerable uncertainty. Nowhere in the engineering field does this fact manifest itself more strongly than in earthquake engineering. In view of this, the uncertainty issues associated with earthquake engineering, particularly with earthquake ground accelerations as loads to mechanical and structural systems, are used as an example to demonstrate the complexity of the problem associated with the uncertainty in loading conditions.

There are many ways in which strong-motion earthquake phenomena can be modeled from the engineering point of view. Each model consists of a number of component models that address themselves to particular phenomena of seismic events. For example, a succession of earthquake arrival times at a site may be modeled as a stationary or nonstationary Poisson process, and the duration of significant ground motion in each earthquake may be modeled as a random variable with its distribution function specified. Also, temporal and spatial ground-motion characteristics may be idealized as a trivariate and three-dimensional nonstationary and nonhomogeneous stochastic wave with appropriate definitions of intensity. Although further study is definitely needed, the progress made in this area has been rather remarkable. Some of the current models are able to reflect the randomness in the seismic source mechanism, propagation path, and surface layer soil amplification.

The difference in ground motion and resulting structural response estimates arising from the use of various models represents modeling as well as parametric uncertainties, since each component model contains a certain number of parameters to which appropriate values must be assigned for numerical analysis. Hence, the total uncertainty consists of modeling uncertainty and parametric uncertainty.

In fact, a number of methods are available and have been used to identify the extent of uncertainty of parametric origin. The process for modeling uncertainty appears to be limited by the extent of the plausible models that can be constructed and the ability to examine the variability of the results from these different models. The degree of variability expressed in terms of range or any other meaningful quantity may be seen as representative of modeling error when, for example, "best estimates" are used for parameters within each model.

SYSTEM STOCHASTICITY

The last several years have seen a resurgence of research interest in the area of system stochasticity. The problem of system stochasticity arises when, among other things, the stochastic variability of the system parameters must be taken into consideration for evaluation of the system reliability under specified loading conditions.

Indeed, the parameters that control the constitutive behavior crack growth and strength of the material tend to be intrinsically random and/or uncertain due to a lack of knowledge. The stochastic variability of these parameters is idealized in terms of stochastic fields, multivariate and multidimensional as appropriate, for continuous systems and by means of a multivariate random variable for discretized systems.

The resurgence of interest appears to have arrived at a time when the finite element method has finally reached its maturity, so that the finite element solution to the problem of system stochasticity can augment existing software packages and thus provide added value. The recent effort in this direction has led to the establishment of a *genre nouveau*, "stochastic finite elements." However, many important issues remain to be addressed. In fact, it was only recently that the basic accuracy and convergence issue arising from the various methods of approximation was addressed in the context of Neuman expansion or Born approximation, primarily when dealing with static problems. Not only that, but the issue of stochastic shape functions has never really been resolved.

From the purely technical point of view, the subsequent comments seem in order with respect to stochastic finite element methods. Exact analytic solutions are available only for simple structures subjected to static loads. Mean-centered perturbation methods are the most widely used, accurate only for small values of variability of the stochastic properties of

the system and inadequate to deal with nonlinear and/or dynamic problems. Solutions based on the variability response function are accurate only for small values of variability of the stochastic properties of the system and inadequate to deal with nonlinear and/or dynamic problems. In the analysis of response variability arising from system stochasticity, however, introduction of the variability response function has provided conceptual and practical novelty. The same analytical procedure can be used as in random vibration analysis where the response variance is obtained from the integral over the frequency of the frequency response function squared multiplied by the spectral density function of the stationary random input function. Monte Carlo simulation techniques are accurate for any variability value of the system's stochastic properties, applicable to nonlinear and dynamic problems, less time consuming, and more efficient in static problems if than Neuman expansion methods are used, and are applicable to non-Gaussian fields.

BOUNDING TECHNIQUES

The primary difficulty associated with probabilistic models dealing with intrinsic randomness and other sources of uncertainty often lies in the fact that a number, for that matter usually a large number, of assumptions must be made in relation to the random variables and/or stochastic processes that analytically idealize the behavior of mechanical and structural systems. In this regard the following statement is in order: When uncertainty problems cloud the process of estimating the structural response, the use of bounding techniques permits estimation of the maximum response, which depends on only one or two key parameters of design and analysis. The maximum response thus estimated provides a good idea as to the range of the structural response variability. Although the applicability of bounding techniques is, at this time, limited to less complicated load and structural models, a strong case can be made for the use and further development of this technique.

The bounding techniques indicated here for system stochasticity are of great engineering significance because these bounds can be estimated without knowledge of the spatial autocorrelation function of the stochastic field, which is difficult, if not impossible, to establish experimentally.

10

CHEMICALLY REACTING FLOW AND COMBUSTION

The unique role of computational methods in combustion cannot be overstated. Reacting flow is a hostile environment for experimental measurements and investigations. Even with modern nonintrusive optical measurement techniques, it is very difficult, expensive, and time consuming to obtain field (vs. point) information about the flow and its intrinsic mechanisms. Due to the lack of predictive methods, extensive and very expensive experimentations are currently the primary means of developing new technology from basic concepts. Even fundamental analysis is encumbered by the fact that most methods are actually postdictive. Developing computer-aided design (CAD)/computer-aided manufacturing (CAM)/computer aided engineering (CAE) methods for reacting flow and combustion should lead to similar time and cost savings as those achieved by the proliferation of finite-element methods in structural mechanics.

BASIC MODELING DIFFICULTIES

Progress in the study of reacting flow and combustion has traditionally been difficult due to several factors, not the least of which are the complexity of the governing equations and the multiplicity of their governing processes. Combustion processes, in most cases, involve changes over a wide range of length and time scales (e.g., flames or detonation fronts are very thin compared with other flow scales). Strong interactions exist between these fronts and the overall flow, and solutions including variations of all scales must be obtained. Even in a homogeneous chemical reaction between a simple hydrocarbon fuel and air, hundreds of elementary reactions occurring over time scales that cover orders of magnitudes are encountered. Most frequently, chemical reactions take place while transport processes (convection, diffusion, and radiation) act to change the local composition and temperature at scales different than those corresponding to the reaction.

Most practical combustion applications occur in a turbulent field. With the current lack of universal models of turbulence, even in much simpler nonreacting flows, one is faced with a unique challenge when turbulent combustion is a dominant process in the system. Reacting flows are intrinsically compressible flows with large, highly localized density and temperature variations. At high Mach number, as in supersonic combustion, strong pressure waves are also encountered. In diesel engines, jet engines, some rocket engines, and furnaces, combustion occurs in a turbulent multi-phase flow. In many manufacturing processes, turbulent reacting flow occurs on complex catalytic surfaces. It is no surprise, therefore, that computational methods have, and will continue to play, a significant role in this field. Some of the most important achievements of computational methods are described below in their order of complexity and outstanding challenges, identified by physical phenomena and numerical complexities, are listed. Recommendations are suggested.

RECENT ADVANCES

Some of the most visible achievements of the application of computational methods have been the development of off-the-shelf standard codes to perform some basic and widely used calculations. Computations of the equilibrium composition of a reacting mixture and one-dimensional reacting flow at any given thermodynamic state is one such accomplishment. It is currently used in undergraduate courses all the way to design and analysis in the automotive, aerospace, and utility industries. This was made possible by the development of efficient and accurate methods for the solution of nonlinear systems of algebraic equations. Similar progress has been made in chemical kinetics, which is concerned with homogeneous reacting systems far from thermodynamic equilibrium. In chemical kinetics it is necessary to integrate large systems of stiff, nonlinear, coupled ordinary differential equations. Though not as fast as those used in chemical equilibrium computations, codes are currently available for such applications.

In laminar combustion, where chemical kinetics is coupled with the transport of heat and mass by molecular diffusion, it is now possible, sometimes at a large cost, to compute the detailed structure and speed of propagation of laminar flames. The equations governing this problem are of the steady reaction-diffusion type, and progress in finite difference methods for one-dimensional initial and boundary value problems has allowed for accurate solution of these

systems. Extensions to laminar-reacting boundary layer flows have also been achieved using adaptive one-dimensional grids in two-point boundary value problems with Newton iteration. Attempts to extend such computations to study the interaction between a laminar flame and the field of a single vortex, modeling one component of a reacting flow, are beginning to yield interesting results. However, in this case, models have been limited to simplified chemical kinetics schemes of one or a few reactions.

The picture is not as bright for turbulent combustion, the most frequently encountered in practice. In this case, solutions of the full Navier-Stokes equations coupled with the equations governing chemical kinetics and state are, while mostly unattainable, necessary. However, one cannot ignore the enormous progress in the computation of high Reynolds number nonreacting shear and separating flows. Both flows are fundamental to combustion since they are used in the mixing of reactants prior to their burning or in combustion stabilization in premixed streams. The results of these computations are called simulations since they are based on the unsteady, unaveraged Navier-Stokes equations. They have been instrumental in shedding light on some important instability modes of shear flows, the large-scale structure of turbulent flow, processes that lead to the transfer of energy from the mean flow to turbulence, and the decay of turbulence.

These simulations were made possible by the development of numerical schemes that are highly accurate in temporal and spatial discretizations and stable at high Reynolds numbers. In particular, spectral methods and vortex methods have been particularly successful in flows dominated by highly concentrated regions of vorticity that change their shapes as the flow evolves between different transition modes. Spectral methods have mostly been applied to flows with simple boundary conditions. Recent progress in domain decomposition and global finite element discretization promises to lead to the application of spectral simulation to flows with complex boundaries. Vortex methods, on the other hand, conform to complex domains at the cost of increasing the number of computational elements. The formulation of fast particle solvers and multipole expansion methods has contributed to overcoming the cost problem in vortex methods.

OPEN PROBLEMS AND RESEARCH NEEDS

Applying computational methods for the simulation of turbulent reacting flows has been limited to cases in which the flow's time and length scales and the chemical reaction are of the same order of magnitude. Extensive research is required to extend their applicability to more interesting problems where the length scales vary over a wide range. Methods that rely on multidimensional adaptive gridding, such as adaptive projection methods; Lagrangian methods based on resolving the gradients of the gas dynamic variable, such as vortex and transport element methods; methods based on numerical asymptotic expansions; and hybrid methods that combine the convenience of kernel function discretization with spectrally accurate core functions should all be pursued for this purpose. It is important to mention that the problem of turbulence-combustion interactions—the crux of turbulent combustion— even in its most rudimentary form involving flame-strain and flame-vortex interactions, has not been satisfactorily explored due to the lack of powerful numerical schemes.

Increasing the accuracy and efficiency of time-dependent, multidimensional, adaptive numerical methods is one of the major tasks in the next decade. This includes further exploration of methods based on the Lagrangian formulation of the conservation equations as a natural means of achieving spatial adaptivity. Both grid-based and grid-free Lagrangian methods are natural candidates for this development. Methods that combine the convenience of grids with the accuracy of grid-free schemes, such as those that reduce the dimensionality of the grid by utilizing Green function solutions in the direction where the steepest gradients exit, should be explored further. Finite difference schemes based on embedding solutions of the governing equations at the elementary cell level, such as Godonov methods in which the solution of the Reimann problem is implemented within each cell, offer higher accuracy than conventional methods and should be pursued further.

It is unlikely that all the spatial scales of a turbulent flow, let alone turbulent reacting flow, can be represented in a practical system using stationary or moving grids. Thus, progress in modeling subgrid or subscale effects in a form compatible with the discretization of the resolvable scales will be instrumental in continued development in flow simulation. This includes methods that consider structures of scales smaller than the smallest grid size (e.g., flame fronts) as discontinuities. It is not yet clear how to embed all the physics

of the phenomena that occur within the discontinuity into the formulation of the equations governing its motion or how to implement all the interactions between the discontinuity and the "external" flow (e.g., how the turbulent flow wrinkles a discontinuity and how changes within the discontinuity affect the turbulent flow). Methods based on the fractal representation of these discontinuities have been proposed, but work is still needed to determine the relationship between the fractal properties of turbulized surfaces and the physics of turbulent combustion.

The physics and numerical implementation of small structures of turbulence that cannot be resolved using discretization schemes (e.g., folding hairpin vortices encountered in the interial subrange of turbulence) are currently under investigation using vortex methods. Also, methods are being pursued for representation of the effects of high wavenumber structures in spectral simulations using methods of renormalization group theory. Research on the relationship between the two discretization methods, vortex and spectral methods (particle vs. wavelike representation of the dynamics), is also needed to show ways of optimizing the application of hybrid methods. Wavelet methods offer an opportunity in this regard. Using such statistical methods as probability density function formulation to model high-order moments has also been proposed in connection with solutions of the averaged equations. However, these methods rely heavily on experimental input, and their theoretical foundations are not yet clear.

Other phenomena, which have emerged from engineering applications of combustion and present research opportunities in the coming decade, include supersonic combustion involving shock wave reaction/front turbulence interaction (which is the mode of propulsion in the hypersonic transport vehicles); different modes of combustion instability that encumber the optimization of energy conversion processes (knock in internal combustion engines, chucking in rocket engines, screech in turbojet and ramjet engines); flow and combustion of sprays, slurries, and two-phase media (coal-air, coal-water); combustion of non-Newtonian fluids (which arises in incineration); effects of radiation on combustion dynamics (encountered in, e.g., the formation and destruction of ozone in the atmosphere); and coagulation of heavy particulates (such as carbon) in reacting streams.

Numerical methods in computational fluid dynamics have evolved in two directions: (1) general methods that can be applied to "any" partial differential equation irrespective of its type, such as finite difference and finite element methods, and (2) special methods formulated to obtain solutions for

equations of certain forms that describe a narrow range of phenomena, such as the method of characteristics, vortex, and particle methods. Each class of methods offers some advantages and suffers from certain drawbacks. Progress in both directions has recently led to the formulation of some hybrid methods, especially the coupling of Eulerian and Lagrangian methods, which combine the best of the two approaches; the results have been encouraging. Sample results obtained from such methods will be shown at the end of this chapter. Continued research in the construction and application of hybrid schemes is necessary for the investigation of complex phenomena in reacting flow (e.g., shock/flame-flow/boundary layer interactions that have not yet been satisfactorily simulated). For these problems it is potentially desirable to couple interface-tracking algorithms with grid-based or grid-free solutions of the Navier-Stokes equations.

Hybrid methods, such as combined domain decomposition and interacting particle methods, also permit effective utilization of modern computer architecture (e.g., medium and fine-grain parallel processing in which accurate, but limited, interaction between various parts of the solution domain is necessary in order to take advantage of distributed computing). The availability of many computing environments with different capabilities and characteristics in terms of speed and memory combination, vector length, and number of processors has necessitated, and will continue to necessitate, the development of new algorithms, as well as ways of combining different schemes to take advantage of particular hardware technology. In this regard, advanced graphics software used in pre-and postprocessing of multidimensional, time-dependent solutions represents another important research area that promises to make computational fluid and combustion dynamics more accessible to design engineers. Modern CAD/CAM/CAE systems will have to incorporate such software. Both pre-and postprocessing software will have to be interactive, and the graphics interface may have to be run in real time (i.e., exhibit the solution as it is being computed). Here also, the coupling between the algorithms used to generate the solution of the governing equations (Eulerian or Lagrangian) and those used to interpret them in a way compatible with the intuition of the design engineer (particle trajectories or streamlines) poses another challenge.

POTENTIAL PAYOFF

The need to continue the development of computational methods in combustion becomes more critical when considering the wide range of applications of turbulent reacting flow. The availability of predictive methods will reduce the design and testing time and cost of new products in the automotive, aerospace, electric utilities, and many manufacturing industries. Predictive methods allow one to perform "computer experiments" to investigate the fundamentals and outcome of new processes, thus realizing how clean, efficient, and fast these processes can be without actually building the hardware and running laboratory experiments. Computer experiments can also be used to test different concepts to exercise more effective control over reacting flow systems if they operate at off-design conditions. The continuous improvement of existing products and the introduction of new and innovative products are expected to play a major role in global industrial competitiveness. This can only be achieved via the development of tools to model and analyze the fundamental thermofluid processes that govern the operation of products utilizing combustion processes. For example, as new materials are used in the production of stationary and mobile engines, it will be possible to increase the combustion temperature attained in these engines, which will call for analysis of high-temperature turbulent combustion stabilization and efficiency. Computational methods can substantially limit the time and cost required for such an analysis.

The utilization of computational models to study combustion dynamics is not limited to situations in which there is a need to economize the time and cost. There are cases when one cannot afford to perform experiments or to wait for results in order to reach decisions based on a trial-and-error strategy. Such cases include the aquatic and atmospheric transport of pollutants, their effect on climatic change, and the ecological and biological systems of the earth. Complex chemical reactions accompany this transport and may lead to the formation or destruction of compounds detrimental or necessary for the critical balance that has been maintained over the decades for life to continue and flourish. The depletion of ozone in the upper atmosphere, the formation of acid rain, the release and buildup of formaldehydes, the chemical contamination of soil, ground water supplies and harbors, etc., have been recognized only after years of damage caused partly by a lack of understanding of reacting flow physics. The current effort to reverse the trend by eliminating the sources and cleaning up some of the afflicted areas using newly discovered

processes has been supported by, and should continue to witness, extensive research activities in the field of reacting flow. It is expected, as mentioned above, that this research will benefit extensively and directly from continued developments in computational fluid mechanics and combustion.

Figure 10.1 shows results obtained by a method based on combining the better features of Lagrangian vortex methods and Eulerian finite element methods to study the evolution of a reacting shear layer in three-space dimensions.[9] Besides being a canonical problem in turbulent combustion physics, reacting shear layers pose a severe challenge to computational methods since they possess all the properties that cause the breakdown of simple schemes: time-dependence, changing spatial length scales and overall topology, and the generation of strong strain fields and multiplicity of physical instabilities. Figure 10.1 shows the distortion of the computational grid, which also reveals the evolution of the reaction front as different modes of flow instability grow into their nonlinear stages, and confirms the adaptivity of the method. This adaptivity is crucial for the long time stability and accuracy of the computation at the lowest possible cost.

[9] Knio and Ghoniem, 1990, Journal of Computational Physics, Vol. 86, p. 75.

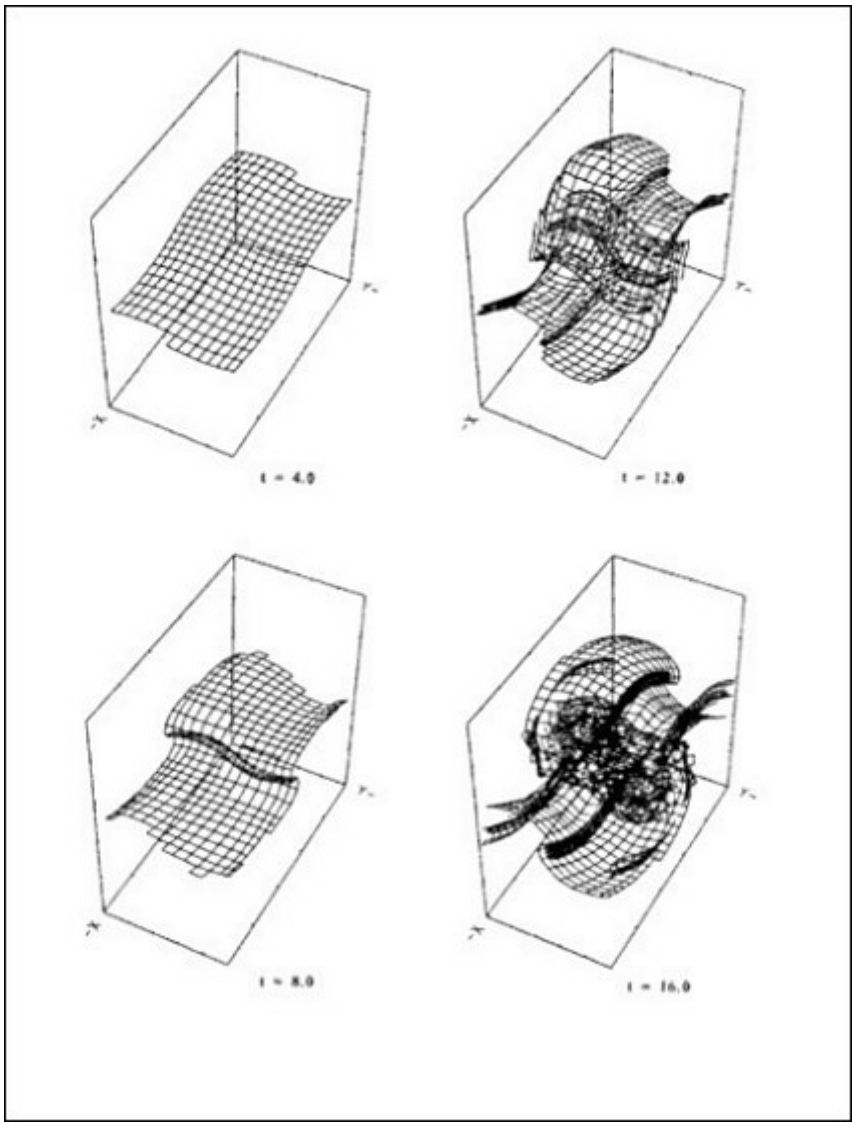

Figure 10.1

11

NUMERICAL MODELING OF TURBULENCE

In volume two of the *Annual Review of Fluid Mechanics* (1970), Professor Howard Emmons reviewed the possibilities for numerical modeling of fluid dynamics and concluded that, "The problem of turbulent flows is still the big holdout. This straight-forward calculation of turbulent flows—necessarily three-dimensional and non-steady—requires a number of numerical operations too great for the foreseeable future." However, within a year of the article's publication, the field of direct numerical simulation of turbulence began with the achievement of wind tunnel flow simulations at moderate Reynolds numbers. In the past 20 years the field has developed remarkably.

The importance of computational turbulence study is its position at the crossroads of advanced technological applications and cutting-edge science. It is an enabling technology in applications ranging from engine design, weather forecasting, and the dynamics of the universe. (Nearly all terrestrial, atmospheric, and marine fluid dynamical processes involve turbulence). Owing to the current knowledge limitations about turbulence problems, research progress is limited for the entire range of flow problems, from many thousands of kilometers to centimeters or smaller scales.

In all fluids (including air, water, and solar gases) the most important physical consequence of turbulence is its enhancement of transport momentum, energy, and particles. A related turbulence feature, which is indeed the fundamental characteristic that makes it so theoretically and computationally difficult, is that it exhibits far more small-scale structure than its nonturbulent counterparts. This small-scale structure is responsible for the enhanced turbulent transport phenomena and is itself evidence of enhanced transport in the sense that small scales develop from the degradation of large-scale excitations that are maintained by energy transport from one scale to another. In the cascade from large to small scales, the nonlinearity of the underlying dynamical equations, the Navier-Stokes equations, plays a pivotal role in mediating interactions among these scales.

One way to measure the effective nonlinearity of the Navier-Stokes equations is by a nondimensional quantity R, called the Reynolds number, defined as $R = UL/\nu$, where U is

a typical velocity change across a typical large length scale L and v is the kinematic viscosity of the fluid. It can easily be shown that the Reynolds number is a typical measure of the ratio of the nonlinear terms in the dynamical equations to the linear (dissipative) term. For typical flows, R is usually quite large. For example, since v ` 0.01 cm^2/s for water, even U = 1 m/s (2 mph) and L = 1 m gives $R = 10^6$.

One of the main results of turbulence theory is that the range of significantly excited scales of motion in both space and time is of the order $R^{3/4}$ Therefore, to calculate a high Reynolds number flow in three space dimensions plus time, it is necessary to perform order $(R^{3/4})^4 = R^3$ computational work. This means that calculating a flow at Reynolds number $2R$ requires roughly 10 times more work than calculating at Reynolds number R!

Another important characteristic of turbulent flows is their apparent randomness and instability in the face of small perturbations, a feature readily noticeable in these complex flows. Two turbulent flows that are at some time nearly identical in detail do not remain so on the time scales of dynamical interest. Instability of turbulent motion is related to the limited predictability of, say, atmospheric motions. The onset character of this randomness (i.e., the transition to turbulence) is a subject of much current interest and involves the study of such interesting dynamical phenomena as those of "strange attractors."

Over the past two decades it has become clear that substantial progress in understanding turbulent flows will require the largest and most powerful computer resources available. Several kinds of computer studies are important.

1. Full numerical solution of the Navier-Stokes equations for turbulent flows to answer fundamental fluid dynamical questions.
2. Numerical tests of theories of turbulence.
3. Numerical tests of turbulent transport approximations for use in large-scale computer models of engineering systems, the ocean, atmosphere, etc. (Here a transport approximation means that only the largest scales of motion are calculated dynamically; all smaller scales are modeled, usually in terms of eddy transport coefficients, like eddy viscosity.)
4. Numerical studies of turbulence dynamics using large-eddy simulations. (In a large-eddy simulation all fluid motions on scales larger than the grid scale in a numerical simulation are calculated in detail, while motions on scales smaller than the grid scale are modeled, often by a turbulence transport approximation.)

5. Studies of the origin of turbulence, including investigations of possible routes leading to chaos (apparent random behavior).

RECENT HISTORY OF TURBULENCE COMPUTATIONS

The first numerical simulations of two-dimensional turbulence were performed on a CDC 6600 computer by Lilly in 1969; the first numerical simulations of three-dimensional turbulence were performed by Orszag and Patterson in 1971 using the CDC 6600 and 7600; the first large-eddy simulation of shear and convective turbulence was performed by Deardorff in 1970; and the first large-eddy simulation of turbulence in a stratified shear flow was performed by Deardorff in 1972. Since then the field has developed so rapidly that for many flows with simple geometry and a moderate Reynolds number, data can now be gathered more cost effectively, more quickly, and more comprehensively by computation than by laboratory experimentation. This feature is especially true in the field of transition to turbulence, where computation has proven to be an exceptionally useful tool. It is expected that this trend will continue and may, indeed, accelerate over the next decade, especially with the advent of new cost-effective parallel computers.

In the following paragraphs, we review the progress on certain specific key problems in turbulence—theory and application—and note the current unsolved issues that numerical simulations have raised. We start with a brief statement on the state of the art circa 1975, we shall point out how, for these particular problems, the next generation of supercomputers, with the reservations and uncertainties noted above, will yield new insights and results. We will then be in a better position to discuss the kinds of development to expect in the 1990s.

In 1975, before the introduction of the Cray-1, state-of-the-art turbulence simulations involved full numerical solutions of the Navier-Stokes equations with up to $32 \times 32 \times 32$ degrees of freedom in three space dimensions and 128×128 degrees of freedom in two dimensions. Low Reynolds number inertial-range (small-scale) dynamics was already studied in two dimensions; however, the study of three-dimensional inertial-range dynamics seemed well beyond the power of existing computers. Numerical studies of thermal convection in two dimensions had been done, as well as some isolated, low-resolution, three-dimensional convection studies. However, there were no systematic studies of the origin of chaotic time

dependence in flows, nor was it clear then that numerical methods would prove definitive in understanding the transition to turbulence. Large-eddy simulations were well on their way, given the pioneering work of Deardorff, but there were no attempts at that time to use the computer to understand the basic fluid dynamics of wall layers in turbulent shear flows. Thus, in 1975, turbulence computations that accounted for all relevant scales of motions were only in their infancy, and scientists were just beginning to perceive the usefulness of such methods to provide answers to fundamental fluid dynamics questions.

With the availability of Cray computers in the late 1970s, studies could be done with computer codes using several orders of magnitude and more resolution than was possible with the CDC 7600. Generally, the CRAY-1 brought a range of three-dimensional turbulence problems within the realm of numerical scrutiny, and it permitted the rudiments of inertial ranges to be defined for simple geometries. For two-dimensional turbulence, considerable progress was made in understanding detailed inertial range effects.

Let us now review some direct numerical simulation studies that have yielded new insights. As yet, most of these problems are not completely solved; they need CRAY 3/MP-level, or higher, computing capabilities before high Reynolds number asymptotics can be ascertained. The survey given here is very limited, but it does highlight some of the range of problems involved.

TAYLOR-GREEN VORTEX, VORTEX RECONNECTION, AND 3-DIMENSIONAL INERTIAL-RANGE DYNAMICS

There are certain simple initial value problems that have yielded much insight into the way in which the nonlinearities in the Navier-Stokes equations lead to ever smaller scales of motion and the rapidity with which this happens. Two problems come immediately to mind in this connection: the Taylor-Green problem, which is a regular system of three-dimensional vortices, and the evolution of a pair of closely spaced, well-defined vortices. Improved resolution on Cray 2-level computers together with new algorithms has allowed an increase in resolution up to 800^3 spatial gridpoints for numerical simulation of the Taylor-Green problem. While the Taylor-Green vortex is a special flow, it appears that the special symmetries it invokes do not inhibit small-scale dynamics in the inertial range. Indeed, in the early 1980s it was for this particular problem that the first direct calculation

of a three-dimensional inertial range was possible. In addition, recent studies have permitted the investigation of intermittency effects of the dissipation-fluctuation spectrum. These computations address very fundamental issues concerning the nature of the effects of nonlinearities in Navier-Stokes dynamics; for example, does frictionless (inviscid) flow stretch vorticity an infinite amount in a finite time? Resolution of the existence of this singularity for the inviscid Euler equations in three-dimensional flow is still undecided, but recent calculations suggest that there may be no singularity. If this is the case, numerical simulations call into question conventional assumptions about the universal nature of high Reynolds number turbulence.

The vortex reconnection problem consists of studying the time development of a pair of vortices that evolve by reconnecting, so that their flow topology changes. This change is induced by the action of viscosity that allows vortex lines to be broken and reconnect. The process has the potential not only to build an understanding of boundary phenomena but to be a mechanism in inertial-range dynamics. So far, resolution 256^3 calculations have demonstrated cleanly the dynamics of such processes and have displayed the subtleties of dissipation (and acoustic noise generation) during the collision process.

TWO-DIMENSIONAL AND QUASI-GEOSTROPHIC FLOWS

Two-dimensional flows are the prototype of the large scales of the atmosphere in which the rotational constraint is of primary consideration. In the past decade, calculations using improved resolution (up to 1024^2) have profoundly affected our understanding of these flows; for example, an initial state of random vortices tends, in time, to become a system of isolated vortices whose dynamics are controlled by their distant interactions with an occasional collisional interaction. Their scale-size distribution is set by their initial distribution. The classical picture of a universal scale-size distribution $E(k) \sim k^{-3}$ has very limited validity at intermediate times. These ideas are being pressed toward the study of quasi-geostrophic flows (fully three-dimensional but still rapidly rotating), and it will be most important to see how much of the strictly two-dimensional picture carries over. Undoubtedly, the next computer generation will do for quasi-geostrophic flows (with various additional effects, such as topography) what the last generation did for strictly two-dimensional turbulence. With respect to meteorology, these

studies assume even more relevance than the two-dimensional studies mentioned above because of the improved realism.

TURBULENT MAGNETOHYDRODYNAMICS

Early studies of turbulent magnetohydrodynamics (MHD) at 64^3 resolution have been extended to 128^3 in the past few years. For the two-dimensional case a considerable inertial range is computable, but for three dimensions the inertial-range effects are unresolved. Currently, calculations have been able to explore issues of anisotropic flows and, in particular, have revealed unsuspected large-scale instabilities of these flows, which indicated that helical flows (flows with spiral structure) are able to organize—under certain conditions—a significant amount of their energy into large-scale structures. Very large computers are necessary here to extend and improve these studies in order to provide the large-scale separation implied by theoretical analysis. Another effect brought under numerical scrutiny is the problem of magnetic field line reconnection (in two dimensions), which is a dissipative process, at extremely small scales. It is analogous to the vortex-reconnection problem but considerably more complicated. New directions in MHD are study of the effects of compressibility and study of three-dimensional field line reconnections. These computations have applications in fusion research, as well as heating of the corona.

LARGE-EDDY SIMULATIONS

Large-eddy simulations (LESs) are of great practical importance. They are calculations in which the effects of scales smaller than the grid scale on those retained in the calculation are statistically modeled. In reality this method is the only hope we have to detail flow models of real-world engineering complexity, and it has wide applications ranging from calculating detailed flows over aircraft wings to global models of the planetary boundary layer. The current state of the art here is 1283 calculations in relatively simple geometries. In the atmospheric sciences, applications include simulation of the full planetary boundary layer, including radiation and condensation. Stratus-topped planetary boundary layers may now be studied via LES simulations, and their input into global climate models will be of great significance in global warming studies. For this latter step, new-generation computers are required. Equally important is the extension of LESs

to boundaries via new algorithms derived from the renormalization group method. Such procedures have the potential to eliminate much uncertainty in the small-scale modeling required for LESs. On the next generation of computers, LESs should be applied to study flows in complex geometries and with complex physics, such as combustion and multiphase flow effects. Other extensions of LESs include such problems as cloud-top entrainment instability mechanisms. Here, fine resolution is essential and, at present, unavailable.

CONVECTIVE FLOWS

Numerical simulation of thermal convection (and stably stratified flows) has proceeded to resolution sufficiently high (192^3) that fully turbulent, high aspect ratio convection may be simulated (at least at moderate Prandtl numbers). Of interest here is not only verification (or repudiation) of traditional scaling laws but also the statistics (distribution function level) of various flow quantities such as vorticity and temperature differences. Here recent theoretical/numerical developments by Sinai and Yakhot are particularly significant. We should note that recent high-precision experiments have shown a sequence of transitions toward progressively "stronger" turbulence through which the flow proceeds as the basic buoyancy driving force is increased. Recent numerical simulations have been able to track these transitions. With increased computing power, it should be possible to explore the asymptotics of this important class of flows.

In summary, John von Neumann's 1949 prediction that computers would prove particularly useful to the study of turbulent flows has come true. Because of still limited theoretical understanding of nonlinear phenomena, engineers, atmospheric scientists, astrophysicists, fluid dynamicists, and others are in great need of computer simulations. To continue recent advances in many aspects of turbulence study, the power and memory of new supercomputers and massively parallel computers are required. In basic research the need for such capabilities will continue until we can achieve resolutions in which asymptotic regimes are manifest, probably not less than 10243 resolution and often much more, depending on the problem.

12

GENERAL COMPUTATIONAL FLUID DYNAMICS

Elsewhere in this report several important areas of computational fluid dynamics (CFD) are discussed in some detail. In particular, aside from general methods and approaches applicable to all areas of computational mechanics (such as adaptive methods, parallel computing, artificial intelligence, nonlinear equation solvers, and stochastic processes), the importance of future research efforts on computer simulation of pollution, chemically reacting flows, combustion, and turbulence is addressed in separate appendices. In this appendix, additional topics on CFD research are discussed that are expected to be crucial to further developments in several key technological areas.

HISTORICAL COMMENTS AND CURRENT STATUS

CFD has always been driven primarily by applications. Aerodynamics, numerical weather prediction, acoustics and fluid-structure interaction, propulsion systems, and nuclear reactor design are among the major applications that have encouraged CFD research.

In the aerodynamics field, particularly aircraft design, CFD was originally regarded merely as a complement to wind tunnel experiments, essentially limited to the simulation of incompressible potential flow. For many years, aircraft-design-oriented CFD was mostly concerned with the solution of potential flow equations by panel methods.

Panel methods were widely used to simulate external flow around aircraft. To evaluate viscous effects, the flow field predicted by those methods was also used as a boundary condition for boundary layer calculations based on some form of the Prandtl equations. This methodology clearly had difficulties with separation and shock phenomena. To simulate such phenomena, it was necessary to use more complicated mathematical models than potential flow, the obvious candidates being the Euler and Navier-Stokes equations. However, with the computers available in the late 1960s and early 1970s,

using Euler or Navier-Stokes models to simulate flow around a full aircraft was impossible, and efforts concentrated on a compressible variant of potential flow embodied in the full potential models of gas dynamics.

During the 1970s upwinded finite difference, finite volume, and finite element methods were developed, leading to the creation of several industrial codes. In addition to upwinding, these codes use multigrid and other fast direct or iterative solvers. By the 1980s transonic flow past a complete aircraft was simulated and used as a production tool in computer design subsonic airplanes with supercritical wings.

During the 1980s the development of compressible Euler solvers, again based on efficient combinations of upwind or artificial viscosity methods, finite element or finite volume approximations, flux-limiting concepts, and multilevel techniques, was an essential step in the simulation of more complicated flows. With these solvers some simulations of three-dimensional inviscid flow around complete aircraft or space vehicles, at reasonably large incidence angles and Mach numbers (including the hypersonic range), have been made. At a recent European/U.S. meeting on CFD, excellent agreement of the results of compressible Euler test cases was reported despite the fact that they were three-dimensional hypersonic flow simulations and that participants were using a large variety of numerical methods. On the other hand, discrepancy among the results of the viscous flow simulations remains large, and three-dimensional unsteady simulations remain out of reach of contemporary CFD capabilities. Additionally, difficulties in modeling pure convection phenomena still exist and researchers point to open questions that remain in developing reliable Euler codes for unsteady three-dimensional flows.

The analysis of propulsion systems, such as the space shuttle main engine and cooling systems for nuclear reactors, has encouraged much research on new efficient incompressible viscous flow simulators. Today codes exist that can simulate three-dimensional flow at Reynolds numbers of the order of several thousand, and several of these codes have turbulent flow simulation options, usually based on k- turbulence models. As yet, there are no viscous flow simulators able to simulate accurately and routinely three-dimensional incompressible and compressible viscous flow at Reynolds numbers greater than 10^5 in complicated geometries. The study of computer simulation of air-and water-borne acoustical phenomena has come to the forefront of computational mechanics research in recent years. When fluid-structure interactions are considered, such as the interaction of submerged elastic struc

tures with acoustical fluids as encountered in modern submarine design, the deficiencies of contemporary methods and mechanics are evident. Significant advances in acoustical simulation techniques are needed to resolve pressing problems in acoustical-structural interactions. To achieve these types of flow simulations within the next decade, research on parallel computing and smart algorithms must accelerate.

The impact of numerical weather prediction on CFD clearly provides a strong motivation to develop better turbulence models and, certainly during the past 15 years, was a practical motivation behind the strong development of spectral and pseudo-spectral methods. More detailed comments are given below on specific areas in CFD that require research advances during the next decade.

COMPUTATIONAL AERODYNAMICS

Modern aircraft, missiles, and reentry vehicles operate under a wide spectrum of conditions. These conditions range from very low speed incompressible flows where small general aviation aircraft operate to the very high Mach number flight regimes of the National Aerospace Plane (NASP) machines such as the aero-assisted orbital transfer vehicles (AOTV). Delineation of the different flight regimes usually proceeds with a comparison between the mean free molecular path and the characteristic length of the flow field. This ratio is the Knudsen number. When the mean free path is much smaller than the characteristic length of flow, the Navier-Stokes equations are considered to be applicable and the fluid is considered to be a continuum. In the short history of computational aerodynamics, the largest research effort has been expended in the continuum regime.

When the Knudsen number is of order one, the flow is said to fall into the slip flow regime. Here the Navier-Stokes equations may not be applicable, although some success in predicting gas flow in this regime has been achieved by solving the Navier-Stokes equations with modified boundary conditions.

When the mean free path is large compared to the characteristic body length, the flow regime is said to be "free molecule." This is an environment, experienced by orbiting vehicles, where molecules are observed as discrete particles. The study of this flow regime is sometimes referred to as superaerodynamics, a name coined in the early literature on free molecule flow.

A major area of concentration in computational aerodynamics research has been the development of better ways to numerically model convection phenomena characterized by the so-called convective terms in the Navier-Stokes equations. Methods have been derived that provide an accurate description of the flow physics where a series of Riemann problems are solved to obtain changes in flow variables in each cell. Central to this approach is the problem of establishing the correct flux terms at cell boundaries. In computing these fluxes, either flux splitting or flux difference splitting schemes are used with modern upwind methods. A number of deficiencies in these ideas remain and need to be investigated further.

The Riemann problem is defined for one-dimensional flow. As employed in present methods, the fluxes and the solution for the dependent variables are determined by the ensuing wave field produced when two gases at different states are allowed to interact. In using Riemann solvers for one-dimensional problems, solutions can be computed that can include shock waves with as few as one transition zone. However, the extension to two and three dimensions is presently accomplished by assuming a series of one-dimensional waves, and a truly satisfactory three-dimensional Riemann solver presently does not exist. Basic considerations attest to the importance of such solvers. For example, vorticity is nonexistent in one dimension where the classic Riemann methods are derived. Yet when multidimensional applications are made, shear waves naturally appear. The implication is that the multidimensional solutions using such one-dimensional modeling ideas are inappropriate.

Along this line, the development of effective three-dimensional solvers requires that substantial information be available about any shock present in the flow. It is necessary to deduce both wave orientation and propagation information from the given solution. This is a result of the nonuniqueness of the local solution to the Riemann problem in several space dimensions. These issues lead to questions regarding the comparison of classical shock fitting and solutions with three-dimensional Riemann solvers. Both approaches need to be pursued. In addition, the flux limitation necessary to produce monotone shock transition needs to be studied in detail. This issue becomes especially important when time asymptotic solutions are computed. Typical limiting problems are evidenced by convergence rates that reach a plateau and level out at a reasonable level. The convergence rate and level depend on both the form of limiter and the particular variable

limited in the solver. Further research is required to provide insight into this behavior.

The complex modeling requirements and delineation of the various flight regimes lead one to question several current approaches used to solve the equations governing fluid flow, particularly in low-density hypersonics. A more satisfying approach may be to attempt to model these flows with more general flow theories. In this light it is worthwhile to expend effort in direct attacks on solutions of the Boltzmann equation. Perhaps some simplification can be achieved by using model distribution functions, which retain the essential features in the flow regime of interest.

ROTORCRAFT FLOWS

Vortex-dominated flows represent another area of major concern. Correct representation of fluid physics is critical in such applications as high-angle-of-attack aerodynamics, helicopter rotor flows, and turbomachinery aerodynamics. To date, conventional numerical schemes have been used to compute flows in the category with limited success. In turbomachinery flows some additional modeling has been incorporated to make three-dimensional calculations feasible. However, in helicopter rotor flows, improved methods are required for solutions to both the hover and forward flight cases. In no other problem is the correct vorticity transport as critical. Methods that use vorticity conservation as an auxiliary constraint would be of great value. Improved induced velocity fields and wake characterization would provide better information for rotor analysis and design than is presently available.

Numerical methods applied to high-angle-of-attack aerodynamics problems lead to difficulties in lee-side flow where flow separation occurs. Present methods exhibit deficiencies that need to be addressed. In inviscid calculations a surprising amount of misunderstanding exists regarding calculations for problems such as vortex roll up over delta wing configurations. Emphasis in this research area will illustrate the need to pose properly such problems and interpret the results as those due to an inviscid solution. Separation at high angle of attack also remains as a critically difficult problem area. Flows over intakes at high angle of attack separate and cause significant losses in thrust and propulsion system efficiency. Turbulence models and the turbulence closure problem are major stumbling blocks in computer modeling in many of these cases.

FULL-SCALE AIRCRAFT SIMULATION

Large-scale simulation of flow over entire aircraft is necessary to maximize the impact of computational aerodynamics on the design process. To date, very few entire aircraft flow fields have been computed. There needs to be a continued emphasis on applying current methods to compute the flow over full configurations. Here mesh generation is a major issue and remains a major deterrent to full-scale computer simulation. While this will become easier to resolve as faster machines with more available memory are introduced, more emphasis should be given to mesh-generation technology.

MONTE CARLO AND BOLTZMANN EQUATIONS METHOD

In the superaerodynamics area, Monte Carlo methods have been used as the major simulation tool. In both slip flow and free molecular regime, new ideas and approaches are needed. Present emphasis on hypersonic military and commercial vehicles will necessitate development of better techniques for flow calculations in these regimes. A better understanding of nonequilibrium chemical effects is also important. With the emphasis on planetary probes, exotic gases usually not considered must be used in planetary entry calculations. Radiation heat transfer is also a major factor in these cases. While the numerical modeling of chemistry problems is covered elsewhere, their importance in the hypersonic regime cannot be overemphasized. For some flow problems involving complicated physics and/or rarefied gas, simulation methods based on Boltzmann equations and Monte Carlo particle methods seem quite appropriate. Indeed, impressive results have been obtained by Japanese investigators using a Boltzmann-equation-based method to simulate multidimensional hypersonic flow. With the anticipated advances in super and parallel computers, these approaches may be attractive methods in the future.

UNSTEADY FLOWS

Virtually all flow in nature is unsteady. However, to date, simulations have largely been concentrated on computing solutions to steady flow problems. Rocket launch booster separation, store separation from aircraft, and forward flight of a helicopter rotor are all examples where the unsteady effects

cannot be neglected. Little work has been done in this area partly due to inadequacies in models and partly to lack of computer memory and processing power. Continuing research is needed, particularly to compute solutions using the unsteady Navier-Stokes equations to better understand transient effects on aerodynamics of vehicle components.

ACOUSTICS AND FLUID-STRUCTURE INTERACTION

Simulation of air-borne and water-borne acoustical phenomena by computational methods has become a subject of considerable interest in recent years. With increased performance demands for modern submarines, simulation of acoustical scattering and radiation in complex submerged deformable bodies has emerged as a crucial research area. Today, realistic computer simulations of acoustical phenomena connected with moving submarines is impossible, and several breakthroughs must be made before these very large scale problems can be treated with confidence.

Research is needed on boundary element methods, which seem to be naturally well suited for the exterior problems of acoustics, parallel computing, modeling techniques for the effects of turbulent boundary layers, and new algorithms for treating resonance and problems with multiple scales. Many research areas discussed elsewhere in this report impact on acoustical simulations and must be pursued if there is to be any progress.

WEATHER PREDICTION

A long-time goal of CFD has been reliable weather prediction—a goal that has not yet been met satisfactorily. Much of the research on turbulence, if successful, could have a positive impact on our ability to predict weather.

13

NONLINEAR DYNAMICS OF MULTIPARTICLE SYSTEMS

There exist numerous systems of scientific and engineering interest, such as bulk granular solids and particulate suspensions, which involve large numbers of mechanically interacting solid particles. The need to understand better their rheological behavior and their interaction with processing equipment has long been recognized, but progress in the field is hampered by the difficulty of comprehensive experimental measurement. Advances in the basic science could provide significant economic benefits to a wide variety of energy and mineral recovery industries and other technologies utilizing or producing solid materials.

Recent years have seen a vigorous growth of activity in computer simulation of multiparticle systems by direct solution of the underlying equations of motion that, because of the inherent nonlinearity and numerous degrees of freedom, are otherwise mathematically intractable. Much of the current work involves the mechanics of soils, granular materials, and fluid-particulate suspensions.

In many respects, particulate simulations can be viewed as an outgrowth of similar work in the fields of the statistical physics, statistical mechanics, and kinetic theory of molecular solids and fluids, where such nonequilibrium molecular dynamics (NEMD) simulations represent a well-established, continually evolving, computational science.[10]

In contrast to the conservative Hamiltonian structure of most molecular systems, many particulate systems of engineering interest are highly irreversible due to frictional forces operating at the particulate level, which not only makes for additional mathematical complexity but also raises new and interesting scientific questions about the statistical mechanics

[10] D.J. Evans and W.G. Hoover, 1986, Annual Review of Fluid Mechanics, Vol. 18, p. 243; and B.J. Adler, 1989, "Challenge in Computational Statistical Mechanics," Lawrence Livermore National Laboratory Report UCRL 100753.

of strongly dissipative systems. Among these questions is the extent to which certain well-established molecular concepts carry over to particulate systems, where one can discern phenomena such as granular "temperature," particle diffusion, and viscoelasticity of granular flow.[11]

Despite the obvious differences from molecular statistical mechanics, one can nevertheless exploit a now generally accepted paradigm in that field based on the triad of continuum modeling, computer simulation, and experimental testing.[12] Computational simulations on idealized particle systems offer a relatively direct method to test, improve, or reject continuum models before undertaking much more costly and difficult experimental measurements on real systems. Also, with computer simulation many detailed properties of idealized systems can be "observed" that are virtually inaccessible to direct experimental measurement. Apart from the practical utility, this may serve to spawn scientific questions and ideas that would otherwise be slow to emerge. Following is a summary of a few important research efforts now under way in the simulation of multiparticle systems.

DISCRETE PARTICLE SIMULATION OF GRANULAR MATERIALS

Three methodologies to model the mechanical behavior of deforming or flowing systems comprising many distinct granules can be discerned. First and most well established is a continuum description based on suitable phenomenological constitutive relations that, with the advent of high-speed computation, can now be supplemented by discrete models of the type to be discussed briefly here.

The Particle Dynamics Method

Numerical techniques from NEMD, which have already served to elucidate transport phenomena in molecular systems, are now being adapted to systems of macroscopic particles.

[11] C.S. Campbell, 1990, Annual Review of Fluid Mechanics, Vol. 22, pp. 57–92.

[12] C.G. Gray, and K. Gubbins, 1984, Theory of Molecular Fluids (International Series of Monographs on Chemistry), Vol. I, Oxford.

Among these techniques are trajectory-tracking algorithms that allow the motion of individual particles to be followed as they interact with inelastic frictional contact forces and gravity. In molecular-dynamics-type calculations, the Particle Dynamics Method (PDM),[13] the Newtonian equations of motion are solved for a large number of interacting grains (typically a few thousand). In the quasi-static limit appropriate to certain problems in soil and rock mechanics, one recovers the so-called Distinct Element Method of Cundall[14] and others. Displacements, linear and angular velocities, and forces on individual particles are all computable, and the temporal and spatial averaging of these quantities allows direct computation of densities, flow speeds, strains, and stresses in a granular material. Figure 13.1 illustrates the kind of complex flow that can be treated by current simulations.

While numerous quantities may be computed directly, one may also attempt to extract continuum constitutive relations and then switch to a continuum description for actual applications. This approach, often desirable because of the relatively small number of particles that can be monitored by direct integration of the Newtonian equations, has led to new continuum models for granular materials based on concepts from dense-gas kinetic theory. However, the inelasticity of microscopic collisions and local energy loss is found to have dramatic effects on the distribution of microscopic velocity fluctuations and the resulting non-Newtonian fluid behavior.

By means of periodic-cell models, numerical simulations with samples as small as a few hundred particles have been able to provide valuable new information on bulk material properties. Figure 13.2 shows a periodic-cell computation involving 512 particles per cell. Application of these same methods to quasi-static deformation of inelastic frictional particle assemblies reveal long-range spatial correlations reflected in the formation of stress networks ("bridges or arches") extending over many particle diameters, somewhat reminiscent of critical-point phenomena in molecular systems. The meaningful simulation of bulk material properties in such situations may require large numbers of particles in order to span the typical "arch" or coherence length (perhaps as high as

[13] O. Walton, et al., 1988, Micromechanics of Granular Materials, Elsevier, New York.

[14] E.T. Brown, ed., 1987, Analytical and Computational Methods in Engineering Rock Mechanics, Unwin Hyman, Inc., Cambridge, MA.

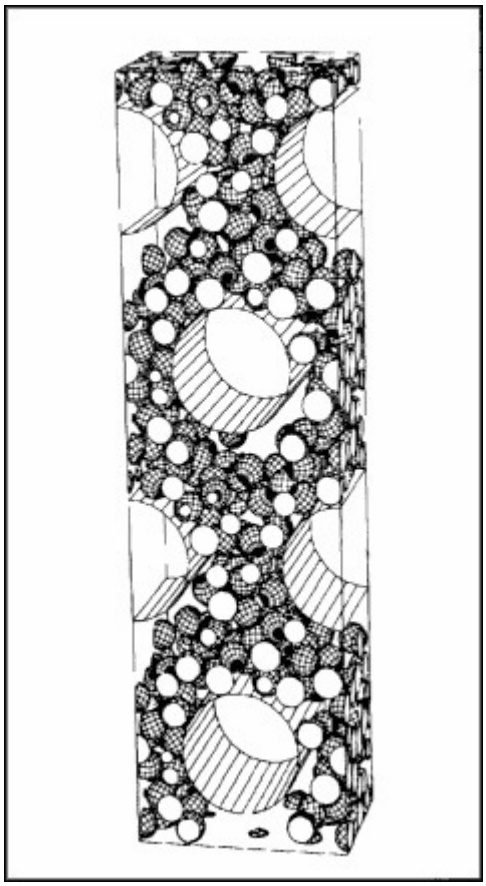

Figure 13.1
Perspective view of spherical particles in a simulation of glass beads falling through a chute with cylindrical obstacles, in a configuration similar to proposed design of burner and cooler for a commercial oil-shale retort. Open circles show intersection of particles with periodic boundary parallel to plane of figure. Source: After O. Walton, 1988, *Energy and Technology Review*, Lawrence Livermore National Laboratory, Report UCRL-52000-88-9, Livermore, CA). Reprinted with Permission.

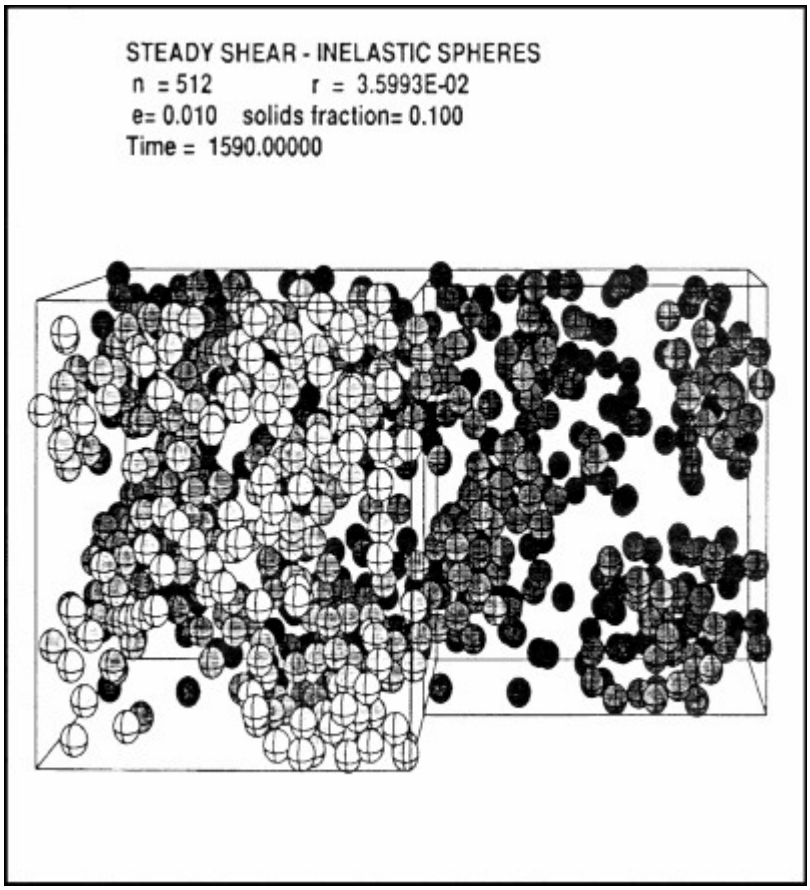

Figure 13.2
Computer-generated snapshot of 512 particles in one and one-half neighboring periodic cells, from a discrete particle simulation of steady shearing flow of a dry granular medium. Periodic-image cells above the primary cell move from left to right and those below from right to left at constant velocity to stimulate the shear. Source: After O. Walton, 1988, *Energy and Technology Review*, Lawrence Livermore National Laboratory, Report UCRL-52000-88-9, Livermore, CA). Reprinted with Permission.

thousands of particles for three-dimensional simulations). Modern supercomputers coupled with vectorized coding of simulation algorithms are just beginning to achieve adequate sample sizes in sufficiently short times to make such simulations practicable. Very little simulation work with such large sample sizes has actually been done to date since most researchers are still concentrating on the refinement of interparticle-force models and methods for efficiently calculating the response of frictional and irregularly shaped particles. The advent of a new, fast, and inexpensive generation of massively parallel computers could have a significant impact on the scale of problems accessible to direct simulation, with dramatic improvements in both the quantity and quality of results. In the future we can expect to see such simulation of granular materials become more and more commonplace, as the computing power of widely available workstations continues to increase and other hardware and software developments allow for the treatment of ever more realistic and larger systems.

Lattice-Gas (Cellular-Automata) Models

To overcome some of the current computational limitations of PDM, an alternative, inspired by the Lattice-Gas Model (LGM) or Cellular-Automata Model of statistical mechanics, the Lattice Grain Model (LGrM), has recently been introduced [15] to allow for explicit tracking of the trajectories of millions of grains. For such particle-tracking methods, neither complex geometric boundaries nor explicit time dependence of flow poses great difficulty in the calculation, in contrast to the numerical solution of the corresponding continuum partial differential equations (when those are known).

The LGrM trades off particle complexity against particle number. While in many PDM applications a fairly realistic contact force is employed, the LGrM description of particle interactions usually entails only momentum conservation, together perhaps with a one-parameter ("coefficient of restitution") model of energy loss. Thus, the update of particle states due to collisions can be performed much more quickly than in the PDM. Because particles are tied to the lattice, the problem of contact detection, a significant computational burden in the

[15] G.M. Gutt, and P.K. Haff, 1990, Proceedings of the 5th Distributed Memory Computer Conference, University of South Carolina, Charleston, April 9–12.

PDM, is highly simplified with the result that the LGrM can follow the motion of vastly larger numbers of particles.

In the LGrM, energy is not conserved, and the corresponding continuum limit of the automata motion is not as well understood as the Navier-Stokes limit of the LGM. Nevertheless, LGrM calculations involving approximately 16,000 particles on an NCUBE parallel computer show encouraging agreement with experiments on time-dependent granular flows in highly complex geometries.[16] LGrM simulations like these are highly suited to parallel computation because of the local nature of motion and collision on the lattice. Other parallelized simulations involving upwards of 10^6 particles in granular Couette and Poiseuille flows have been reported.

NUMERICAL MODELING OF PARTICULATE SUSPENSIONS

The flow of suspensions and slurries is complex and unpredictable, especially near solid walls. A major cause of this complexity are the hydrodynamic forces between solid particles transmitted by the fluid, which, together with aggregation of the particles, lead to many interesting effects, such as dramatically large viscosity. Unfortunately, no simple pairwise additive force rules are adequate to describe these interactions when the solid concentrations exceed more than a few volume percent, so there is an urgent need for novel computational methods.

Stokesian Dynamics

For particles suspended in viscous fluids at zero Reynolds number (Stokes) flow, a new numerical method called Stokesian Dynamics (SD) [17] is currently under development. The method entails solution of the fluid equations throughout the space between particles and determination of the induced forces on the surfaces of the suspended particles.

[16] Gary M. Gutt, 1989, Ph.D. thesis, *The Physics of Granular Systems* , Division of Physics, Mathematics, and Astronomy, California Institute of Technology.

[17] J.F. Brady, and G. Bossis, 1988, Annual Review of Fluid Mechanics , Vol. 20, pp. 111–158.

The crux of the technique is the computation of the induced forces which for spheroidal particles can be done numerically to any desired accuracy by multipolar expansions of particle-surface force distributions.

SD calculations of the theological properties of suspensions have been shown to give excellent agreement with analytical theories for the limiting case of regular arrays as well as with high-frequency experimental viscosity measurements on random dispersions. Such simulations can provide accurate benchmark computations of theological properties, for both small clusters of spheres and large random assemblages. These computations serve as valuable limiting cases for the theory development and testing of various numerical simulations of full (inertial) dynamics behavior of suspensions. The SD method is computationally intensive, especially for simulations of high solids packing.

The computational efficiency of SD has been systematically improved over the past couple of years, and periodic cells containing upwards of a hundred particles can now be treated. To improve efficiency, the method has been modified to include the effects of short-range lubrication forces between nearest-neighbor particles as separate, pairwise forces, significantly reducing computational effort in the high-concentration regime. This method has now been applied to such problems as determining the transport properties of random sphere assemblies, including viscosity, permeability, sedimentation velocity, and diffusion coefficient at various concentrations. With current computer limitations the inherent N^2 dependence of the method precludes its application to systems of more than a few hundred particles.

Lattice-Gas Suspension Models

The Lattice-Gas Model discussed above also provides an alternative approach to the simulation of a molecular fluid containing supramolecular suspended particles.[18] For a sufficiently large number of particles, this model presumably reproduces the Navier-Stokes equations for the suspending fluid, but with the significant advantage that thermal (Brownian) fluctuations for colloidal-sized particles can also be included. One goal of current lattice-gas suspension modeling is to understand the relationship between the theological pro

[18] A.J.C. Ladd, M.E. Colvin, and D. Frenkel, 1988, Physics Review Letters, Vol. 60, p. 975.

perties of suspensions and their microstructure under shear. The nonequilibrium structures that arise from complex hydrodynamic interaction, possibly competing with Brownian effects, can be very different from those at rest.

While the earliest three-dimensional LGMs were not strictly isotropic, newer models (based on the projection of a face-centered four-dimensional hypercube onto a three-dimensional cubic array) can now eliminate artifacts such as anisotropic viscosity. New local "microrules" have also recently been developed to simulate no-slip fluid-particle boundary conditions. Initial tests of the newly developed simulations show that hydrodynamic interactions between particles are quantitatively reproduced by the LGM, even when the particles are almost touching. The significant improvements in efficiency and realism of these new LGM fluid simulations and their inherent applicability to a much wider range of problems suggest that much future effort should be concentrated on further development and application to particulate suspensions.

One exciting aspect of the lattice-gas approach is the prospect of simulating suspensions with about the same computational effort as that now used for dry granular solids by means of special very large-scale integrated (VLSI) chips, which are particularly well suited to simple cellular automata. It is estimated that the massively parallel computer hardware necessary to update a lattice-gas cellular automata at Cray supercomputer speeds could be produced for about $10,000. Prototype machines for such calculations are now under development in the Netherlands.

Such hardware developments hold promise for the scientific and engineering aspects of suspension dynamics, and possible future applications include studies of anisotropic molecular shapes, colloidal forces, and polydisperse-size effects. Such modeling need not be limited to periodic-cell models. A variety of realistic geometries to consider are pipes, chutes, and nozzles.

14

MANUFACTURING PROCESS, DESIGN, AND DEVELOPMENT

The manufacture of material products involves many unit processes that alter the external form and surface as well as the interior structure of the material components and final assembly. These include such processes as grinding, melting, chemical processing, atomizing, agglomerating, casting, molding, thermal processing, deformation processing, forming, machining, coating, and joining. From a scientific viewpoint, these processes involve such subjects as fracture mechanics, heat transfer, fluid mechanics, electromagnetodynamics, chemical reactions, solid mechanics (especially plasticity), and metallurgical reactions. In the past these subjects have been useful as qualitative guides for the design and development of various manufacturing processes, but such activities have progressed primarily by empirical trial and error.

MATHEMATICAL ASPECTS

Today, the subject of design and development of manufacturing processes is undergoing a dramatic revolution. There is a long way to go, but the direction is increasingly clear. Physically realistic mathematical simulations of unit processes are being developed. Increasingly, these simulations are being used to replace many of the physical trials that were necessary in the past. The times and costs of process development are being reduced. The final designs are better because they are based in quantitative physical science. The vision is a physical-science-based, mathematically expressed design methodology that can be used by manufacturing engineers to create designs of manufacturing processes with the same confidence that structural engineers now use elasticity-based methods to create designs of bridges and buildings.

The mathematical expressions of manufacturing processes are considerably more complex than those of elastic structures, and the processes are often inherently three-dimensional and nonsteady. Yet the development of large-scale computers has now brought the exercise of these mathe

matical simulations within reach, and it is clear that further development of computer hardware and software will render them increasingly available. The issue of feasibility is past.

Mathematical expressions of manufacturing processes involve the classical physical laws of conservation of mass, energy, and momentum. They also involve constitutive relations describing the physical response of particular materials. Specific process simulations are obtained by applying specific boundary conditions to the resulting (usually quasi-linear) partial differential equations. The boundary conditions themselves often must be constitutive, describing the surface behavior of the materials being pressed and the tools being utilized. Because the physical nature of the materials being processed is often changed by the process, the constitutive equations (both interior and boundary) must involve not only state equations expressing, say, the relationship of deformations to stresses and temperatures in the current material state but also evolution equations expressing how the material state is changing.

Development of the constitutive equations for manufacturing processes is itself a rich field for computational mechanics. Such constitutive equations can be developed empirically by fitting data from many carefully controlled laboratory experiments. Increasingly, however, the development of these equations is also being guided by micro-mechanical models that idealize the microstructure of the materials being processed and calculate the response of idealized structures to various thermomechanical responses to create the macroscopic constitutive relations. This leads directly to constitutive relations with dependent variables representing specific microstructural features. In fact, it is now becoming possible to develop macroscopic constitutive equations based on a hierarchy of increasingly fine substructural scales, reaching from featureless continua to atomic structures. This is an extremely rich area for the integration of computational mechanics, materials science, and solid-state physics.

These complex and physically detailed constitutive relations provide a rich basis for concurrent design of materials and their processes, but such detail is usually unnecessary for good design of a process for a particular material. In fact, it is often desirable, because of the computational simplification that can be achieved, to perform preliminary design with the simplest possible constitutive models such as Newtonian fluids, linear hypoelastic solids, and rigid/perfectly plastic solids. Then final designs can be achieved by iteration about the preliminary design using increasingly complex models.

To summarize, a revolution is occurring in the way manufacturing processes are being designed and developed. Costly and time-consuming empirical methods of the past are being replaced by more accurate, less costly, and faster physical-science-based computer simulations. The material aspects of the computer simulations are contained in constitutive equations that describe the evolution of the material as well as its current response. Development of the constitutive equations as well as the full process simulation involves computational mechanics. Simple constitutive equations are advocated for preliminary process design and more detailed equations for fine tuning.

Permission to reproduce some figures in this document was given by the following:

1. Pergamon Press, New York, N.Y., Figs. 6.1 and 6.2 from and 1988 and 1989.
2. Elsevier-Science Publishers, Amsterdam, The Netherlands, Fig. 6.3 from *Mechanics Materials*, Vol. 6, 1987.
3. Fig. 6.4; Innovations in Ultrahigh-Strength Steel Technology, *Proceedings of the 34th Sagamore Army Materials Research Conference,* (ed. by G.B. Olson, M. Azrin, and E.W. Wright), pp. 3–66.
4. Kluwer Academic Publishers, Amsterdam, The Netherlands, Fig. 6.5 from *International Journal of Fractures*, 1988.
5. Lawrence Livermore National Laboratory, Fig. 13.1 and Fig. 13.2, 1988.