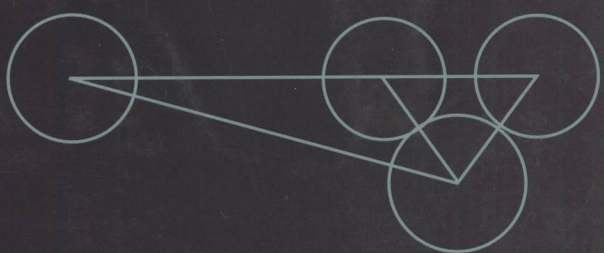
WILEY PROFESSIONAL PAPERBACK SERIES

MOLECULAR DYNAMICS SIMULATION

Elementary Methods



J. M. HAILE

MOLECULAR DYNAMICS SIMULATION

Elementary Methods

J. M. HAILE
Clemson University
Clemson, South Carolina

Wiley Professional Paperback Edition Published 1997

A Wiley-Interscience Publication
JOHN WILEY & SONS, INC.
New York • Chichester • Weinheim • Brisbane • Singapore • Toronto

Library of Congress Cataloging in Publication Data:

Haile, J. M.

Molecular dynamics simulation: elementary methods/J. M. Haile.

p. cm.

"A Wiley-interscience publication."

Includes bibliographical references and index.

ISBN 0-471-81966-2

ISBN 0-471-18439-X Wiley Professional Paperback Edition

1. Molecular dynamics—Computer simulation. 2. Causality (Physics)—Computer simulation. 3. Chemistry. Physical and theoretical—Computer simulation. I. Title.

QC168.7.H35 1992 91-31963

532'.05—dc20 CIP

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

MOLECULAR DYNAMICS SIMULATION

...there is a never ending interest
in the definite mathematical problem
of the equilibrium of motion
of a set of points endowed with inertia
and mutually acting upon one another
with any given force.

Sir William Thomson, *Baltimore Lectures on Molecular Dynamics and the Wave Theory of Light*, Lecture XI (1884)

And these toys did so firmly possess
his imagination with an infallible opinion
that all that *machina* of dreamed inventions
which he read was true,
as he accounted no history in the world
to be so certain and sincere as they were.

Miguel de Cervantes, *Don Quixote* (1605)

PROLEGOMENA

Molecular dynamics methods are now orthodox means for simulating molecular-scale models of matter. The methods were originally devised in the 1950s, but they only began to receive widespread attention in the mid-1970s, when digital computers became powerful and affordable. Today, these simulation methods continue to attract attention from researchers with new problems and from students new to molecular theory.

The essence of molecular dynamics is simply stated: numerically solve the N-body problem of classical mechanics. Science has its fashions, just as do literature and music and architecture. Since the time of Newton, the N-body problem has been viewed as important, but the *reasons* for its importance have evolved. At the present time, its importance stems from attempts to relate collective dynamics to single-particle dynamics, attempts motivated by the hope that the puzzling behavior of large collections of particles can be explained by examining the motions of individual particles. For example, how does flow of fluid around an object produce a turbulent wake? How do atoms on a protein molecule move together so the protein folds in life-supporting ways? How does swirling fluid produce a long-lived vortex, such as the great red spot on Jupiter? How do chain molecules in solution self-assemble into structures such as micelles, vesicles, and lamellae? How does a local disturbance of a few molecules (say by a laser pulse) propagate throughout a system? How do individual molecules combine to form new molecules?

Such questions suggest that molecular dynamics can enlighten diverse research areas; however, its very diversity of application may have hampered development of an introductory literature that helps train new simulators. Specialists have been too busy exploring new territory to worry much about introducing their colleagues to the basic methodology. In addition to new

research, molecular dynamics is now serving as an educational tool in courses in materials science, solid and liquid state physics, statistical mechanics, and physical chemistry. The methods would also be appropriate for study in courses on classical mechanics, advanced engineering thermodynamics, and transport phenomena. It seems that students and instructors, as well as new researchers, need a modest introduction to molecular dynamics methods.

Molecular-scale computer simulation involves a three-step procedure: (a) model individual particles, (b) simulate the movements of a large number of the model particles, and (c) analyze the simulation data for the required collective phenomenon. The emphasis of this book is on step (b). Essentially nothing is discussed concerning step (a), and discussions of step (c) are limited to traditional static and dynamic properties. Steps (a) and (c) depend on the problem to which the simulation is being applied; this book is not about applications but about methodology. Specifically, the book is restricted to molecular dynamics applied to hard spheres and Lennard-Jones atoms because these two substances encompass the two forms of molecular dynamics algorithms. This book contains nothing about molecular substances, quantum effects, multibody interactions, long-range interactions that would require Ewald sums, spectroscopic properties, nonequilibrium simulations, or parallel processing. Instead, the objective is to present a core of fundamentals that should be common to all users of molecular dynamics. From this small core, the reader may build to special applications.

This book is organized around two questions: (a) Why does molecular dynamics work (Chapters 1–5)? (b) How is molecular dynamics used to estimate properties (Chapters 6 and 7)? The emphasis is on *why* because a thorough understanding of fundamentals will enhance the possibility of success when molecular dynamics is applied to situations not explicitly covered here. It is, therefore, perhaps worthwhile to consider what should be included in a thorough response to *why*.

In Book II of his *Physics*, Aristotle proposed a theory of causation in which a complete explanation for a thing necessarily encompasses four causes: (i) material cause, which explains a thing in terms of its component parts; (ii) formal cause, which explains by invoking the design or relation among the components; (iii) efficient cause, which explains a thing as a consequence of an agent that implements the design; and (iv) final cause, which explains by virtue of the use made of the thing. The structure of this book revolves around these modes of explanation:

Chapter 1 considers *final cause*. What are the uses of molecular dynamics? How does molecular dynamics stand in relation to other ways of performing molecular simulation and other ways of doing science? It is commonly held that simulation serves as a competitor for theory and experiment; however, it seems to me that simulation stands in opposition, not necessarily to theory or experiment, but to the broader attitude of reductionism.

Chapters 2 and 4 review *material cause*. Molecular dynamics works because it is a confluence of diverse ideas from nonlinear dynamics, kinetic theory, sampling theory, and numerical methods.

Chapters 3 and 5 present *formal cause*. Molecular dynamics works because of the way it organizes the components discussed in Chapters 2 and 4. One organization of those parts produces an algorithm for simulating hard spheres (Chapter 3), while a different organization produces an algorithm for soft spheres (Chapter 5).

Appendices I, K, and L illustrate *efficient cause*. Molecular dynamics works because computational hardware and software are available for implementing the designs discussed in Chapters 3 and 5. This book contains nothing about hardware and descriptions of software are limited to the codes given in the appendices. This is not to undervalue the role of programming in molecular dynamics; rather, it reflects the attitude that programs themselves can serve as means for communicating among people as well as between people and machines. People can learn to read programs, just as they can learn to read foreign languages or music. The documentation tries to explain what the code is doing or why it is doing it, but *how* the code is accomplishing its tasks is left for the code itself to reveal.

In large measure the objective of the book is for a novice to understand the molecular dynamics codes given in the appendices. *Understand* means that the user will (i) be able to explain what the code does, (ii) be able to have a machine execute the code, and (iii) be able to judge whether the run and its results are meaningful. The programs are not to be used as black boxes; rather, they should be modified, tinkered with, improved—such exploratory play reinforces a learning process. The codes are not written to execute extremely fast: they are intended to be understood. In only a few instances are the codes compromised in favor of execution speed over comprehension. The programs are in Fortran and should execute under any Fortran 77 compilation.

Aristotle enumerated not only four modes of explanation, but he also worried about chance as a causative agent. Although molecular dynamics is often touted as a completely deterministic method, in fact, chance has a healthy and decisive role to play in molecular dynamics. Chance events—random unpredictable events—occur not only as consequences of nonlinearities in the differential equations that prescribe particle trajectories, but, more subversively, as consequences of the hardware and software used in performing the simulations. The role of chance in molecular dynamics is discussed in Chapters 2 and 3.

Molecular dynamics simulation brings to bear a startling range of ideas from several disciplines. This is one reason the topic proves valuable in a classroom. Readers are expected to have some knowledge of classical mechanics, vector analysis, numerical analysis, thermodynamics, and Fortran programming. Previous exposure to kinetic theory, statistics, transport phenomena, and statistical mechanics would also be beneficial.

Computer simulation has become an important and respected research tool, complementing the traditional approaches of theory and experiment. As with any tool, its effectiveness is limited by the skill and dedication of the craftsman who guides its use. But I have come to realize that there are

unusual ways in which simulation is something other than simply a tool. Use of a tool suggests one knows at least the goal, if not the path. However, in research one is often diverted from an original goal, and computer simulation can motivate such a diversion: simulation can produce revelation. By this I mean that simulation has the potential to yield the unexpected. Most often the unexpected is readily understood: "Oh yes, of course! Why didn't I think of that before?" But the fact is, we didn't think of it before, and chances are we might never have thought of it, certainly not nearly so soon, without the simulation.[†]

I think this revelatory aspect of simulation is a major reason why it has become so popular and enthralling. People enjoy pleasant surprises, especially if they are also educational. I think there is a latent danger here, though. It is easy to simulate models so abstract that the results have only intrinsic interest: they convey no insight into any aspect of reality. In fact, it is usually easier to simulate *interesting* models rather than realistic ones, and therein lies the danger. Davis and Hersh[‡] have warned of the dangers of abstraction in mathematics and digital computation. This warning applies, as well, to molecular dynamics.

The use of elaborate abstraction to interpret reality was the focal point of a protracted, fourteenth-century battle that pitted both Church and academia against a solitary Franciscan scholar. The principal weapon in that scholar's logical armory has come down to us transmuted into a cornerstone of modern scientific method: *Entities must not be multiplied beyond necessity*. This attitude—Ockham's razor—serves as one theme of this book: it applies not only to interpretations of results, but also to the methods used to obtain results. It should guide us in the models we build, the algorithms we use, and the programs we write. Computer simulation presents exhaustive and exhausting opportunities for making errors. Simple approaches, approaches that work and are readily understood, reduce the chances for error and enhance the chances for finding errors when they are made. Complex, sophisticated approaches are often touted as executing quickly on a computer, but in truth, they rarely prove to be great time savers.

In the fall of 1884 at Johns Hopkins University Sir William Thomson presented a series of 20 lectures under the title *Molecular Dynamics and the Wave Theory of Light*.[§] The technical content of those lectures have little in common with the volume you now hold in your hands: in his lectures Thomson sought to understand light within the framework of classical mechanics. As such, the lectures were washed away in the tide of quantum mechanics, and they failed to contribute to the development of modern

[†]W. Thomson, Lecture X, p. 103: "...I never heard of anomalous dispersion until after I found it lurking in the formulas. I said to myself, 'these formulas would imply that, and I never have heard of it.' And when I looked into the matter I found to my shame that a thing which had been known by others for eight or ten years I had not known until I found it in the dynamics."

[‡]P. J. Davis and R. Hersh, *Descartes' Dream*, Houghton Mifflin Co., Boston, 1986.

[§]W. Thomson, reprinted in *Kelvin's Baltimore Lectures and Modern Theoretical Physics*, R. Kargon and P. Achinstein, Eds., MIT Press, Cambridge, MA, 1987.

science. Nevertheless, if you read the lectures, not for technical content but to gain insight into the workings of a nineteenth-century scientific mind, you find that Thomson's mind-set—the way he tackled problems—was modern indeed. Or, to twist this observation around, the way we tackle problems is not at all modern—it's merely the power of the computer that's new and that allows us to push an established methodology farther than ever before. I think there is value in placing what we do, including computer simulation, within a historical context, and so I have embellished this text with quotes from Thomson's Baltimore lectures. In 1892 Thomson was named Baron Kelvin of Largs.

Buried within this book is the attitude that simulation only makes manifest possible realities that are latent within the human mind. There are, after all, no little spheres doing their dance in the computer; to the machine, a simulation is merely an elaboration of electronic impulses responding to the directions of a rudimentary arithmetic. Computer graphics do not reveal what is happening in the machine, but rather what humans *interpret* as happening. To contrive a beneficial understanding of these interpretations, these latent realities, we must make contact with the world of our senses: when isolated from theoretical and experimental forms of science, simulation is sterile. It is my hope that this book will help at least one novice keep clear the distinction between the simulated and the real; that it will stimulate one enthusiastic novice to pause and think, then simulate, then think again; that it will provoke one novice to abandon simulation for something more worthy of his talents; that it will challenge one novice to improve any one of the methods described herein. By such small steps does science advance.

Acknowledgments

The following individuals unselfishly took time from their productive schedules to read and criticize portions of the manuscript: J. J. Erpenbeck, C. G. Gray, P. A. Monson, F. H. Ree, and W. A. Steele. I am grateful to them all. Special thanks to Sohail Murad and Ariel Chialvo who read the entire manuscript and corrected scores of errors.

But in spite of the constructive comments provided by these experts, I have not always followed their advice. This is an unconventional book to the extent that objective discussions of methods and results are mingled with subjective views that illuminate those objective perceptions. It presumes a skeptical reader for, as Umberto Eco reminds us in *The Name of the Rose*,

"Books are not made to be believed, but to be subjected to inquiry."

HAILE@CLEMSON

Aboard the Sloop Passepartout
Lake Keowee, South Carolina
June 1991

A NOTE TO THE USER

At the end of each chapter I have posed a variety of problems: some exercise the material in the chapter, some go beyond the material explicitly presented, some try to connect to ideas from previous chapters, some try to connect to ideas from supporting disciplines. To meet the needs of a diverse readership, the problems vary in complexity and topic. Few of the problems are especially difficult, but properly executed, many are time consuming. Rather than work many problems cursorily, it seems to me more profitable to work a few problems thoroughly. The test for thoroughness is that students begin to use a problem as a springboard for devising their own problems.

To warn the reader (and instructor) that certain problems require special consideration, I have marked them with a flag from the set of international maritime signals. The flags are explained in the following table.

Flag	Alphabetic Character	Maritime Meaning	Usage in This Book
⊠	V (victor)	*Require assistance*	A computer is needed for these problems. Aside from the effort needed in attacking the problem, sufficient time should be allowed by the reader (and instructor) for coping with the vagaries of machine and programming language.
■	U (uniform)	*Standing into danger*	These problems will challenge the average reader, either because of difficulty or because of the dedicated effort required.

MOLECULAR DYNAMICS SIMULATION

CONTENTS

1

INTRODUCTION

Molecular dynamics simulations compute the motions of individual molecules in models of solids, liquids, and gases. The key idea here is *motion*, which describes how positions, velocities, and orientations change with time. In effect, molecular dynamics constitutes a motion picture that follows molecules as they dart to and fro, twisting, turning, colliding with one another, and, perhaps, colliding with their container.

This usage is not unique: *molecular dynamics* may also refer to the motions of real molecules when studied primarily by molecular beam [1] or spectroscopic [2] techniques. This terminological confusion is compounded by *lattice dynamics* [3], which refers to the study of vibratory motions of atoms in solids, and by *molecular mechanics* [4], also called force field calculations, which refers to quantum mechanical calculations of the structure of individual molecules. This book is concerned with molecular dynamics solely in the sense of simulation.

Molecular dynamics simulation is the modern realization of an old, essentially old-fashioned, idea in science; namely, the behavior of a system can be computed if we have, for the system's parts, a set of initial conditions plus forces of interaction. From the time of Newton to the present day, this deterministic mechanical interpretation of Nature has dominated science [5]. In 1814, roughly a century after Newton, Laplace wrote [6]:

> Given for one instant an intelligence which could comprehend all the forces by which nature is animated and the respective situation of the beings who compose it—an intelligence sufficiently vast to submit these data to analysis—it would embrace in the same formula the movements of the greatest bodies of

> the universe and those of the lightest atoms; for it, nothing would be uncertain and the future, as the past, would be present to its eyes.

If this approach is thwarted by the complexities of reality, then we replace reality with a model. In one of his Baltimore lectures (Lecture XI), roughly a century after Laplace, Thomson observed [7]:

> It seems to me that the test of "Do we or not understand a particular subject in physics?" is, "Can we make a mechanical model of it?"

Today, roughly a century after Thomson, we remain undeterred from Laplace's dream: the requisite "intelligence" is provided by the digital computer, the "respective situation" is a set of initial positions and velocities, "the same formula" though not literally true could be interpreted as the same algorithmic program, and Laplace's universe has given way to model universes. Now, deterministic mathematical models pervade not only the physical sciences and engineering, but the life and social sciences [8] as well.†

This attitude is old-fashioned in the sense that, while often successful, it is nevertheless simplistic. In spite of Laplace's claim, we can still identify systems that are unpredictable—stock markets and the weather, for example. Why should this be? If deterministic mathematical models can help us successfully land *Apollo XI* on the moon, why can't they help us predict next month's weather on earth?

The resolution of this dilemma is based on the kind of forces acting among system components: when a system contains objects that interact *nonlinearly*, the system's behavior may be unpredictable. In the past few years studies in nonlinear dynamics have decoupled deterministic from predictable [9]. *Deterministic* situations have system outputs causally connected to system inputs. *Calculable* situations are those deterministic situations in which an algorithm allows us to compute system outputs if the inputs were known. *Predictable* situations are those calculable situations in which the algorithm can be numerically implemented to actually compute the outputs. Calculable situations may be unpredictable because of the large number of inputs needed, because of an unrealistically high precision with which the inputs must be known, and/or because the algorithm's stability is sensitive to intermediate calculations. In pool, *Eight ball in the side pocket* is deterministic, calculable, and predictable; however, whether it will rain in two weeks is deterministic but unpredictable.

The overriding theme of this book is predicated on the decoupling of predictability from determinism. Be warned—that you use a machine to

†In fairness, paleontologists, at least, have discovered deterministic unpredictability. Thus, Stephen Jay Gould [10] posits that if the tape of life were rewound to some previous, sufficiently removed condition and then replayed, the result would be life unlike life as we know it. For more technical conjectures on connections between life and deterministic unpredictability, see Fox [11].