LECTURE NOTES IN PHYSICS

M. Ferrario
G. Ciccotti
K. Binder
(Eds.)

Computer Simulations in Condensed Matter: From Materials to Chemical Biology – Vol. 2

Springer

Lecture Notes in Physics

Editorial Board

R. Beig, Wien, Austria
W. Beiglböck, Heidelberg, Germany
W. Domcke, Garching, Germany
B.-G. Englert, Singapore
U. Frisch, Nice, France
P. Hänggi, Augsburg, Germany
G. Hasinger, Garching, Germany
K. Hepp, Zürich, Switzerland
W. Hillebrandt, Garching, Germany
D. Imboden, Zürich, Switzerland
R. L. Jaffe, Cambridge, MA, USA
R. Lipowsky, Golm, Germany
H. v. Löhneysen, Karlsruhe, Germany
I. Ojima, Kyoto, Japan
D. Sornette, Nice, France, and Zürich, Switzerland
S. Theisen, Golm, Germany
W. Weise, Garching, Germany
J. Wess, München, Germany
J. Zittartz, Köln, Germany

The Lecture Notes in Physics

The series Lecture Notes in Physics (LNP), founded in 1969, reports new developments in physics research and teaching – quickly and informally, but with a high quality and the explicit aim to summarize and communicate current knowledge in an accessible way. Books published in this series are conceived as bridging material between advanced graduate textbooks and the forefront of research to serve the following purposes:

• to be a compact and modern up-to-date source of reference on a well-defined topic;
• to serve as an accessible introduction to the field to postgraduate students and nonspecialist researchers from related areas;
• to be a source of advanced teaching material for specialized seminars, courses and schools.

Both monographs and multi-author volumes will be considered for publication. Edited volumes should, however, consist of a very limited number of contributions only. Proceedings will not be considered for LNP.

Volumes published in LNP are disseminated both in print and in electronic formats, the electronic archive is available at springerlink.com. The series content is indexed, abstracted and referenced by many abstracting and information services, bibliographic networks, subscription agencies, library networks, and consortia.

Proposals should be sent to a member of the Editorial Board, or directly to the managing editor at Springer:

Dr. Christian Caron
Springer Heidelberg
Physics Editorial Department I
Tiergartenstrasse 17
69121 Heidelberg/Germany
christian.caron@springer.com

Mauro Ferrario Giovanni Ciccotti Kurt Binder

Computer Simulations in Condensed Matter Systems: From Materials to Chemical Biology Volume 2

Springer

Editors

Professor Mauro Ferrario
Dipartimento di Fisica
Università di Modena e Reggio Emilia
Via Campi, 213/A
41100 Modena, Italy
E-mail: mauro.ferrario@unimore.it

Professor Kurt Binder
Institut für Physik
Universität Mainz
Staudinger Weg 7
55128 Mainz, Germany
E-mail: kurt.binder@uni-mainz.de

Professor Giovanni Ciccotti
Dipartimento di Fisica, INFN
Università di Roma La Sapienza
Piazzale Aldo Moro 2
00185 Roma, Italy
E-mail: giovanni.ciccotti@roma1.infn.it

M. Ferrario et al., *Computer Simulations in Condensed Matter Systems: From Materials to Chemical Biology Volume* 2, Lect. Notes Phys. 704 (Springer, Berlin Heidelberg 2006), DOI 10.1007/b11761754

Library of Congress Control Number: 2006927292

ISSN 0075-8450
ISBN-10 3-540-35283-X Springer Berlin Heidelberg New York
ISBN-13 978-3-540-35283-9 Springer Berlin Heidelberg New York

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilm or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer. Violations are liable for prosecution under the German Copyright Law.

Springer is a part of Springer Science+Business Media
springer.com
© Springer-Verlag Berlin Heidelberg 2006

The use of general descriptive names, registered names, trademarks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

Typesetting: by the authors and techbooks using a Springer LaTeX macro package
Cover design: WMXDesign & production GmbH, Heidelberg

Printed on acid-free paper SPIN: 11761754 54/techbooks 5 4 3 2 1 0

Preface

The school that was held at the Ettore Majorana Foundation and Center for Scientific Culture (EMFCSC), Erice (Sicily), in July 2005, aimed to provide an up-to-date overview of almost all technical advances of computer simulation in statistical mechanics, giving a fair glimpse of the domains of interesting applications. Full details on the school programme and participants, plus some additional material, are available at its Web site, `http://cscm2005.unimore.it`

Computer simulation is now a very well established and active field, and its applications are far too numerous and widespread to be covered in a single school lasting less than 2 weeks. Thus, a selection of topics was required, and it was decided to focus on perspectives in the celebration of the 65th birthday of Mike Klein, whose research has significantly pushed forward the frontiers of computer simulation applications in a broad range, from materials science to chemical biology. Prof. M. L. Klein (Dept. Chem., Univ. Pennsylvania, Philadelphia, USA) is internationally recognized as a pioneer in this field; he is the winner of both the prestigious Aneesur Rahman Prize for Computational Physics awarded by the American Physical Society, and its European counterpart, the Berni J. Alder CECAM Prize, given jointly with the European Physical Society. The festive session held on July 23rd, 2005, highlighting these achievements, has been of a particular focus in this school. In the framework of the EMFCSC International School of Solid State Physics Series, the present school was the 34th course of its kind.

However, this school can be considered as being the third (and perhaps last?) event in a series of comprehensive schools on computer simulation, 10 years after the COMO Euroconference on "Monte Carlo and Molecular Dynamics of Condensed Matter systems," and 20 years after the VARENNA Enrico Fermi Summer School on "Molecular Dynamics of Statistical Mechanical Systems." Comparing the topics emphasized upon in these schools, both the progress in achieving pioneering applications to problems of increasing complexity, and the impressive number of new methodological developments are evident. While the focus of the Varenna School was mostly on Molecular Dynamics (MD) and its applications from simple to complex fluids, the Como school included both Monte Carlo (MC) simulations of lattice systems (from

quantum problems to the advanced analysis of critical phenomena in classical systems like the simple Ising model), and the density functional theory of electronic structure up to the Car-Parrinello ab initio Molecular Dynamics techniques (CPMD). At the Erice school, a new focus was put on the paradigma of "Multiscale Simulation", i.e. the idea to combine different methods of simulation on different scales of length and time in a coherent fashion. This method allow us to clarify the properties of complex materials or biosystems where a single technique (like CPMD or MD or MC etc.) due to excessive needs of computer resources is bound to fail. Good examples presented at this school for such multiscale simulation approaches included MD studies of polymers coupled with a solvent, which is described only in a coarse-grained fashion by the lattice Boltzmann technique and hybrid quantum mechanical/molecular mechanics (QM/MM) methods for CPMD simulations of biomolecules, etc.

As a second "leitmotif," emphasis has been put on rapidly emerging novel simulation techniques. Techniques that have been dealt with at this school include the methods of "transition path sampling" (i.e. a Monte Carlo sampling not intending to clarify the properties of a state in the space of thermodynamic variables, but the properties of the dominating paths that lead "in the course of a transition" from one stable state to another), density of state methods (like Wang-Landau sampling and multicanonical Monte Carlo, allowing an elegant assessment of free energy differences and free energy barriers, etc.) and so on. These techniques promise substantial progress with famous "grand challenge problems" like the kinetics of protein folding, as well as with classical ubiquitous problems like the theory of nucleation phenomena. Other subjects where significant progress in methodological aspects was made included cluster algorithms for off-lattice systems, evolutionary design in biomedical physics, construction of coarse-grained models describing the self-assembly and properties of lipid layers or of liquid crystals under confinement and/or shear, glass simulations, novel approaches to quantum chemistry, formulation of models to correctly describe the essence of dry friction and lubrication, rare event sampling, quantum Monte Carlo methods, etc. The diversity of this list vividly illustrates the breadth and impact that simulation methods have today.

While the most simple MC and MD methods have been invented about 50 years ago (the celebration of the 50th anniversary of the Metropolis algorithm was held in 2003, the 50th anniversary of the Alder-Wainwright spectacular first discovery by MD of the (then unexpected) phase transition in the hard sphere fluid is due in 2007), even the "second generation" of scientists, who started out 30-40 years ago as "simulators" are now already the "old horses" of the field, either close to the end of their scientific career, or, in the best case, near it. Thus, we can clearly observe that the task of developing the computer simulation methodology is further taken over with vigor by the "third generation" of well-established younger scientists who have emerged in the field. Because two of the organizers of the school (KB, GC) do belong to the "old horse" category, it was clearly necessary to get an energetic younger

co-organizer involved (MF), and we also felt it was the appropriate time that the most senior experts need not give the main lectures of the school, but rather the younger generation who are now most actively driving forward the frontier of research. Of course, it was crucial to involve the very valuable experience and knowledge of our senior colleagues into the school as well, and we are very glad that so many of them have accepted our invitation to give one-hour seminars providing tutorial introductions to various advanced research topics, which is at the heart of the research interests of the speakers. In this way, it was possible to produce an exciting event on the forefront of research on computer simulation in condensed matter, in a very stimulating and interactive atmosphere, with plenty of fruitful discussions.

It is with great pleasure that we end this preface with several acknowledgments. This school, of which the lecture notes are collected here, could not have taken place without the generous support of the European Community under the Marie Curie Conference and Training Courses, Contract No. MSCF-CT-2003-503840. We are grateful to the coordinators of this program, Michel Mareschal and Berend Smit, for their help in securing this support. We also wish to thank the CECAM secretaries, Emmanuelle Crespeau and Emilie Bernard.

We thank the Ettore Majorana Foundation and Centre for Scientific Culture in Erice, Sicily, for providing their excellent facilities to hold this school, and also Giorgio Benedek, Director of the International School of Solid State Physics, for the opportunity to hold our school as its 34th course: for his enthusiastic support during the school, and for his personal scientific participation. We are particularly grateful to him for providing the beautiful facilities of Erice.

MF thanks Davide Calanca, INFM-S3, Modena, for his valuable help in setting up the Web site of the school.

We thank the director of the physics department of the University of Rome "La Sapienza", Guido Martinelli, and the Administrative Secretary of the Department, Mrs. Maria Vittoria Marchet and her assistant, Mrs. Maria Proietto, for helping us in the difficult duty of managing all the financial matters. Mrs. Fernanda Lupinacci deserves grateful appreciation for her devoted and untiring presence and skillful help in overcoming all practical difficulties related to the organizational needs, and for providing a hospitable atmosphere to all the participants.

We are very grateful to Daan Frenkel, Mike Klein, and Peter Nielaba for their very valuable input when setting up the scientific program of the school, to all the lecturers, for their willingness to engage in the endeavor, and to all the participants, for their engagement and enthusiasm.

Mauro Ferrario
May 2006 *Giovanni Ciccotti*
Kurt Binder

Contents

Computer Simulations of Supercooled Liquids
W. Kob 1

Numerical Simulations of Spin Glasses: Methods and Some Recent Results
A.P. Young 31

Dipolar Fluctuations in the Bulk and at Interfaces
V. Ballenegger, R. Blaak, and J.-P. Hansen 45

Theory and Simulation of Friction and Lubrication
M.H. Müser 65

Simulation of Nanodroplets on Solid Surfaces: Wetting, Spreading and Bridging
A. Milchev 105

Monte Carlo Simulations of Compressible Ising Models: Do We Understand Them?
D.P. Landau, B. Dünweg, M. Laradji, F. Tavazza, J. Adler, L. Cannavaccioulo, and X. Zhu 127

Computer Simulation of Colloidal Suspensions
H. Löwen 139

Phase Transitions of Model Colloids in External Fields
P. Nielaba, S. Sengupta, and W. Strepp 163

Computer Simulation of Liquid Crystals
M.P. Allen 191

Coarse-Grained Models of Complex Fluids at Equilibrium and Under Shear
F. Schmid 211

Mesoscopic Simulations of Biological Membranes
B. Smit, M. Kranenburg, M. M. Sperotto, and M. Venturoli 259

Microscopic Elasticity of Complex Systems
J.-L. Barrat 287

Mesoscopic Simulations for Problems with Hydrodynamics, with Emphasis on Polymer Dynamics
B. Dünweg 309

Polymer Dynamics: Long Time Simulations and Topological Constraints
K. Kremer 341

Reaction Kinetics of Coarse-Grained Equilibrium Polymers: A Brownian Dynamics Study
C.-C. Huang, H. Xu, F. Crevel, J. Wittmer, and J.-P. Ryckaert 379

Equilibration and Coarse-Graining Methods for Polymers
D.N. Theodorou 419

Drug-Target Binding Investigated by Quantum Mechanical/Molecular Mechanical (QM/MM) Methods
U. Rothlisberger and P. Carloni 449

Redox Free Energies from Vertical Energy Gaps: Ab Initio Molecular Dynamics Implementation
J. Blumberger and M. Sprik 481

Advanced Car–Parrinello Techniques: Path Integrals and Nonadiabaticity in Condensed Matter Simulations
D. Marx 507

Evolutionary Design in Biological Physics and Materials Science
M. Yang, J.-M. Park, and M.W. Deem 541

Monte-Carlo Methods in Studies of Protein Folding and Evolution
E. Shakhnovich 563

Index 595

Contents
LNP 703 Computer Simulations in Condensed Matter Systems: From Materials to Chemical Biology Volume 1

Introduction: Condensed Matter Theory by Computer Simulation
G. Ciccotti, K. Binder, and M. Ferrario

Introduction to Cluster Monte Carlo Algorithms
E. Luijten

Generic Sampling Strategies for Monte Carlo Simulation of Phase Behaviour
N.B. Wilding

Simulation Techniques for Calculating Free Energies
M. Müller and J.J. de Pablo

Waste-Recycling Monte Carlo
D. Frenkel

Equilibrium Statistical Mechanics, Non-Hamiltonian Molecular Dynamics, and Novel Applications from Resonance-Free Timesteps to Adiabatic Free Energy Dynamics
J.B. Abrams, M.E. Tuckerman, and G.J. Martyna

Simulating Charged Systems with ESPResSo
A. Arnold, B.A.F. Mann, and Christian Holm

Density Functional Theory Based Ab Initio Molecular Dynamics Using the Car-Parrinello Approach
R. Vuilleumier

Large Scale Condensed Matter Calculations using the Gaussian Augmented Plane Waves Method
J. VandeVondele, M. Iannuzzi, and J. Hutter

Computing Free Energies and Accelerating Rare Events with Metadynamics
A. Laio and M. Parrinello

Transition Path Sampling Methods
C. Dellago, P.G. Bolhuis, and P.L. Geissler

Sampling Kinetic Protein Folding Pathways using All-Atom Models
P.G. Bolhuis

Calculation of Classical Trajectories with Boundary Value Formulation
R. Elber

Transition Path Theory
E. Vanden-Eijnden

Multiscale Modelling in Molecular Dynamics: Biomolecular Conformations as Metastable States
E. Meerbach, E. Dittmer, I. Horenko, and C. Schütte

Transport Coefficients of Quantum-Classical Systems
R. Kapral and G. Ciccotti

Linearized Path Integral Methods for Quantum Time Correlation Functions
D.F. Coker and S. Bonella

Ensemble Optimization Techniques for Classical and Quantum Systems
S. Trebst and M. Troyer

The Coupled Electron-Ion Monte Carlo Method
C. Pierleoni and D.M. Ceperley

Path Resummations and the Fermion Sign Problem
A. Alavi and A.J.W. Thom

List of Contributors

J. Adler
Department of Physics
Technion-Israel Institute
of Technology
Haifa, Israel
phr76ja@tx.technion.ac.il

Michael P. Allen
Department of Physics
and Centre for
Scientific Computing
University of Warwick
Coventry CV4 7AL, U.K.
m.p.allen@warwick.ac.uk

Vincent Ballenegger
Laboratoire de Physique
Moléculaire
UMR CNRS 6624
Université de Franche-Comté
La Bouloie
25030 Besançon cedex, France
vincent.ballenegger@univ-fcomte.fr

Jean-Louis Barrat
Laboratoire de Physique de
la Matière Condensée et
Nanostructures
Université Claude Bernard
Lyon I and CNRS
6 rue Ampère
69622 Villeurbanne Cedex, France
barrat@lpmcn.univ-lyon1.fr

Ronald Blaak
Department of Chemistry
Lensfield Road
Cambridge CB2 1EW, U.K.
rb419@cam.ac.uk

Jochen Blumberger
Center for Molecular Modeling
and Department of Chemistry
University of Pennsylvania
231 S. 34th Street
Philadelphia, PA 19104-6323
jochen@cmm.upenn.edu

L. Cannavaccioulo
IFF, Forschungszentrum Juelich
Juelich, Germany
l.cannavacciuolo@fz-juelich.de

François Crevel
Institut Charles Sadron
6 Rue Boussingault
67083 Strasbourg, France
crevel@ics.u-strasbg.fr

Michael W. Deem
Rice University
6100 Main Street–MS 142
Houston, TX
77005-1892 U.S.A.
mwdeem@rice.edu

Burkhard Dünweg
Max-Planck-Institut für
Polymerforschung
Ackermannweg 10
55128 Mainz, Germany
duenweg@mpip-mainz.mpg.de

Jean-Pierre Hansen
Department of Chemistry
Lensfield Road
Cambridge CB2 1EW, U.K.
jph32@cam.ac.uk

Chien-Cheng Huang
Physique des polymères
CP223, Université
Libre de Bruxelles
Bv du Triomphe
1050 Brussels, Belgium
chien-cheng.huang@ulb.ac.be

Walter Kob
Laboratoire des Colloïdes
Verres et Nanomatériaux
UMR 5587, CNRC
Université Montpellier II
Place E. Bataillon
34095 Montpellier, France
kob@lcvn.univ-montp2.fr

Marieke Kranenburg
Van 't Hoff Institute
for Molecular Sciences
University of Amsterdam
Nieuwe Achtergracht 166
1018 WV Amsterdam
The Netherlands
marieke@science.uva.nl

Kurt Kremer
Max-Planck-Institut für
Polymerforschung
Ackermannweg 10
55128 Mainz, Germany
kremer@mpip-mainz.mpg.de

D.P. Landau
Center for Simulational Physics
The University of Georgia
Athens GA 30622, U.S.A.
dlandau@hal.physast.uga.edu

M. Laradji
Dept. of Physics
University of Memphis
Memphis, TN, U.S.A.
mlaradji@memphis.edu

Hartmut Löwen
Heinrich-Heine-Universität
Düsseldorf
Universitätsstraße 1
40225 Düsseldorf, Germany
hlowen@thphy.uni-duesseldorf.de

Dominik Marx
Lehrstuhl für Theoretische Chemie
Ruhr–Universität Bochum
44780 Bochum, Germany
dominik.marx@theochem.ruhr-uni-bochum.de

Andrey Milchev
Institute of Physical Chemistry
Bulgarian Academy of Sciences
G. Bonchev Str. Block 11
1113 Sofia, Bulgaria
milchev@ipc.bas.bg

Martin H. Müser
Department of Applied Mathematics
University of Western Ontario
London, ON, Canada N6A 5B7
mmuser@uwo.ca

P. Nielaba
Physics Department
University of Konstanz
78457 Konstanz, Germany
peter.nielaba@uni-konstanz.de

Jeong-Man Park
Rice University
6100 Main Street–MS 142
Houston, TX, 77005-1892 U.S.A.
and
Department of Physics
The Catholic University of Korea,
Puchon 420–743, Korea
jmanpark@catholic.ac.kr

Jean-Paul Ryckaert
Physique des polymères
CP223, Université Libre
de Bruxelles
Bv du Triomphe
1050 Brussels, Belgium
jryckaer@ulb.ac.be

Friederike Schmid
Fakultät für Physik
Universität Bielefeld
Universitätsstraße 25
33615 Bielefeld
schmid@physik.uni-bielefeld.de

S. Sengupta
S.N. Bose National Center
for Basic Sciences
Calcutta 700098, India
surajit@bose.res.in

Eugene Shakhnovich
Department of Chemistry
and Chemical Biology
Harvard University
12 Oxford Street
Cambridge MA 02138, U.S.A.
Shakhnovich@chemistry.harvard.edu

Berend Smit
Centre Européen de Calcul
Atomique et Moleculaire (Cecam)
Ecole Normale Supérieure
46 Allée d'Italie
69007 Lyon, France
and
Van 't Hoff Institute
for Molecular Sciences
University of Amsterdam
Nieuwe Achtergracht 166
1018 WV Amsterdam
The Netherlands
b.smit@cecam.fr

Maria Maddalena Sperotto
Biocentrum
The Technical University
of Denmark
Kgs. Lyngby, Denmark.
maria@cbs.dtu.dk

Michiel Sprik
Department of Chemistry
University of Cambridge
Cambridge CB2 1EW, U.K.
ms284@cus.cam.ac.uk

W. Strepp
Physics Department
University of Konstanz
78457 Konstanz, Germany
strepp@hamilton.physik.uni-konstanz.de

F. Tavazza
Center for Simulational Physics
The University of Georgia
Athens GA 30622, U.S.A.
ftavazza@nist.gov

Doros N. Theodorou
School of Chemical Engineering
National Technical University of Athens
9 Heroon Polytechniou Street
Zografou Campus
157 80 Athens, Greece
doros@central.ntua.gr

Maddalena Venturoli
CCS, University College
C. Ingold Labs
20 Gordon St.
London WC1H 0AJ, U.K.
m.venturoli@ucl.ac.uk

Joachim Wittmer
Institut Charles Sadron
6 Rue Boussingault
67083 Strasbourg, France
jwittmer@ics.u-strasbg.fr

Hong Xu
LPMD, Inst. Physique-Electronique
University Paul Verlaine-Metz
1bd Arago
57078 Metz cedex 3, France
hongxu@univ-metz.fr

Ming Yang
Rice University
6100 Main Street–MS 142
Houston, TX, 77005-1892 U.S.A.
myang@rice.edu

A. Peter Young
Physics Department
University of California Santa Cruz
Santa Cruz, CA 95064
peter@bartok.ucsc.edu

Computer Simulations of Supercooled Liquids

W. Kob

Laboratoire des Colloïdes, Verres et Nanomatériaux, UMR 5587, CNRS Université Montpellier II, Place E. Bataillon, 34095 Montpellier, France
kob@lcvn.univ-montp2.fr



Walter Kob

W. Kob: *Computer Simulations of Supercooled Liquids*, Lect. Notes Phys. **704**, 1–30 (2006)
DOI 10.1007/3-540-35284-8_1 © Springer-Verlag Berlin Heidelberg 2006

1 Introduction 3
2 Salient Features of Glass-Forming Systems 4
2.1 Theoretical Approaches 11
3 Computer Simulations of Glass-Forming Liquids 18
4 Conclusions 28
References 28

We review some of the salient features of glass-forming systems. After a brief discussion of the theoretical approaches that are used to describe the slowing down of the dynamics of these system, notably the theory of Adam and Gibbs and the mode-coupling theory of the glass transition, we present some results of computer simulations that have been done to check the validity of these theories.

1 Introduction

The phenomenon of the glass transition is a feature that can be observed in a multitude of complex systems. The paradigm of this transition is of course the freezing of a glass-forming liquid such as silica, SiO_2, into an amorphous solid if the melt is cooled with a sufficiently high cooling rate, e.g. faster than a few K/s.[1] However, it is found that also many other systems can be frozen into a disordered structure, i.e. a configuration that is very different from the one of the ground state, if an external parameter (temperature, magnetic field, chemical potential, etc.) is changed sufficiently quickly, and prominent examples include polymers, colloidal suspensions, metals, superconductors, proteins, and optimization problems. All these systems have one key feature in common: The tendency to optimize the structure on the local scale with the result that on larger scales the structure is frustrated and therefore disordered. Examples are the packing of spheres for which the densest *local* packing is an icosahedron, a geometrical structure that does not allow to fill space in a regular manner. Another important example are commissions in which the members have to find a consensus on a given subject, i.e. optimize a "cost function": It is easy to find a good solution for a small number of people, but very often impossible to find one that satisfies everyone. (Therefore one has the golden rule for an efficient functioning of a commission that the number of members has to be odd and smaller than three.)

The presence of this frustration on large length scales (or with a large number of degrees of freedom) makes it very difficult to come up with a reliable theoretical description that goes beyond mean field concepts and therefore progress in the theory of disordered systems has been very slow. Therefore it is presently not possible to present a coherent theory that is able to describe the multitude of features that can be found in glass-forming systems. Thus the goal of the present text is to give just a concise overview of the typical features of glass-forming systems, to discuss briefly some of the theoretical approaches that have been proposed to describe these systems, and to review some of the numerical simulations that have been done to test these theories.

[1] This rate is in fact not very high and therefore silica is considered to be a "good" glass former. For other systems, such as mixtures of metals with only 2–3 components, one needs rates on the order of 10^6 K/s in order to avoid crystallization.

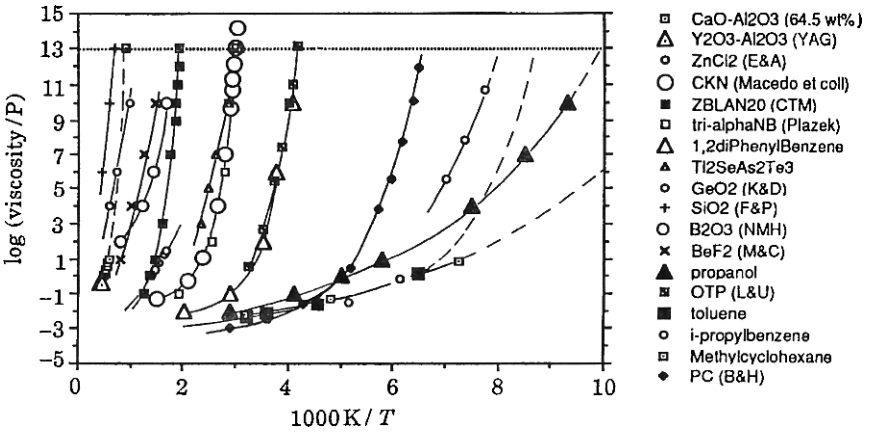


Fig. 1. Arrhenius plot of the temperature dependence of the viscosity of various glass-forming liquids. From [9]

2 Salient Features of Glass-Forming Systems

In this section we will discuss some of the main features of glass-forming systems. In order to keep the language simple we will restrict ourselves to the case of structural glasses and thus will not discuss other important glassy systems such as spin glasses, protein folding, or optimization problems. In addition we will be mainly concerned with the *temperature* dependence of the properties of these systems, even if a similar behavior can be found if, e.g., one changes the pressure. More extensive discussions of these subjects can be found in [1–8].

At sufficiently high temperatures the viscosity of liquids is relatively small, i.e. η is of the order 10^{-2} Pa s, the viscosity of water at room temperature. Experimentally it is found that with decreasing temperature the viscosity increases rapidly and can reach values of the order of 10^{14}Pa s and more. This is demonstrated in Fig. 1 in which we show the logarithm of η as a function of inverse temperature for many different types of liquids. The interaction between the atoms in these various liquids can be of very different nature (van der Waals, ionic,...) and therefore the relevant temperature scale depends also strongly on the liquid and makes a comparison of the different curves somewhat difficult. In order to overcome this problem it is customary to scale the T-axis by T_g, a temperature that is defined by requiring that $\eta(T_g) = 10^{12}$Pa s [2] [10,11]. The resulting graph, which is usually called "Angell-plot", is shown in Fig. 2 from which we now recognize more easily that the T-dependence of η is indeed very strong but still smooth. Some of the curves (the

[2] The reason for choosing this somewhat arbitrary value is related to the fact that such a viscosity usually corresponds to a relaxation time that is on the order of 10s, i.e. a time scale that is well adapted to human beings.

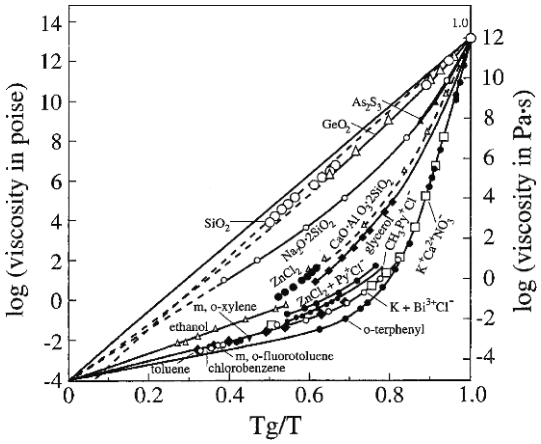


Fig. 2. Viscosity of various glass-forming liquids as a function of rescaled temperature T_g/T. After [11]

one at the bottom of the graph) seem to show a crossover from an Arrhenius dependence with a small activation energy at high T to an Arrhenius-law with a higher activation energy at low T. Liquids that show such a behavior are called "fragile glass-formers" [11]. On the other hand there are liquids that seem to show in the whole accessible temperature range an Arrhenius-law with the same activation energy (curves at the top) and these materials are called "strong glass-formers" [11].

Although Fig. 2 clearly allows to distinguish between strong and fragile glass-formers, it is presently not clear whether the different shapes of the $\eta(T)$ curves is not only due to the way the data is represented [12, 13]. E.g. Hess et al. have proposed a different type of scaling plot in which one scales the T-axis not only by T_g but in addition by a factor F that makes that the *slope* of the curves at $T = T_g$ are the same. The resulting plot is shown in Fig. 3 and it demonstrates that it is indeed possible to obtain in the regime of high η a nice collapse of the data onto a master curve, the shape of which is slightly bend. Thus this is evidence that in fact *all* $\eta(T)$ curves are bend and that the reason that in Fig. 2 this is not seen very well is just a problem of the representation of the data which makes that the high T data is squeezed in a small interval of the abscissa.

The temperature above which the data no longer falls onto this master curve has been found to be very close to the so-called critical temperature of mode-coupling theory of the glass transition (MCT) [14–17]. As we will discuss in more detail below, this theory is indeed able to rationalize many features of the dynamics of glass-forming liquids and can presently be regarded as the most successful theoretical approach to describe the dynamics of glass-forming liquids.

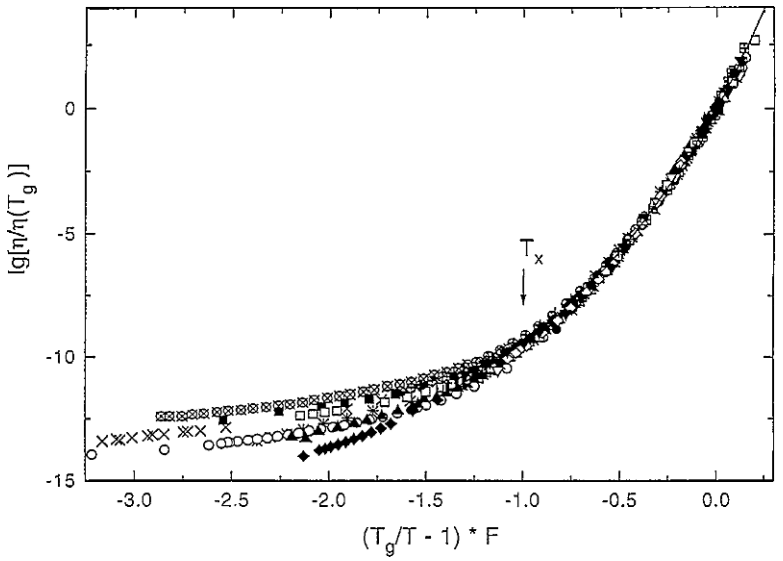


Fig. 3. The viscosity of various glass-forming systems (silica, B_2O_3, salol, propylene carbonate, CKN ($40Ca(NO_3)_260KNO_3$), etc.). T_g is the calorimetric glass transition temperature. The data is plotted as a function of $(T_g/T-1)\cdot F$ with $F = T_x/(T_x - T_g)$. Here T_x is a temperature that has been adjusted for each liquid such that the viscosity data falls at low T on a master curve. Adapted from [13]

From Fig. 2 it also becomes clear that the T-dependence of η does not change if the temperature crosses the melting temperature T_m. Although the latter is not included in the graph, there is the rule of thumb that $T_g/T_m \approx 2/3$, i.e. in the figure T_m is located between 0.6 and 0.7 T_g/T. Since the $\eta(T)$ curves are very smooth, it is evident that the mechanism responsible for the dramatic slowing down of the dynamics cannot be related to the fact that the liquid is supercooled. In fact, the case of SiO_2 shows that this liquid is already very viscous *at* its melting temperature ($T_m = 2000\,\mathrm{K}$ and $T_g = 1450\,\mathrm{K}$, thus $T_g/T_m = 0.725$ and $\eta(T_m) \approx 10^7$ Pa s). This is the reason why it is not really appropriate to speak in the context of glass-forming liquids as the dynamics of "super-cooled liquids". Instead one prefers to use the terms "glassy liquids" to emphasize their slow dynamics.

Figure 2 allows to recognize immediately *the* relevant question in the field: What is the reason for the strong slowing down of the dynamics (at least 15 decades!) upon a relatively mild change in temperature? The puzzling observation is that although the quantities characterizing the dynamics, such as the viscosity discussed here, or the relaxation times introduced below, show a very strong T-dependence, whereas the thermodynamic quantities (density, specific heat,...) or structural quantities (structure factor, angular distribution functions,...) show only a very mild T-dependence. This can be recognized, e.g., from Fig. 4 which shows the static structure factor of a binary

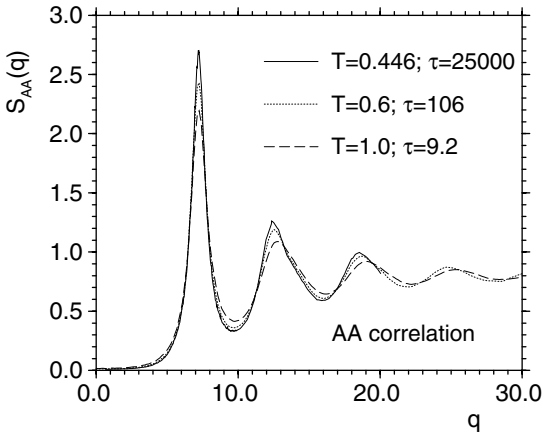


Fig. 4. Wave-vector dependence of the partial structure factor $S_{\mathrm{AA}}(q)$ for a mixture of Lennard-Jones particles. The different curves correspond to different temperatures. Also included are the α-relaxation times τ (in reduced units)

Lennard-Jones system at various temperatures. Although the relaxation time τ, given in the figure, changes by orders of magnitude, the structure factor shows only a very weak dependence on T in that with decreasing temperature the peaks become slightly more pronounced. This weak T-dependence of the thermodynamic and structural quantities is in stark contrast to the case of second order phase transitions, another phenomenon in which one finds a strong increase of the relaxation times, and poses a real challenge to come up with a theoretical description that is able to describe the slowing down of the dynamics.

The form of the temperature dependence of η, or τ, is another important question. At low temperatures the experimental data can often be fitted very well by the so-called "Vogel-Fulcher-Tammann-law" which is given by

$$\eta(T) = \eta_0 \exp[A/(T - T_0)] \,, \tag{1}$$

where T_0 is the temperature at which the extrapolation (!) of the viscosity seems to diverge, and which is called "Vogel-temperature".

In order to characterize the slowing down of the dynamics on a more microscopic level, it is useful to consider the time and temperature dependence of time correlation functions such as the van Hove function or the intermediate scattering function [8,18,19]. The typical time dependence of such a correlator Φ is shown in Fig. 5 which shows $\Phi(t)$ vs. $\log(t)$. Let us first discuss the correlator at high temperatures. At very short times $\Phi(t)$ shows a quadratic time dependence, which can be understood immediately by recalling that at short times the trajectories of the particles can be expanded in a Taylor series in time and that, due to the time reflection symmetry of Newton's equation of motion, any time correlation function must be even in time. After this

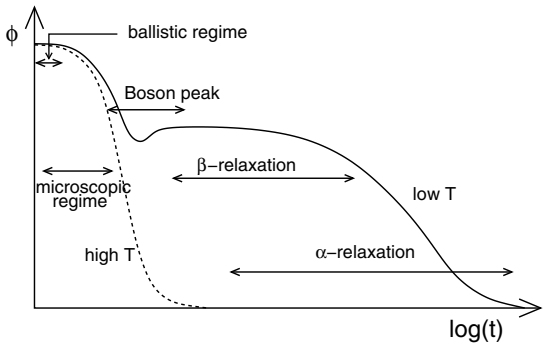


Fig. 5. Typical time dependence of a time correlation function. The curve to the left corresponds to the case of a normal liquid and the one to the right to a glassy liquid

"ballistic regime" the curve decays to zero very quickly (for atomic liquids on the time scale of pico-seconds) and the time dependence of the correlator can be approximated well by an exponential function, i.e. a Debye-law. This final decay is usually denoted by "α-relaxation".

Also at low temperatures the correlator show a ballistic regime at short times. However, for intermediate times $\Phi(t)$ shows a plateau, i.e. on this time scale the correlation function does not decay. The reason for this is that each particle is (temporarily) trapped in the cage formed by its neighbors, a phenomenon that is called "cage-effect" and which is in (rough) analogy to the situation found in a crystal. Although in this time window the time dependence of the correlator is weak, it is not negligible and it is usually called "β-relaxation". At the beginning of the mentioned plateau one often sees some strongly damped oscillations which are due to low frequency vibrational modes in the system. Although this feature is not found in all glass-forming systems, it is quite common and is referred to as the so-called "Boson-peak", since in the dynamic susceptibility it is seen as a peak [20][3] At these temperatures the α-relaxation is no longer given by a Debye-law. Although the exact functional form is not really known, the so-called stretched exponential, or Kohlrausch-Williams-Watts function (KWW), given by

$$\Phi(t,T) = A(T)\exp[-(t/\tau(T))^{\beta(T)}]\ , \tag{2}$$

gives often a very good description of the correlator. Here $A(T)$, $\beta(T)$, and $\tau(T)$ are fit parameters. The quantity $\tau(T)$ is often used to define a relaxation time, although other definitions, e.g. the area under the correlator, can be used as well.

[3] The dynamics susceptibility $\chi''(\omega)$ of a time correlation function $\Phi(t)$ is given by the imaginary part of the time Fourier transform of $\Phi(t)$ multiplied by the frequency ω: $\chi''(\omega) = \frac{\omega}{k_B T}\Phi''(\omega)$ [18].

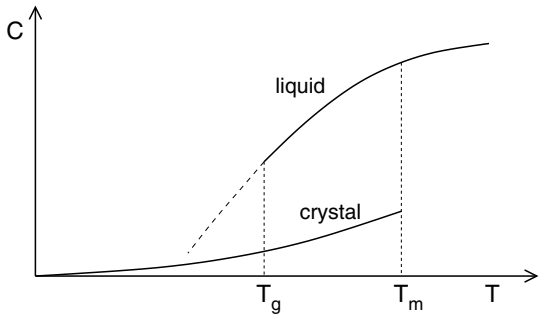


Fig. 6. Schematic plot of the temperature dependence of the specific heat of a glass-forming liquid (*upper curve*). The dashed line is an extrapolation of this curve to temperatures below the glass transition temperature T_g. The lower curve is the specific heat for the crystalline state of the same material that ends at the melting temperatures T_m

Above we have mentioned that structural and thermodynamic quantities show only a very mild variation with temperature. This observation applies also to the specific heat $C(T)$. Since at low T one has a separation of the time scale for the vibrational motion from the one for the motion giving rise to the α-relaxation, see Fig. 5, it is clear that $C(T)$ can be decomposed into two parts: One that is related to the vibrational dynamics inside the cage, $C_{\rm vib}(T)$, and another part that is related to the configurational degrees of freedom related to the flow of the particles, $C_{\rm conf}(T)$. It is found that $C_{\rm vib}(T)$ is only a mild function of temperature and quite similar to the specific heat found in a crystal, i.e. $C_{\rm vib}(T) \approx C_{\rm cryst}(T)$, see Fig. 6. Therefore it is possible to study the T-dependence of $C_{\rm conf} \approx C(T) - C_{\rm cryst}(T)$ and in a seminal paper Kauzmann showed that this quantity decreases rapidly with decreasing T [21]. This observation has an important consequence for the entropy of the system which can be obtained from the relation $S(T) = S(T_m) - \int_T^{T_m} \frac{C(T)}{T}\, dT$, where T_m is the melting temperature. Similar to the specific heat, $S(T)$ can be split into a vibrational and configurational part and hence one can determine $\Delta S(T)$, the difference between the entropy of the liquid and the one of the crystal.

The T-dependence of ΔS, normalized by its value at the melting temperature T_m, is shown schematically in Fig. 7. Thus we recognize that, at least for fragile glass-formers, this difference decreases rapidly with decreasing temperature and if one makes an extrapolation (!) to low T the difference seems to vanish at a *finite* temperature T_K, the so-called "Kauzmann temperature". The implication of this observation is that the dynamics of the system should come to an arrest for $T \to T_K^+$, since there are no more configurations into which the system can move. (Recall that we can expect that $S_{\rm conf} = k_B T \ln(\Omega_{\rm conf})$, where $\Omega_{\rm conf}$ is the number of accessible configurations.) Interestingly the experimental data shows that often T_K is very close

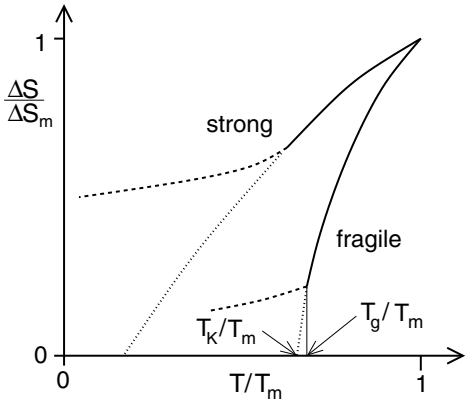


Fig. 7. Schematic plot of the temperature dependence of the difference between the entropy of a glass-forming liquid and the one of the corresponding crystal, normalized by its value at the melting point. The two *solid lines* correspond to the typical behavior of a strong and fragile glass-former. The *dashed lines* are the non-equilibrium data in the glass, whereas the *dotted lines* are an extrapolation of the equilibrium data to lower temperatures. These latter lines cross the abscissa at the Kauzmann temperature T_K

to T_0, the Vogel temperature, see (1), the temperature at which the dynamics seems to come to a halt [22]. Thus this gives evidence that there is a connection between the thermodynamic properties of glass-forming systems and their relaxation dynamics.

So far we have already encountered *three* relevant temperatures: *i*) The glass-transition temperature T_g at which the viscosity attains the (somewhat arbitrary) value of 10^{12} Pa s and which is a misnomer since at T_g does *not* undergo a glass transition; *ii*) The Vogel temperature T_0 at which the extrapolation of the T-dependence of the viscosity seems to predict a divergence; *iii*) The Kauzmann temperature T_K at which the extrapolation of the configurational part of the specific heat seems to go to zero. A further important temperature is T_c, the critical temperature of MCT at which the transport properties of the glassy liquid changes their nature, see below for details. Last not least there is the (kinetic) glass transition temperature $T_{\rm kin}$, but which is also sometimes denoted by T_g (!), at which the system falls out of equilibrium. The reason for this is that with decreasing temperature the relaxation times increase rapidly, see Fig. 2. Thus if one wants to equilibrate a system at a given temperature T (e.g. by performing a quench with a given cooling rate), it is usually necessary to wait for a time $\tau(T)$ in order to allow the system to search for the configurations of particles that are typical at this temperature. Before this process is not finished the system is out of equilibrium and should therefore be termed "glass". Thus $T_{\rm kin}$ is nothing else than the temperature at which the relaxation time $\tau(T)$ of the system crosses the time scale of the experiment. (Note that the latter can be very short, such as in laser melting

experiments which are on the order of micro seconds, or very large, such as cooling of lava on the time scales of thousands of years.)

In the examples we have discussed so far, the slowing down of the dynamics was due to the decrease in temperature. This is, however, by no means the only possibility to induce a slow dynamics. In colloidal gels the interaction of the particles is changed by the addition of polymers, making that the depletion forces increase. In chemical gels one adds molecules that serve as crosslinkers between polymer chains. In atomic or molecular liquids an increase of pressure, at constant T, will also give rise to a slow dynamics. Furthermore one can generate *glasses* by vapor depositions upon a cold substrate, by a chemical reaction (sol-gel transition), irradiation of a crystal by heavy ions, and many more (see [9] for a more exhaustive list). All these processes lead to an amorphous structure that is relatively similar to the one of a liquid, but no longer flows. One must realize, however, that, *even for the same material!*, the details of the structure will depend on its production history and therefore it is not appropriate to refer to the "structure of the glass" without having specified how this glass has been obtained.

As summary from this section we can conclude that glassy systems show in general the following features: Very strong dependence of the relaxation dynamics on an external parameter (temperature, pressure, external fields, chemical potential,...); no obvious presence of long range order; some kind of frustration that prevents to find the optimal structure; stretching of the time correlation functions.

Although we have restricted ourselves here to the case of structural glasses, it is evident that many other systems will have the features of this list. Examples are spin glasses, the vortex lines in superconductors, granular materials, foams, protein folding, optimization problems, etc. This large variety of systems explains why the field of glass physics is currently so popular since an advance in any one of these subclasses of systems is likely to lead also to a breakthrough in the other subclasses as well.

2.1 Theoretical Approaches

Due to the importance of glassy systems in science and applications, many theoretical approaches have been proposed that aim to rationalize the above mentioned key features of these systems (free volume theory, Adam-Gibbs theory, mode-coupling theory, replica theory, theories based on dynamic facilitation or traps, etc.). Due to the lack of space we will not be able to discuss here all these theories and refer instead to some of the relevant literature in which more details can be found [2, 5, 6, 8, 14]. Instead we will focus here on two popular theories: The one by Adam and Gibbs and the mode-coupling theory of the glass transition. However, also for these cases the discussion will be rather brief.

One of the oldest theories for the glass transition is the one proposed by Adam and Gibbs already in 1965 [23]. The basic physical ingredient of

this theory is the idea that in a glassy system the particles are not able to leave the cage formed by their neighbors, i.e. to do the α-relaxation, without that the particles forming the cage allow/help to do so. Thus on the time scale of the α-relaxation the motion of the particles must be expected to be highly cooperative. Adam and Gibbs assumed that it is possible to define regions in space in which the particles do indeed behave cooperatively (the so-called "cooperatively rearranging regions", CRR) and which are the relevant entities for the α-relaxation. Of course these CRRs constitute only domains that have a finite life time, i.e. are created via a density fluctuation in space and subsequently disappear on the time scale of the α-relaxation. (That the precise nature of these density fluctuations is not known is one of the drawback of this theory.) It is assumed that these regions are *independent* and sufficiently large, say more than O(20) particles, so that one can treat each of them as a small system in which the laws of statistical mechanics already apply.

Under these assumptions it is easy to show [5, 8, 23] that the probability that a CRR of size z makes a cooperative rearrangement, is given by

$$W(z,T) = A \exp(-\beta z \delta\mu) \; , \tag{3}$$

where A is a constant, $\beta = 1/k_B T$, and $\delta\mu$ is related to the *effective* energy barrier that the CRR has to overcome in order to relax, i.e. to go to a new local state. At sufficiently low temperatures we have $\beta\delta\mu \gg 1$ and therefore the most important contribution to the relaxation dynamics of the *whole* system will come from the CRRs that have the lowest number of particles. Note that this lowest number, which we will call z^*, will be significantly larger than 1, since the motion is cooperative. In order to estimate z^*, Adam and Gibbs used the fact that due to the separation of the time scales for the vibrational dynamics and the α-relaxation, also the entropy can be split into a part $S_{\rm vib}$ that is related to the vibrations and a part $S_{\rm conf}$ that is related to the exploration of new configurations, i.e. the α-relaxation (see the above discussion on the Kauzmann temperature). We now assume that the whole system, which has N particles, can be decomposed into distinct cooperatively rearranging regions of size z^*. Thus the number of such regions is given by $n(z^*,T) = N/z^*$. If the configurational entropy of the whole system is $S_{\rm conf}$, the one of the subsystems are given by $s_{\rm conf} = S_{\rm conf}/n(z^*,T)$. (Note that $s_{\rm conf}$ must be of order $k_B \ln 2$, since a subsystem must have at least two distinct configurations.) Thus we find that the size of a cooperatively rearranging region is given by

$$z^* = \frac{N}{n(z^*,T)} = \frac{N s_{\rm conf}}{S_{\rm conf}} \; . \tag{4}$$

Together with (3) we thus obtain

$$W(T) = A \exp\left[-\frac{\beta N s_{\rm conf}\delta\mu}{S_{\rm conf}}\right] = A \exp\left[-\frac{C}{T S_{\rm conf}}\right] \; , \tag{5}$$

where C is a constant.

If we assume that the relaxation time $\tau(T)$, or the viscosity $\eta(T)$, is proportional to $W(T)^{-1}$, we thus obtain immediately the final result of the theory of Adam and Gibbs:

$$\tau(T) \propto \eta(T) \propto \exp\left[\frac{C}{TS_{\text{conf}}}\right] , \tag{6}$$

i.e. that the vanishing of the configurational entropy leads to a divergence of the relaxation time.

Following the procedure discussed above in the context of the Kauzmann temperature, measurements of the specific heat will allow to determine the configurational entropy S_{conf}. Therefore the relation given by (6) can be tested. It is found that, at least for fragile glass-formers, a plot of the logarithm of the relaxation time as a function of $1/TS_{\text{conf}}$ does indeed give a straight line for a rage of $\tau(T)$ that covers many decades [22,24]. Thus this is evidence that the theory of Adam and Gibbs, as ad hoc and approximate as it seems, does indeed capture some of the relevant physics that is responsible for the slow dynamics at low temperatures.

Last not least we mention that if one makes the assumption that the specific heat of a liquid shows a T-dependence of the form K/T, where K is a constant, a behavior that can indeed be found in fragile glass-forming liquids [25], one can easily calculate the entropy of the system. Putting the so obtained result into the Adam-Gibbs expression, (6), one finds easily that

$$\tau(T) \propto \exp\left[\frac{CT_K/K}{T - T_K}\right] , \tag{7}$$

which is nothing else that the Vogel-Fulcher-law from (1), *if* one identifies the Vogel temperature T_0 with the Kauzmann temperature T_K. Experimental data shows that often this predicted equality of T_K and T_0 does indeed hold [22,26,27] and therefore gives further support to the theory.

Although the theory of Adam-Gibbs looks quite appealing and seems to be able to rationalize some experimental findings, it has several caveats: First of all the nature of the CRRs is never specified and no prescription is given how these regions can be measured in an experiment or computer simulation. Secondly, it is far from obvious that the CRRs can indeed be considered as *independent* entities, i.e. that their interaction can indeed be neglected. Last not least it is doubtful that one can indeed focus only on *one* size of CRR, i.e. a fixed z^*. This approximation might (perhaps!) be appropriate at *very* low temperatures, but the experiments show that often the Adam-Gibbs relation holds already at temperatures where the dynamics of the liquid is not yet very slow and where it is unlikely that the condition $\beta\delta\mu \gg 1$ really holds, in which case it would be necessary to take into account also other types of CRRs.

A very different approach for the description of the dynamics of glassy liquids is given by the mode-coupling theory of the glass transition (MCT).

This theory is based on the assumption that all the structural and thermodynamical properties of a glass-forming liquid show only a smooth variation with temperature. Using the formalism by Zwanzig and Mori, described below, it is possible to derive (approximative) equations that give the time and wave-vector dependence of the space-time correlation functions. Due to the T-dependence of the structure, the solution of these equations show a singular behavior in that they predict, among many other things, a rapid increase of the relaxation times. Due to the complexity of MCT as well as the multitude of the predictions it makes it is not possible to give here a complete description of the theory. Instead we refer the reader to the more specialized literature [8, 14–17].

Let us start with a *brief* reminder of the Zwanzig-Mori projection operator formalism. A more complete discussion can be found in [18, 28–31].

We consider a system of N classical particles. The equations of motion of an arbitrary phase space function g is given by

$$\dot{g} = i\mathcal{L}g = \{H, g\} \; , \tag{8}$$

where $\mathcal{L}$ is the Liouville operator, H is the Hamiltonian of the system, and $\{.\,,.\}$ are the Poisson brackets.

The set of all possible phase space functions form a vector space in which one can define a scalar product between two vectors g and h as follows:

$$(g|h) = \langle \delta g^* \delta h \rangle \; , \tag{9}$$

where $\langle . \rangle$ is the usual canonical average, and $\delta g = g - \langle g \rangle$, i.e. we are considering only the fluctuating part of g and h.

We now assume that due to the physical properties of the system considered we can identify k phase space variables that are "slow", i.e. their time dependence is significantly slower than the one of the other variables. We will denote these slow variables $A_1, A_2, \ldots, A_k$ and collect them into a column vector $\mathbf{A}$.

We now define an operator $\mathcal{P}$ that projects an arbitrary function g on the space spanned by the subset A_n:

$$\mathcal{P}g = (\mathbf{A}|g)(\mathbf{A}|\mathbf{A})^{-1}\mathbf{A} = \sum_{n,m} |A_n)[(A_i|A_j)]^{-1}_{nm}(A_m|g) \; , \tag{10}$$

where $[(A_i|A_j)]^{-1}_{nm}$ is the n, m element of the inverse of the matrix $(A_i|A_j)$. It is easy to verify that $\mathcal{P}$ and $1-\mathcal{P}$ are projectors, i.e. that they satisfy $\mathcal{P}^2 = \mathcal{P}$, $(1-\mathcal{P})^2 = 1-\mathcal{P}$, and that $\mathcal{P}(1-\mathcal{P}) = (1-\mathcal{P})\mathcal{P} = 0$.

Some straightforward algebraic manipulations allow then to show that

$$d\mathbf{A}/dt = \exp(i\mathcal{L}t)[\mathcal{P} + (1-\mathcal{P})]i\mathcal{L}\mathbf{A} \tag{11}$$

$$= i\mathbf{\Omega} \cdot \mathbf{A}(t) + \exp(i\mathcal{L}t)(1-\mathcal{P})i\mathcal{L}\mathbf{A} \; , \tag{12}$$

where the "frequency matrix" $i\mathbf{\Omega}$ is defined by $i\mathbf{\Omega} = (\mathbf{A}|i\mathcal{L}\mathbf{A}) \cdot (\mathbf{A}|\mathbf{A})^{-1}$.

A simple but somewhat lengthy calculation [8,30] allows to transform this equation of motion for $\mathbf{A}$ into the following one:

$$\dot{\mathbf{A}} = i\boldsymbol{\Omega} \cdot \mathbf{A}(t) - \int_0^t d\tau \mathbf{M}(\tau) \cdot \mathbf{A}(t-\tau) + \mathbf{f}(t) \; , \tag{13}$$

Here $\mathbf{f}(t)$ is the so-called "fluctuating force" and is given by

$$\mathbf{f}(t) = \exp[i(1-\mathcal{P})\mathcal{L}t] i(1-\mathcal{P})\mathcal{L}\mathbf{A} \; , \tag{14}$$

and the matrix $\mathbf{M}(t)$ is the "memory function" which is given by

$$\mathbf{M}(t) = (\mathbf{f}|\mathbf{f}(t)) \cdot (\mathbf{A}|\mathbf{A})^{-1} \; . \tag{15}$$

Thus we see that the equation of motion for $\mathbf{A}$ has been transformed into a generalized Langevin equation, (13), with $\mathbf{f}(\mathbf{t})$ playing the role of the fluctuating forces.

To obtain an equation of motion for the time correlation functions $\mathbf{C}(t) = \langle \mathbf{A}^*(0)\mathbf{A}(t) \rangle$, note that this is a $k \times k$ matrix, we take the scalar product of (15) with $\mathbf{A}$ and use the fact that $(\mathbf{A}|\mathbf{f}(t)) = 0$ [8,30]. This gives

$$\dot{\mathbf{C}}(t) = i\boldsymbol{\Omega} \cdot \mathbf{C}(t) - \int_0^t d\tau \mathbf{M}(\tau) \cdot \mathbf{C}(t-\tau) \; . \tag{16}$$

We emphasize that this equation for $\mathbf{C}(t)$ is *exact*, since no approximations have been made. Its solution can be readily obtained by introducing the Laplace transformed quantities

$$\widehat{\mathbf{C}}(z) = i \int_0^\infty dt \exp(izt) \mathbf{C}(t) \text{ and } \quad \widehat{\mathbf{M}}(z) = i \int_0^\infty dt \exp(izt) \mathbf{M}(t) \; , \tag{17}$$

from which one obtains

$$\widehat{\mathbf{C}}(z) = -\left[z\mathbf{I} + \boldsymbol{\Omega} - i\widehat{\mathbf{M}}(z) \right]^{-1} \cdot \mathbf{C}(0) \; , \tag{18}$$

where $\mathbf{I}$ is the unit matrix.

Although (18) gives a formal solution for the correlator $\mathbf{C}(t)$, in reality the problem is not yet solved since we have no explicit expression for the highly complex time dependence of $\mathbf{M}(t)$ which is given by (14) and (15). The basic idea is now to approximate $\mathbf{M}(t)$ by a simple time dependent function (e.g. an exponential) or by a functional of the correlator $\mathbf{C}(t)$. In both cases (17) becomes a self-contained integro-differential equation for $\mathbf{C}(t)$ which can be solved by analytical techniques or numerically.

Note that due to the generality of the Zwanzig-Mori formalism it is possible to apply it also to the memory function itself, i.e. to derive an equation of motion for $\mathbf{M}(t)$ that has exactly the same form as (17) and which involves the "second-order memory function", i.e. the memory function of the memory

function. In the context of the dynamics of glass-forming liquids it has indeed been found to be useful to follow this approach. In this case the slow variables, i.e. the A_i, are chosen to be the coherent intermediate scattering function $F(q,t)$, i.e. the density correlator for wave-vector q:

$$F(q,t) = \frac{1}{N}\langle \delta\rho(\mathbf{q},t)\delta\rho^*(\mathbf{q},t)\rangle \ . \tag{19}$$

Here we have introduced the density fluctuations

$$\delta\rho(\mathbf{q},t) = \sum_{j=1}^{N} \exp[i\mathbf{q}\cdot\mathbf{r}_j(t)] \ . \tag{20}$$

Note that $F(q,t)$ is a space-time correlation function that can directly be measured in light- or neutron scattering experiments and is also one of the key quantities in the theory of liquids. Therefore it is highly interesting to have a theory at hand that makes prediction for this correlator, and MCT is such a theory.

As mentioned above, the Zwanzig-Mori formalism requires to have an explicit expression for the memory function (in the present case, the second-order memory function). Such an expression can be obtained by making the so-called "mode-coupling approximations". The heart of these approximations is to replace in the expression for the (second order) memory function the operator $(1-\mathcal{P})\mathcal{L}$ by the normal Liouville operator $\mathcal{L}$ and to factorize a four-point correlation function into a product of two-point correlation functions which are just $F(q,t)$ [8,31].

At the end the resulting equations of motion have the following form:

$$\ddot{\Phi}(q,t) + \Omega_q^2\Phi(q,t) + \Omega_q^2 \int_0^t M_{MC}(q,t-t')\dot{\Phi}(q,t')dt' = 0 \ . \tag{21}$$

Here $\Phi(q,t)$ is defined as $F(q,t)/F(q,0) = F(q,t)/S(q)$, where $S(q)$ is the static structure factor. m is the mass of the particles and the frequency Ω_q is given by

$$\Omega_q^2 = q^2 k_B T/(mS(q)) \ . \tag{22}$$

The memory function $M_{MC}(q,t)$ is given by

$$M_{MC}(q,t) = \frac{1}{2(2\pi)^3}\int d\mathbf{k} V^{(2)}(q,k,|\mathbf{q}-\mathbf{k}|)\Phi(k,t)\Phi(|\mathbf{q}-\mathbf{k}|,t) \tag{23}$$

where the "vertex" $V^{(2)}$ is given by

$$V^{(2)}(q,k,|\mathbf{q}-\mathbf{k}|) = \frac{n}{q^2}S(q)S(k)S(|\mathbf{q}-\mathbf{k}|)\left(\frac{\mathbf{q}}{q}\cdot[\mathbf{k}c(k)+(\mathbf{q}-\mathbf{k})c(|\mathbf{q}-\mathbf{k}|)]\right)^2 . \tag{24}$$

Here $n = N/V$ is the particles density and $c(k)$ is the so-called "direct correlation function" that is related to the structure factor via $c(k) = n(1-1/S(q))$

[18, 19]. Note that in (23) the function M_{MC} is a bilinear function of the "modes" $\Phi(q,t)$, i.e. the modes are "coupled". This is the reason why this approach is called "mode-coupling theory".

Equations (21)–(24), commonly referred to as "mode-coupling equations, constitute a set of equations that describe the time and wave-vector dependence for $F(q,t)$. The only input required are *static* quantities, such as the particles density, the temperature, and the static structure factor. Therefore we see that these *static quantities* determine the *relaxation dynamics* of the system, a conclusion that is somewhat surprising.

An inspection of the structure of the mode-coupling equations shows that they have the form of a retarded damped oscillator, see (21). The damping is given by $M_{MC}(q,t)$, which increases with decreasing temperature since, see (23) and (24), $S(q)$ becomes (usually!) more peaked with decreasing T. Therefore one recognizes immediately that the relaxation dynamics will slow down with decreasing T. In addition the damping is also a bi-linear function of the correlator, see (23), and hence a slower decay of $\Phi(q,t)$ will lead to a damping that lasts longer. This *non-linear* feedback effect makes that with decreasing temperature the slowing down of the dynamics is not a smooth function of temperature, but shows in fact a singular behavior. This singularity occurs at a temperature T_c at which, due to the increase of the vertices $V^{(2)}$, the feedback mechanism is so strong that the solutions of the mode-coupling equations do no longer go to zero even at infinite time, which means that the system is no longer ergodic. The temperature T_c is usually referred to as the "critical temperature of mode-coupling theory". Note that it is an intrinsic *equilibrium* temperature of the system at which the relaxation dynamics is predicted to show a dramatic change upon a variation of T and therefore should be independent of cooling rate of the sample etc.

Due to their complexity, no analytical solutions of the mode-coupling equations is known. However, for temperatures close to T_c it is possible to simplify the mode-coupling equations since one has a strong separation of the time scales between the microscopic ones, which are on the order of Ω_q^{-1}, and the ones of the α-relaxation, which is predicted to diverge for $T \to T_c^+$. In the following section we will discuss some of the predictions that can be obtained from these simplified equations although we will not have the space to derive these predictions. Instead we refer the reader to the more specialized literature on this subject [8, 14, 16, 17].

Before we end this section we mention that the mode-coupling equations given by (21)–(24) do not only hold for $F(q,t)$ but that it is possible to derive also similar equations for all other observables $X(t)$ that couple to the density fluctuations $\delta\rho(q,t)$, see (20), i.e. for which the thermal average $\langle \delta X(t) \delta\rho^*(q,t) \rangle$ is non-zero. This is, e.g., the case for the frequency dependent specific heat, the mean squared displacement, etc. One particularly important example is the self intermediate scattering function $F_s(q,t)$ which is defined as

$$F_s(q,t) = \langle \delta\rho_s(\mathbf{q},t)\delta\rho_s^*(\mathbf{q},t)\rangle \quad \text{with} \quad \delta\rho_s(\mathbf{q},t) = \exp[i\mathbf{q}\cdot\mathbf{r}_j(t)] \; . \tag{25}$$

Also this function is of experimental relevance since it can be measured, e.g., in light-and neutron scattering.

The mode-coupling equations that one obtains for $F_s(q,t)$ are given by

$$\ddot{F}_s(q,t) + \frac{q^2 k_B T}{m} F_s(q,t) + \int_0^t M^s_{MC}(q,t-t')\dot{F}_s(q,t')dt' = 0 \; . \tag{26}$$

Here the memory function is given by

$$M^s_{MC}(q,t) = \frac{nk_BT}{(2\pi)^3 m} \int d\mathbf{q}' \left(\frac{\mathbf{k}'\cdot\mathbf{q}}{q}\right)^2 c(q')S(q')\Phi(q',t)F_s(|\mathbf{q}-\mathbf{q}'|,t) \; . \tag{27}$$

Thus we see that although the structure of (26) is similar to the one for $F(q,t)$, see (21), the expression for the memory function is very different in that the one for $F_s(q,t)$ involves also the correlators $F(q,t)$. Thus in order to determine the relaxation dynamics of $F_s(q,t)$ one needs first to obtain the one for $F(q,t)$. This result is quite plausible since in order for a particle to move in a dense system, one needs to know how it is moving with respect to its neighbors (collective effects!) and this information is encoded in $F(q,t)$.

3 Computer Simulations of Glass-Forming Liquids

In this section we will present some results on computer simulations of a simple glass-former. These results have been obtained from molecular dynamics simulations using the velocity form of the Verlet algorithm. The reason for the choice of such a simple algorithm are two-fold: Firstly one is interested in the relaxation *dynamics* of glass-forming liquids in order to understand how *real* systems relax and therefore one should use a "realistic" dynamics, i.e. the solution of Newton's equation of motion. Secondly, and more embarrassing, there exist presently no reliable algorithms that allow to equilibrate glass-forming liquids at very low temperatures. Although approaches such as parallel tempering etc. [32] help to some extent [33], it is not obvious whether they also give *correct* results at low T and hence they are not often used in the field of glass-forming liquids.

The system we investigate is a binary (80:20) Lennard-Jones mixture (BLJM) in which we denote the majority component as the "A-particles" and the minority component as "B-particles" [34–36]. All particles have the same mass m and interact via a Lennard-Jones potential of the form

$$V_{\alpha\beta}(r) = 4\epsilon_{\alpha\beta}[(\sigma_{\alpha\beta}/r)^{12} - (\sigma_{\alpha\beta}/r)^6] \quad \text{with} \quad \alpha,\beta \in \{\mathrm{A},\mathrm{B}\} \; . \tag{28}$$

The parameters of the potential are given by $\epsilon_{\mathrm{AA}} = 1.0$, $\sigma_{\mathrm{AA}} = 1.0$, $\epsilon_{\mathrm{AB}} = 1.5$, $\sigma_{\mathrm{AB}} = 0.8$, $\epsilon_{\mathrm{BB}} = 0.5$, and $\sigma_{\mathrm{BB}} = 0.88$. Hence we use σ_{AA} and ϵ_{AA} as the unit

of length and energy, respectively (setting the Boltzmann constant $k_B = 1$). Time will be measured in units of $(m\sigma_{AA}^2/48\epsilon_{AA})^{1/2}$. The number of particles used was 1000 and the cubic simulation box had the size $(9.4)^3$.

In the past the static and dynamical properties of this system has been studied extensively and it has been found that, due to the appropriate choice of the interaction parameters, it is not prone to crystallization [37]. This can be seen, e.g., from Fig. 4 where we show the partial static structure factor $S_{\rm AA}(q)$. This quantity is just the value of the coherent intermediate scattering function $F_{\rm AA}(q,t)$, see (19), at $t = 0$, i.e.

$$S_{\rm AA}(q) = \frac{1}{N_{\rm A}+N_{\rm B}} \sum_{j=1}^{N_{\rm A}} \sum_{l=1}^{N_{\rm A}} \langle \exp[i\mathbf{q}\cdot(\mathbf{r}_j(0)-\mathbf{r}_l(0))]\rangle \ . \tag{29}$$

(Note that we have here generalized the expression from (19) to the case of a binary mixture [18].) Figure 4 shows that within the temperature range investigated $S_{\rm AA}(q)$ shows only a very mild variation with T, i.e. there is no sign of the formation of a crystal (which would give rise to Bragg-peaks). A similar weak dependence has been found for other structural quantities as well for thermodynamic quantities (specific heat, pressure,...) and therefore we have strong evidence that this system is indeed a good glass-former and that it should be of interest to investigate its dynamical properties.

The simplest quantity that can be used to characterize the relaxation dynamics of a system is the mean-squared displacement (MSD) of a tagged particle, i.e.

$$\langle r^2(t)\rangle = \frac{1}{N_\alpha} \sum_{i=1}^{N_\alpha} \langle |\mathbf{r}_i(t)-\mathbf{r}_i(0)|^2\rangle \ . \tag{30}$$

In Fig. 8 we show the time dependence of $\langle r^2(t)\rangle$ for the A particles for different temperatures. We start the discussion of the data with the curves for high T. For short times we see that $\langle r^2(t)\rangle \propto t^2$, since we are in the ballistic regime mentioned in Sect. 2. Once $\langle r^2(t)\rangle$ is of the order 0.04, i.e. the distance is around 0.2, the time dependence changes over to a power-law with exponent 1.0, i.e. the system is diffusive. This change in the t-dependence is due to the collisions of the tagged particle with the neighbors that surround it at $t = 0$. Despite the presence of these neighbors the particle is still able to move away quickly from its initial position, i.e. $\langle r^2(t)\rangle$ increases quickly with time, since the kinetic energy of the particle is high enough to escape the local crowding.

Also for low temperatures we see at early times a ballistic behavior. However, in contrast to the curves for high T, the MSD does at the end of this regime not cross over to the diffusive regime, but instead shows a plateau at intermediate times. This plateau is due to the cage effect mentioned in Sect. 2, i.e. the temporary trapping of the particle by its neighbors. Only for sufficiently long times the particle is able to leave this cage and once it has done so the MSD quickly crosses over to a t-dependence that corresponds to

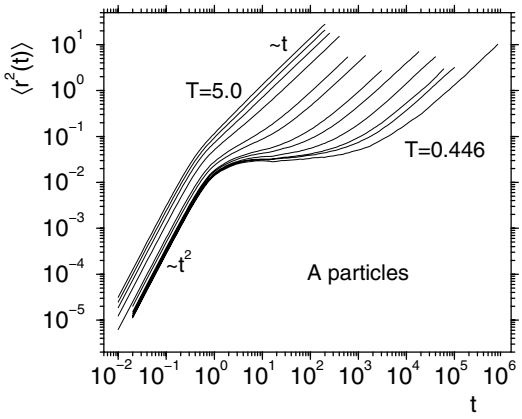


Fig. 8. Time dependence of the mean-squared displacement of the A particles for a BLJM. The curves correspond to different temperatures which are given by $T = 5.0$, 4.0, 3.0, 2.0, 1.0, 0.8, 0.6, 0.55, 0.50, 0.475, 0.466, and 0.446 (from left to right)

a diffusive motion. Note that the height of the plateau is around 0.03, from which it follows that the size of the cage is around 0.17, i.e. it is relatively small compared to the typical nearest neighbor distance which is around 1.0, a value that can be read off from the location of the first peak in the radial distribution function [35].

From this figure we recognize immediately that the slowing down of the dynamics is due to the existence of the plateau at intermediate times, whereas the behavior of the system at very short and very long times is independent of T. Since this plateau is due to the cage-effect, the goal of a theory of the glass transition should be to come up with a *self-consistent* description for this cage-effect, and MCT is indeed attempting to do this.

From the time dependence of the MSD at large t one can obtain, using the Einstein relation $D = \lim_{t\to\infty}\langle r^2(t)\rangle/6t$, the diffusion constant D. Before we discuss the temperature dependence of this quantity, we will, however, study the time and temperature dependence of a further important correlation function, the incoherent intermediate scattering function $F_s(q,t)$, which we have already introduced in (25).

In Fig. 9 we show the time dependence of $F_s(q,t)$ for the A particles. The wave-vector is $q = 7.25$, the location of the main peak in $S_{\rm AA}(q)$, see Fig. 4. From this figure we see that at high temperatures the curves decay, after the ballistic regime, very quickly to zero. This decay is basically exponential and characteristic for the relaxation dynamics of a simple liquid at high temperatures. If T is decreased, one finds that $F_s(q,t)$ starts to show a weak shoulder at around $t = 10$ which becomes more pronounced if T is lowered further. At the lowest temperature this shoulder has developed into a plateau, the reason of which is the cage-effect. At these low temperatures the time correlation function does no longer decay to zero in an exponential manner, but is

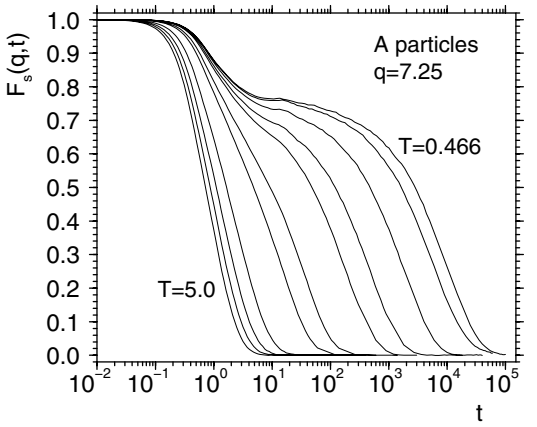


Fig. 9. Time dependence of the incoherent intermediate scattering function for different temperatures (T = 5.0, 4.0, 3.0, 2.0, 1.0, 0.8, 0.6, 0.55, 0.50, 0.475, and 0.466). The wave-vector corresponds to the location of the main peak in the static structure factor

stretched, i.e. it can be fitted well by the KWW function given by (2) with a $\beta < 1$.

MCT predicts that in the α-relaxation regime the shape of the time correlation functions should be independent of T, *if* one is close to T_c. Thus it should be possible to write the time dependence of a correlator $\Phi(t)$ in the following way:

$$\Phi(t,T) = \widehat{\Phi}(t/\tau(T)) \;, \tag{31}$$

where $\widehat{\Phi}$ is a master function and $\tau(T)$ is the relaxation time. This property is called "time-temperature superposition principle" (TTSP) since it allows to predict the t-dependence at one temperature if one knows it at a different T. One possibility to check whether the TTSP holds is to plot the time correlation functions as a function of $t/\tau(T)$ and to see whether they fall onto a master curve.

In Fig. 10 we show the data from Fig. 9 as a function of $t/\tau(T)$, where we have defined τ via the relaxation $F_s(q,\tau) = e^{-1}$. (Although this definition is somewhat arbitrary (instead of e^{-1} one could have chosen another value, or have taken the area under the correlator), the precise nature of the definition is irrelevant as long as the TTSP holds.) From this figure we recognize that this scaling of the time axis does indeed lead to a nice collapse of the curves for low T onto (essentially) *two* master curves. The first one corresponds to the data at high temperatures and is approximated very well by an exponential. The second one is for the correlators at intermediate and low temperatures and its shape can be described well by a KWW-law with an exponent β around 0.83, see dashed line in Fig. 10. Qualitatively similar results are obtained if one considers other wave-vectors, the incoherent function for the B particles, or the coherent functions $F_{\alpha\beta}(q,t)$. (The latter have usually a statistical accuracy

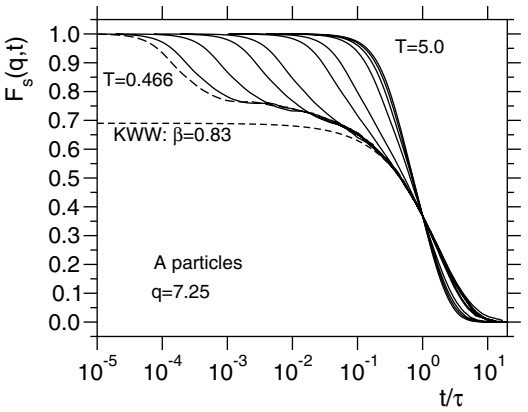


Fig. 10. Incoherent intermediate scattering function as a function of rescaled time $t/\tau(T)$ for the same correlators shown in Fig. 9. Also included is a fit to the curve at the lowest T at long times with the KWW function from (2)

that is somewhat inferior to the one of $F_s(q,t)$ and therefore one often focuses on $F_s(q,t)$.) Thus we can conclude that the prediction of MCT regarding the TTSP does indeed hold, at least for the present system.

As already discussed in the previous section, MCT predicts that each glass-forming system has a temperature T_c at which the relaxation times diverge (and the diffusion constant goes to zero). The functional form of this divergence is given by a power-law,

$$\tau \propto D^{-1} \propto (T - T_c)^{-\gamma} \; , \tag{32}$$

where $\gamma > 1.0$ is an exponent that depends on the system and which can be calculated from the solution of the mode-coupling equations, i.e. from the static structure factor. In Fig. 11 we show in a double logarithmic plot the data for the diffusion constants D_α as well as for τ_α^{-1}, as determined from $F_s(q,t)$ for q at the peak of $S_{\alpha\alpha}(q)$, as a function of $T-T_c$ [38]. Here T_c was chosen such that at intermediate and low temperatures the data is rectified, i.e. follows the T-dependence given by (32). The first conclusion that can be drawn from this figure is that there exists indeed a temperature T_c for which *all* the data can be rectified. Thus this is evidence that the notion of a critical temperature does, at least for this system, indeed make sense. The corresponding power-laws are included in the figure as bold solid lines. Secondly one sees that the exponents of these power-laws, stated in the figure, do depend on the data considered in that they are below 2.0 for the diffusion constants and around 2.4 for the relaxation times. Since MCT predicts that these exponents should all be equal, see (32), we must conclude that this prediction of the theory does not hold for the present system. It is presently believed that the reason for this discrepancy is related to the presence of the so-called "dynamical heterogeneities" [39–41]. By this one refers to the observation that the relaxation dynamics of the

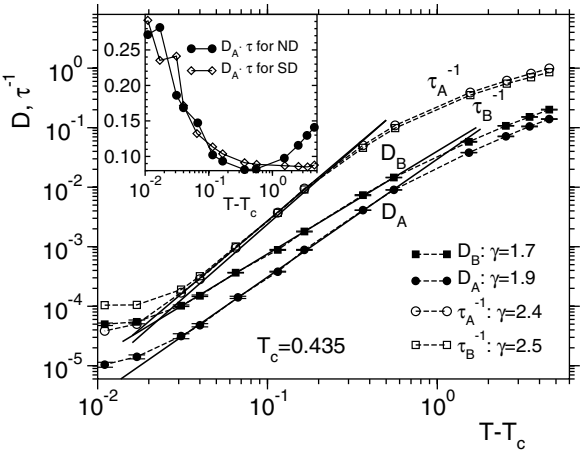


Fig. 11. The diffusion constant D_α and the inverse of the α-relaxation time τ_α as a function of $T - T_c$ for the case of the BLJM. The straight lines are fits to the data with power-laws, see (32). The values of the exponents γ are given in the figure as well. Inset: T-dependence of the product $D_\mathrm{A} \cdot \tau_\mathrm{A}$. The full symbols correspond to a Newtonian dynamics whereas the open ones correspond to a stochastic dynamics

system is not uniform throughout the sample, but instead shows a spacial dependency that is related to the statistical fluctuations in the structure of the liquid. Thus there are regions that relax faster and others that relax slower. (But of course on a time scale that is much longer than the α-relaxation time the structure and hence the dynamics of the sample is homogeneous!) Due to the presence of these heterogeneities the mean squared displacement, i.e. the diffusion constant, and the incoherent intermediate scattering function, i.e. τ, are averaged very differently and hence it is no surprise that they do not show exactly the same temperature dependence. This can be seen, e.g., from the inset of Fig. 11 where we plot for the case of the A particles the product $D \cdot \tau$ as a function of T. This graph clearly shows that with decreasing T the product increases rapidly, since, due to the dynamical heterogeneities, there are always a few particles that have a relatively large MD and which hence increase strongly the diffusion constant. Last not least we also mention that the value found for the exponent γ for the relaxation times, around 2.4, is in excellent agreement with the theoretical prediction for this system. In the context of Fig. 13 we will come back to this point.

The data presented so far has been obtained by using for the simulation the usual Newtonian dynamics. However, for the case of colloidal systems the motion of the colloidal particles is given by a Brownian dynamics and one might wonder to what extent the nature of the microscopic dynamics influences the relaxation dynamics at long times. Since there are presently no efficient numerical algorithms available that correspond to a Brownian dynamics, we chose to make a compromise by simulating a "stochastic dy-

namics" (SD), i.e. a dynamics in which the particles are subject to a random force, a viscous damping force but still have a *finite* mass, i.e. the inertia term cannot be neglected. Thus the resulting equations of motion are given by

$$m\ddot{\mathbf{r}}_i + \nabla_i \sum_l V_{il}(|\mathbf{r}_l - \mathbf{r}_i|) = -\zeta\dot{\mathbf{r}}_i + \eta_i(t). \tag{33}$$

Here V_{il} is the potential between particles i and l, $\eta_i(t)$ are Gaussian distributed random variables with zero mean, i.e., $\langle\eta_i(t)\rangle = 0$, and ζ is a damping constant. The fluctuation dissipation theorem relates ζ to the second moment of η_i, and thus we have $\langle\eta_i(t)\cdot\eta_l(t')\rangle = 6k_BT\zeta\delta(t-t')\delta_{il}$ [42]. The value used for ζ is 10, which is large enough to ensure that the relaxation dynamics at intermediate and long times is independent of ζ, apart from a trivial scaling of the time scale by a factor ζ.

Note that since the interactions between the particles in the Newtonian and the stochastic system are exactly the same, all equilibrium properties must be identical. Since the static structure factor for the system with the Newtonian dynamics is therefore the same as the one with the SD, the remarkable prediction of MCT is that the relaxation dynamics of these two systems should be identical, apart from an overall shift of the time constant, see the discussion on this in Sect. 2. A comparison of the two dynamics shows that this prediction of the theory holds surprisingly well in that, e.g., the T-dependence of the diffusion constants and of the α-relaxation times is the same [43,44]. This conclusion can also be reached from Fig. 12 where we show

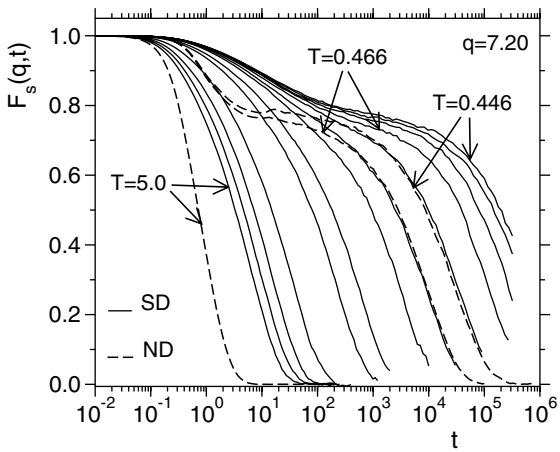


Fig. 12. Time dependence of the incoherent intermediate scattering function for the A particles in a binary Lennard-Jones mixture. The *solid lines* are for the case of the stochastic dynamics at the temperatures $T = 5.0$, 4.0, 3.0, 2.0, 1.0, 0.8, 0.6, 0.55, 0.5, 0.475, 0.466, 0.452, and 0.446 and the *bold dashed lines* are for the Newtonian dynamics at $T = 5.0$, 0.466, and 0.446. The *dashed-dotted lines* is a fit to the SD curve at $T = 0.446$ with the β-correlator from the MCT. From [38]

the t-dependence of the incoherent intermediate scattering function for the A particles as determined from the SD (thin lines). Also included is the data for $F_s(q,t)$ as obtained from the Newtonian dynamics (bold dashed lines). A comparison of these two data sets shows that at low temperatures the height of the plateau, i.e. the size of the cage, is the same. Furthermore also the shape of the curves in the α-relaxation regime, i.e. the stretching, is the same. Thus the main difference between the Newtonian dynamics and the SD is the way the correlation functions approach the plateau. Whereas in the SD this approach is very gentle since the dynamics of the particles is strongly damped, the Newtonian dynamics shows an approach which is much quicker and which is dominated by phonon-like vibrations.

Last not least it is also found that the product of $D \cdot \tau$ is independent of the microscopic dynamics, see Inset of Fig. 11, which is evidence that also the nature of the dynamical heterogeneities is independent of the microscopic dynamics. Therefore we can indeed conclude that the nature of the β-and α-relaxation is independent of the microscopic dynamics, a result that is highly surprising and in agreement with the prediction of MCT.

The tests that we have presented so far in order to check whether or not MCT is able to give a correct description of the relaxation dynamics of the BLJM have the drawback that they concern only *qualitative* aspects of the theory (existence of power-law, TTSP,...). Therefore one uses quite a few fit parameters (T_c, γ,...) and at the end one might wonder whether MCT is not just a theory that gives a good qualitative description of the dynamics but hides its problems in a large number of fit parameters. In the last part of the section we will show that this conclusion is wrong and that the theory is indeed able to make also correct *quantitative* predictions. In order to make such quantitative predictions one must solve the mode-coupling equations given by (21)–(24) numerically, using as input a structure factor that can be obtained from analytical theories or from computer simulations. The first results on such an approach has been obtained by Barrat et al. and by Götze and Sjögren who used analytical structure factors to solve the MCT equations and compared the so obtained solutions with results from computer simulations and experiments of colloidal particles [45, 46]. The comparison showed that the theory is indeed able to reproduce also quantitatively the relaxation dynamics of the glass-forming systems, thus giving evidence that the MCT is not only qualitatively correct but also quantitatively.

This type of comparison has also been made for the case of the BLJM by Nauroth et al. [47–49]. These authors obtained the three partial structure factors from computer simulations and used these functions as input for solving the MCT equations. One of the result of this calculation is presented in Fig. 13 where we show the wave-vector dependence of the non-ergodicity parameter, i.e. the height of the plateau in the coherent intermediate scattering function (solid lines). Also included in the graphs are the non-ergodicity parameters as determined from the simulations. We see that the agreement between theory and simulation is excellent in that the theoretical curve reproduces the data

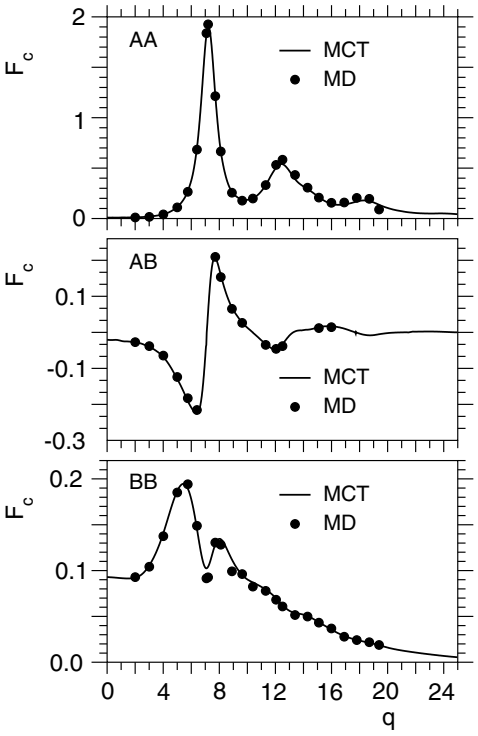


Fig. 13. Wave-vector dependence of the nonergodicity parameters of $F(q,t)$ for the case of a binary Lennard-Jones system. The *dots* are the results from the molecular dynamics simulations from [38]. The *lines* are the prediction of MCT. Adapted from [48]

from the simulation within the accuracy of the latter. We emphasize that for the calculation of the theoretical curve *no* fit parameter of any kind has been used and therefore one can evidently conclude that MCT is able to make also *quantitative* predictions with a very high accuracy.

The same calculation also gave the theoretical prediction for the value of the exponent γ in (32) and it was found that $\gamma = 2.34$. This value is in very good agreement with the values that were obtained in the simulations by considering the relaxation time τ, $\gamma \approx 2.4$, see Fig. 11. (Recall that when we discussed that figure we already mentioned that the value of γ as determined from the diffusion constant is likely to be underestimated, since the dynamical heterogeneities will make that the diffusion constant is higher than expected from the relation $D \propto \tau^{-1}$.) Hence this is further evidence that MCT is indeed able to make reliable *quantitative* predictions on the relaxation dynamics of glass-forming systems.

Many more tests on the validity of the theory have been done and most of them show that the theory is indeed able to give an accurate description

of the relaxation dynamics of glassy systems. The systems investigated so far include colloids, simple liquids, molecular liquids, polymers, as well as network forming liquids such a silica [8, 15–17]. Therefore it is clear that this theory is indeed highly valuable in that it allows to understand many features of glass-forming liquids.

Before we end this section it is appropriate to make some further remarks on the validity of the theory. MCT predicts the existence of a temperature T_c close to which the relaxation dynamics shows a very strong dependence on the external parameters (temperature, pressure, ...) in that the relaxation times are predicted to diverge. A look at the experimental data shows however, see Fig. 2, that in the temperature range accessible to experiments the relaxation times show *no* divergence. So what is going on? The MCT equations (21)–(24) are the ones that are obtained if one uses in the Zwanzig-Mori formalism the density fluctuations as slow variables. However, it is obvious that these are not the only slow variables in the system but other ones will be slow as well (e.g. the current density). If one includes also these other slow variables in the Zwanzig-Mori formalism, one obtains equations that have a very similar structure than the ones discussed in Sect. 2 and which are usually referred to as "extended mode-coupling equations" or "mode-coupling equations with hopping terms", whereas the equation given by (21)–(24) are usually called "ideal mode-coupling equations". Due to these additional variables it is, however, no longer possible to give explicit expressions for the vertices and therefore no quantitative calculation can be made. The structure of the equations allows however to see that the solution of the extended mode-coupling equations are very similar to the solutions of the ideal equations. In particular these solutions will also show a rapid increase of the relaxation time with decreasing T. However, due to the additional terms the increase of τ for temperatures very close to T_c is weaker than the one predicted by the solution of the ideal equations in that it is closer to an Arrhenius law than to the power-law given by (32) [8, 50, 51]. Thus this means that the predicted singularity in the dynamics at a finite T is an artifact of the ideal MCT which is no longer present in the extended version of the theory. Nevertheless, most, but not all, results of the ideal theory still hold even for the extended version of the theory and therefore one focuses in most cases on the ideal theory since it is significantly simpler to handle. Although the singularity predicted by the ideal MCT becomes smeared out by the "hopping processes", i.e. the additional terms in the extended theory, the presence of the singularity still can be seen. In the Arrhenius plot of the relaxation time it can be expected that $\tau(T)$, or $\eta(T)$ does not show a divergence at a finite temperature, i.e. at T_c, but at least a rapid increase. This is indeed the case as can be inferred from Fig. 2 where we see that most glass-formers show at around $\eta \approx 10^2$ Pa s a strong bend. This bend can be interpreted as the ghost of the singularity predicted by MCT and more detailed investigations have shown that this is indeed the case. Thus we can conclude that MCT does indeed predict correctly the crossover in the dynamics from the moderately glassy dynamics at intermediate temperatures

to the very slow dynamics seen at low temperatures. To what extent the theory is also reliable (qualitatively or quantitatively) far below this crossover temperature is presently not really known and remains a subject of research for the future.

4 Conclusions

Many of the properties of glass-forming systems are still not understood very well and many of them do not have a satisfactory theoretical explanation. In this short review we have presented only some of the salient features that are found in these systems. Due to the lack of space we have not been able to discuss all of the other properties, although they are important as well and are a challenge to be understood (two-level systems, Boson peak, nature of the dynamical heterogeneities, aging behavior at low temperatures, confined systems, ...). In addition, and as already mentioned in the Introduction, glass-forming liquids are by no means the only statistical mechanics systems that show a glassy dynamics. There are spin glasses, foams, gels, granular materials etc. and all of them show features that are quite similar to the ones of glass-forming liquids. To what extent this similarity is just superficial or whether there is a more fundamental theoretical reason for it remains presently not understood very well. It seems therefore obvious that the field of disordered materials is still an area that remains to a large extent unexplored and poses many experimental and theoretical challenges.

References

1. K. Binder and A. P. Young (1986) Spin-glasses – Experimental facts, theoretical concepts, and open questions. *Rev. Mod. Phys.* **58**, p. 801
2. K. H. Fischer and J. A. Hertz (1991) *Spin Glasses.* Cambridge University Press, Cambridge
3. J. Zarzycki (Ed.) (1991) *Materials Science and Technology.* vol. 9 VCH Publ., Weinheim
4. A. Feltz (1993) *Amorphous Inorganic Materials and Glasses.* VCH, Weinheim
5. P. G. Debenedetti (1997) *Metastable Liquids.* Princeton University Press, Princeton
6. A. P. Young (Ed.) (1998) *Spin Glasses and Random Fields.* World Scientific, Singapore
7. E.-W. Donth (2001) *The Glass Transition Relaxation Dynamics in Liquids and Disordered Materials.* Springer, Berlin
8. K. Binder and W. Kob (2005) *Glassy Materials and Disordered Solids: An Introduction to their Statistical Mechanics* World Scientific, Singapore
9. C. A. Angell, P. H. Poole, and J. Shao (1994) Glass-forming liquids, anomalous liquids, and polyamorphism in liquids and biopolymers. *Nuovo Cimento D.* **16**, p. 993

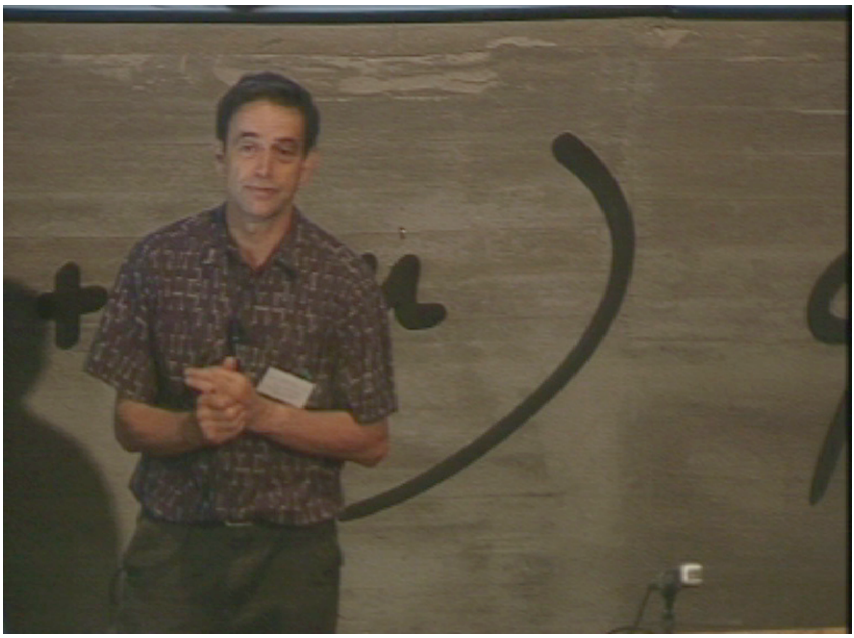
10. D. R. Uhlmann (1972) A kinetic treatment of glass formation. *J. Non-Cryst. Solids.* **7**, p. 337
11. C. A. Angell (1985) Fast Ion Conductors in Viscous Liquids and Glasses. In *Relaxation in Complex Systems*, K. L. Ngai and G. B. Wright (eds.). p. 1, US Dept. Commerce, Springfield
12. K.-U. Hess, D. B. Dingwell, and E. Rössler (1996) Parametrization of viscosity-temperature relations of aluminosilicate melts. *Chem. Geol.* **128**, p. 155
13. E. Rössler, K.-U. Hess, and V. N. Novikov (1998) Universal representation of viscosity in glass forming liquids. *J. Non-Cryst. Solids.* **223**, p. 207
14. W. Götze (1989) In *Liquids, Freezing and the Glass Transition.* Eds.: J. P. Hansen, D. Levesque, and J. Zinn-Justin, Les Houches, Session LI (1991). p. 287, North-Holland, Amsterdam
15. W. Götze and L. Sjögren (1992) Relaxation processes in supercooled liquids. *Rep. Prog. Phys.* **55**, p. 241
16. W. Götze (1999) Recent tests of the mode-coupling theory for glassy dynamics. *J. Phys.: Condens. Matter.* **10**, p. A1
17. S. P. Das (2004) Mode-coupling theory and the glass transition in supercooled liquids. *Rev. Mod. Phys.* **76**, p. 785
18. J.-P. Hansen and I. R. McDonald (1986) *Theory of Simple Liquids.* Academic, London
19. J.-L. Barrat and J.-P. Hansen (2003) *Basic Concepts for Simple and Complex Liquids.* Cambridge University Press, Cambridge
20. E. Courtens, M. Foret, B. Hehlen, and R. Vacher (2001) The vibrational modes of glasses. *Solid State Commun..* **117**, p. 187
21. W. Kauzmann (1948) The nature of the glassy state and the behavior of liquids at low temperatures. *Chem. Rev.* **43**, p. 219
22. R. Richert and C. A. Angell (1998) Dynamics of glass-forming liquids. V. On the link between molecular dynamics and configurational entropy. *J. Chem. Phys.* **108**, p. 9016
23. G. Adam and J. H. Gibbs (1965) On temperature dependence of cooperative relaxation properties in glass-forming liquids. *J. Chem. Phys.* **43**, p. 139
24. J. H. Magill (1967) Physical properties of aromatic hydrocarbons .3. A test of Adam-Gibbs relaxation model for glass formers based on heat-capacity data of 1,3,5-tri-alpha-naphthylbenzene. *J. Chem. Phys.* **47**, p. 2802
25. C. Alba, L. E. Busse, D. J. List, and C. A. Angell (1990) Thermodynamic aspects of the vitrification of toluene, and xylene isomers, and the fragility of liquid hydrocarbons. *J. Chem. Phys.* **92**, p. 617
26. C. A. Angell (1997) Entropy and fragility in supercooling liquids. *J. Res. NIST.* **102**, p. 171
27. H. Tanaka (2003) Relation between thermodynamics and kinetics of glass-forming liquids. *Phys. Rev. Lett.* **90**, 055701
28. R. Zwanzig (1960) Ensemble method in the theory of irreversibility. *J. Chem. Phys.* **33**, p. 1338
29. H. Mori (1965) Transport collective motion and brownian motion. *Prog. Theor. Phys.* **33**, p. 423
30. U. Balucani and M. Zoppi (1994) *Dynamics of the Liquid State.* Oxford University Press, Oxford
31. U. Balucani, M. H. Lee, and V. Tognetti (2003) Dynamical correlations. *Phys. Rep.* **373**, p. 409

32. K. Hukushima and K. Nemoto (1996) Exchange Monte Carlo method and application to spin glass simulations. *J. Phys. Soc.* Japan, **65**, p. 1604
33. R. Yamamoto and W. Kob (2000) Replica-exchange molecular dynamics simulation for supercooled liquids. *Phys. Rev. E.* **61**, p. 5473.
34. W. Kob and H. C. Andersen (1994) Scaling Behavior in the ß-Relaxation Regime of a Supercooled Lennard-Jones Mixture. *Phys. Rev. Lett.* **73**, p. 1376
35. W. Kob and H. C. Andersen (1995) Testing mode-coupling theory for a supercooled binary Lennard-Jones mixture: The van Hove correlation function. *Phys. Rev. E.* **51**, p. 4626
36. W. Kob and H. C. Andersen (1995) Testing mode-coupling theory for a supercooled binary Lennard-Jones mixture II: Intermediate scattering function and Dynamic Susceptibility. *Phys. Rev. E.* **52**, p. 4134
37. W. Kob (1999) Computer Simulations of Supercooled Liquids and Glasses. *J. Phys.: Condens. Matter.* **11**, p. R85
38. T. Gleim (1998) *Ph.D. Thesis.* (Johannes Gutenberg Universität Mainz)
39. H. Sillescu (1999) Heterogeneity at the glass transition: a review. *J. Non-Cryst. Solids.* **243**, p. 81
40. M. D. Ediger (2000) Spatially heterogeneous dynamics in supercooled liquids. *Ann. Rev. Phys. Chem.* **51**, p. 99
41. R. Richert (2002) Heterogeneous dynamics in liquids: fluctuations in space and time. *J. Phys.: Condens. Matter.* **14**, p. R703
42. R. K. Pathria (1986) *Statistical Mechanics.* Pergamon Press, Oxford
43. T. Gleim, W. Kob, and K. Binder (1998) How does the relaxation of a supercooled liquid depend on its microscopic dynamics? *Phys. Rev. Lett.* **81**, p. 4404
44. G. Szamel and E. Flenner (2004) Independence of the relaxation of a supercooled fluid from its microscopic dynamics: Need for yet another extension of the mode-coupling theory. *Europhys. Lett.* **67**, p. 779
45. J.-L. Barrat, W. Götze, and A. Latz (1989) The liquid glass-transition of the hard-sphere system. *J. Phys.: Condens. Matter.* **1**, p. 7163
46. W. Götze and L. Sjögren (1991) Beta-relaxation at the glass-transition of hard-spherical colloids. *Phys. Rev. A.* **43**, p. 5442
47. M. Nauroth and W. Kob (1997) A quantitative test of the mode-coupling theory of the ideal glass transition for a binary Lennard-Jones system. *Phys. Rev. E.* **55**, p. 657
48. F. Sciortino and W. Kob (2001) The Debye-Waller factor of liquid silica: Theory and simulation. *Phys. Rev. Lett.* **86**, p. 648
49. W. Kob, M. Nauroth, and F. Sciortino (2002) Quantitative tests of mode-coupling theory for fragile and strong glass-formers. *J. Non-Cryst. Solids.* **307–310**, p. 181
50. W. Götze and L. Sjögren (1987) The glass-transition singularity. *Z. Phys. B.* **65**, p. 415
51. L. Sjögren (1990) Temperature-dependence of viscosity near the glass-transition. *Z. Phys. B.* **79**, p. 5

Numerical Simulations of Spin Glasses: Methods and Some Recent Results

A.P. Young[1]

Physics Department, University of California Santa Cruz Santa Cruz, CA 95064
peter@bartok.ucsc.edu



A. Peter Young

A.P. Young: *Numerical Simulations of Spin Glasses: Methods and Some Recent Results*, Lect. Notes Phys. **704**, 31–44 (2006)
DOI 10.1007/3-540-35284-8_2
© Springer-Verlag Berlin Heidelberg 2006

1 Introduction 33

2 Model and Theory 36

3 Numerical Aspects 37

3.1 Speeding Up the Dynamics 37
3.2 Test for Equilibration 38
3.3 Finite Size Scaling 39
3.4 Evidence for a Finite Transition Temperature 40

4 Absence of a Phase Transition in a Magnetic Field 41

5 Conclusions 43

References 44

1 Introduction

This talk will provide an introduction to spin glasses [1,2], why they are interesting, and what are the computational challenges in studying them numerically. At the end I will describe one topic that I have recently been working on; whether or not there is a phase transition in a spin glass in a magnetic field, the so-called Almeida-Thouless (AT) line.

It is generally agreed that for a system to be a spin glass it needs two ingredients: randomness and "frustration". Frustration refers to the competition between interactions which means that no configuration of the system will simultaneously minimize each term in the Hamiltonian. A toy example illustrating frustration is shown in Fig. 1. Here the arrows refer to Ising spins which can point up or down, $S_i = \pm 1$, and the interaction between them is ferromagnetic (preferring parallel spin alignment) if it is indicated by a "+" and antiferromagnetic (preferring antiparallel spin alignment) if indicated by a "−". A spin glass therefore has a random mixture of ferro- and antiferromagnetic tendencies. In the figure, no configuration of the spins will minimize the energy of each bond because there is an odd number of negative bonds.

It is clear from this example that if one considers a large lattice, determining the exact ground state of a spin glass is highly non-trivial. In fact, it is an example of what computer scientists call a combinatorical optimization problem. There has been overlap of ideas between spin glasses and the theory of combinatorical optimization with computer scientists providing efficient code for determining spin glass ground states [3] (in certain cases) and spin glass physics providing the inspiration for efficient algorithms [4] for some optimization problems.

Just as the determination of the true ground state is very hard, it is clear that there are many states fairly close by in energy to it and which, it turns

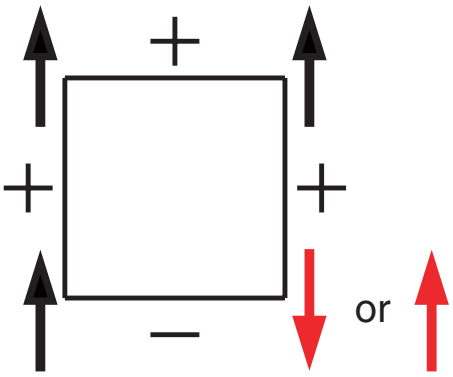


Fig. 1. A toy model which shows frustration. If the interaction on the bond is a "+", the spins want to be parallel and if it is a "−" they want to be antiparallel. Clearly all these conditions can not be met so there is competition or "frustration"

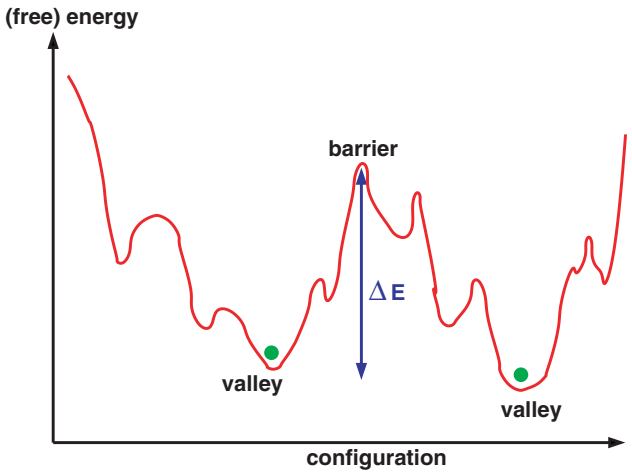


Fig. 2. A cartoon of the "energy landscape" in a spin glass. As the spin configuration is changed, represented by moving along the horizontal axis, the system goes through local minima (which have very similar energy but differ in the orientation of a large number of spins) separated by barriers. At low temperature the system easily gets trapped in one of the minima leading to slow dynamics

out, can differ in the orientation of a large number of spins. As a result an important concept in spin glasses is that of the "energy landscape", which is a projection on to a single axis of the variation of the energy (strictly speaking at finite temperature a suitably defined free energy). This has a complicated structure with minima (valleys) separated by barriers, see Fig. 2.

Experimentally, there are different types of materials with the ingredients of randomness and frustration needed to form a spin glass. For example:

- Metals:
  Diluted magnetic atoms, e.g. Mn, in a non-magnetic metal such as Cu, interact with the RKKY interaction whose sign depends on the distance between the atoms.
- Insulators:
  An example is $Fe_{0.5}Mn_{0.5}TiO_3$, which comprises hexagonal layers. The spins align perpendicular to layers (hence it is *Ising*-like). Since it is an insulator, the interactions are short-range. Within a layer, spins in pure $Fe_{0.5}TiO_3$ are ferromagnetically coupled whereas spins in pure $Mn_{0.5}TiO_3$ are antiferromagnetically coupled. Hence a random mixture of Fe and Mn gives a spin glass.

We should mention, at this point, that two types of averaging have to be done in random systems. First of all we need to perform the usual statistical mechanics average for a given set of random interactions. Precise values of measured quantities will depend on the particular set of random interactions, though many quantities, such as the free energy or energy, will be

"self-averaging", i.e. the sample to sample fluctuation will be small compared with the average value (generally down by of order $1/\sqrt{N}$ where N is the number of spins). However, even for self averaging quantites, the sample to sample fluctuations will not be negiligible for the small lattice sizes that can be simulated, and so we need to perform an average over the disorder, by repeating the simulation many times with different choices for the interactions. We denote the thermal average by $\langle \cdots \rangle$ and the disorder average by $[\cdots]_{\rm av}$.

Next we discuss the main experimental features of spin glasses. At low temperatures, the dynamics of spin glasses becomes very slow, so the system is not in equilibrium. This non-equilibrium behavior has been extensively studied in recent years. Of particular note has been the study of "aging" in spin glasses, pioneered by the Uppsala group [5]. One cools the system to low temperature and waits for a "waiting time" t_w. The system is then perturbed in some way, e.g. by applying a magnetic field, and the subsequent response is measured. It is found that the nature of the response depends on t_w, providing clear evidence that the system was not in equilibrium. However, in this talk, I will not be discussing this interesting non-equilibrium behavior.

If one measures the ac susceptibility $\chi(\omega)$ one finds a Curie, or Curie-Weiss behavior at high temperature followed by a fairly sharp cusp at a freezing temperature $T_f(\omega)$ which depends weakly on the frequency ω. There has been extensive discussion as to whether this represents "gradual freezing", i.e. the slowing down of the dynamics discussed above (of local entities), or whether the cusp indicates a sharp thermodynamic phase transition.

After considerable work, it became clear that a spin glass has a sharp phase transition at temperature $T = T_{SG}$, such that for $T < T_{SG}$ the spins freeze in some random-looking orientation. The spin glass order parameter can be taken to be

$$Q = [\langle S_i \rangle^2]_{\rm av} \ . \tag{1}$$

To evaluate the square of the thermal average in simulations without bias, two copies of the system, "1" and "2", with the same bonds are studied, and $Q = [\langle q \rangle]_{\rm av}$, where the microscopic order parameter (spin overlap) is given by

$$q = \frac{1}{N} \sum_i S_i^{(1)} S_i^{(2)} \ . \tag{2}$$

The transition temperature T_{SG} is equal to the zero frequency limit of the freezing temperature $T_f(\omega)$. As $T \to T_{SG}^+$, the spin glass correlation length, the distance over which spins are correlated, diverges. We should mention, though, that some pairs of spins will be ferromagnetically correlated, whereas others, the same distance apart, will be antiferromagnetically correlated because they have different environments. A related quantity which diverges, therefore, is the *spin glass susceptibility*

$$\chi_{\rm SG} = \frac{1}{N} \sum_{\langle i,j \rangle} [\langle S_i S_j \rangle^2]_{\rm av} \ , \tag{3}$$

(notice the square) which is accessible in simulations. It is also essentially the same as the *non-linear susceptibility*, χ_{nl}, which can be measured experimentally and is defined by the coefficient of h^3 in the expansion of the magnetization m,

$$m = \chi h - \chi_{nl} h^3 + \cdots \ , \tag{4}$$

where h is the magnetic field. We expect that χ_{nl} and $\chi_{\rm SG}$ diverge at T_{SG} like

$$\chi_{nl} \sim (T - T_{SG})^{-\gamma} \ , \tag{5}$$

where γ is a critical exponent. This divergent behavior has been seen in many experiments. A nice example is the data of Omari et al. [6] on 1% Mn in Cu which shows a clear divergence.

2 Model and Theory

Most theoretical work has used the simplest model with the properties of randomness and frustration, which is due to Edwards and Anderson [7]:

$$\mathcal{H} = -\sum_{\langle i,j \rangle} J_{ij} S_i S_j \ , \tag{6}$$

in which the spins S_i lie on the sites of a regular lattice with $N = L^d$ sites with periodic boundary conditions. The interactions J_{ij}, which we assume here to be between nearest neighbors only, are independent random variables with mean and standard deviation given by

$$[J_{ij}]_{\rm av} = 0; \quad [J_{ij}^2]_{\rm av}^{1/2} = J \, (= 1) \ . \tag{7}$$

A zero mean is chosen to avoid any bias towards ferromagnetism or antiferromagnetism, and it is convenient, in the simulations, to take a Gaussian distribution for the J_{ij}, though sometimes a bimodel distribution, $J_{ij} = \pm 1$ with equal probability, is also used. Here we take Ising spins, $S_i = \pm 1$, though classical m-component vector spins are also of interest, see e.g. [8].

The infinite range version of the Edwards Anderson model, proposed by Sherrington and Kirkpatrick [9] (SK), has also been extensively studied. SK argued that the exact solution of this model could be considered the mean field theory for spin glasses, and that the actual behavior of a short range system may not be very different. These assumptions are true for ferromagnets. In fact, the exact solution of the SK model, found by Parisi [10, 11] is very complicated. To average over the disorder analytically, the "replica trick" is used, and the Parisi solution involves "replica symmetry breaking" (RSB) which seems to physically correspond to the phase space being divided up into different regions with infinite barriers in between. Hence, in a finite time, the system will stay confined within a single region ("ergodic component").

One of the surprising features of the SK model is that there is a transition in a magnetic field the Almeida Thouless (AT) [12] line, separating a complicated region with RSB (and hence a spectrum of relaxation times extending to infinity) below the line and a simpler region without RSB (and only finite relaxation times) above the line. Note that there is no symmetry change here since the order parameter q is non-zero on both sides of the line. The AT line therefore represents a pure ergodic-non ergodic transition without any symmetry change.

A lot of discussion has arisen as to the nature of the spin glass state below T_{SG} in short range spin glasses. Two main scenarios have been proposed.

- According to the RSB scenario, the behavior of short range spin glasses is very similar to that of the SK model. In particular there is an AT line in a magnetic field.
- In the "droplet picture" [13, 14], attention is focused on the geometrical aspects of the low energy excitations, which do not exist in an infinite range model. A number of properties are expected to be different in the droplet picture compared with the RSB picture; in particular a magnetic field rounds out the spin glass transition so there is no AT line.

A third, intermediate scenario, called the "TNT" (Trivial-Non Trivial) [15,16], has also been proposed. It is difficult to clearly determine numerically which scenario is correct because they refer to the equilibrium state, which is never achieved for large systems. Hence the system sizes are quite modest, and there is an uncertainty whether they are big enough for the asymptotic (i.e. $L \to \infty$) limit to have been reached.

3 Numerical Aspects

3.1 Speeding Up the Dynamics

We have mentioned that spin glass dynamics is slow at low temperatures on account of the complicated energy landscape. To speed things up most spin glass simulations now use the "parallel tempering" [17, 18] method in which one simulates n copies of the system, with the same bonds, at different temperatures, see Fig. 3. In addition to the usual single spin flip Monte Carlo moves for each copy, one also performs global moves in which configurations at neighboring temperatures are swapped, with a probability which satisfies the detailed balance condition for the ensemble of copies. The detailed balance



Fig. 3. Temperature swaps in the parallel tempering approach

condition ensures that the ensemble eventually comes to thermal equilibrium. However, if one looks at a single set of spins its temperature is not constant but performs a random walk between the minimum and maximum temperatures, $T_{min} \equiv T_1$ and $T_{max} \equiv T_n$ respectively. T_{max} is chosen to be sufficiently high that the system equilibrates fast, and so, if we follow a given set of spins from T_{min} up to T_{max} and back down again to T_{min}, there is no reason for it to be in the same valley when it returns to T_{min} as was at the beginning, because the configuration was completely randomized at T_{max}.

In practice, parallel tempering enables us to simulate intermediate size systems, e.g. $\sim 10^3$ spins, at low temperatures. Unfortunately, it still does not enable us to study very large system sizes.

3.2 Test for Equilibration

Even with parallel tempering, Monte Carlo dynamics is quite slow at low temperatures, and it is desirable to have a sound criterion for deciding whether the system has equilibrated. The usual approach is to increase the length of the simulation until the results don't seem to change. However, this can potentially be unreliable if measured quantities change only very slowly with number of Monte Carlo sweeps; e.g. results from runs of t and $2t$ sweeps could give the same results within error bars but could still be far from equilibrium.

For a spin glass with a Gaussian distribution of bonds an alternative approach has been developed [19]. The average energy U involves terms like $-[J_{ij}\langle S_i S_j\rangle]_{\rm av}$. The disorder average is with respect to the weight $\int dJ_{ij}$ $\exp(-J_{ij}^2/2)$ and so one can integrate the energy by parts with respect to the J_{ij}. This gives

$$U = U(q_l) \equiv -\frac{z}{2}\frac{1-q_l}{T} \ , \tag{8}$$

where q_l is the "link overlap"

$$q_l = \frac{1}{N_b}\sum_{\langle i,j\rangle}[\langle S_i S_j\rangle^2]_{\rm av} \ , \tag{9}$$

the sum is over nearest neighbor pairs, and $N_b = Nz/2$ is the number of such pairs. Here z is the number of neighbors per site, e.g. 6 for the simple cubic lattice.

Clearly U will approach its equilibrium value from above. To evaluate the average of the square in (9) two separate copies ("1" and "2") are simulated and $\langle S_i S_j\rangle^2$ is evaluated as $\langle S_i^{(1)} S_j^{(1)} S_i^{(2)} S_j^{(2)}\rangle$. Starting the spins in the two copies in random directions, it is plausible that q_l approaches its equilibrium value from below. Hence we expect the two sides in (8) to approach their common equilibrium value from opposite sides, and, once they agree, not to change further. This seems to be correct as shown in Fig. 4.

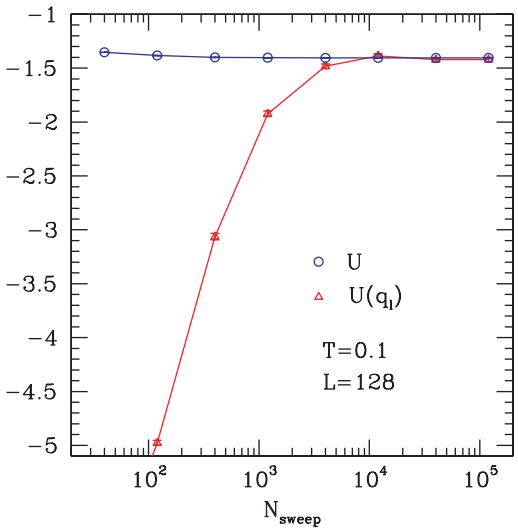


Fig. 4. Equilibration test in a spin glass with Gaussian interactions. This is actually for a one-dimensional model with long-range, power law interactions. (Adapted from [20])

3.3 Finite Size Scaling

In order to locate and analyze a critical point, we use the technique of finite size scaling to extrapolate results from a range of finite sizes to the thermodynamic limit. A particularly useful quantity turns out to be the correlation length of the finite size system ξ_L. One can extract this by Fourier transforming the spin glass correlation function

$$\chi_{\rm SG}(\mathbf{k}) = \frac{1}{N}\sum_{i,j}[\langle S_i S_j\rangle^2]_{\rm av} e^{i\mathbf{k}\cdot(\mathbf{R}_i-\mathbf{R}_j)} \ . \tag{10}$$

Above T_{SG} and a long wavelengths we expect this to have an Ornstein-Zernicke form

$$\chi_{\rm SG}(\mathbf{k}) \propto \left(\xi_L^{-2} + k^2\right) \tag{11}$$

and so we *define* ξ_L for all T to be

$$\xi_L^2 = \frac{1}{2(1-\cos(k_{\rm min}))}\left(\frac{\chi_{\rm SG}(0)}{\chi_{\rm SG}(\mathbf{k}_{\rm min})} - 1\right) \tag{12}$$

where $\mathbf{k}_{\rm min} = \frac{2\pi}{L}(1,0,0)$ is the smallest non-zero wavevector. A factor of $k_{\rm min}^2$ has been replaced by $2(1-\cos(k_{\rm min}))$ since this incorporates the periodicity of $\chi_{\rm SG}(\mathbf{k})$ in the repeated Brillouin zone.

The results for the correlation lengths will be analyzed according to finite-size scaling (FSS). The basic assumption of FSS is that the size dependence comes from the ratio $L/\xi_{\rm bulk}$ where

$$\xi_{\rm bulk} \sim (T - T_{SG})^{-\nu} \tag{13}$$

is the *bulk* correlation length. In particular, the *finite-size* correlation length is expected to vary as

$$\frac{\xi_L}{L} = X\left(L^{1/\nu}(T - T_{SG})\right) \ , \tag{14}$$

since ξ_L/L is dimensionless (and so has no power of L multiplying the scaling function X). In particular, it is reasonable to expect that $\xi_L \propto L$ at the critical point, since L is the only length scale left in the problem ($\xi_{\rm bulk}$ is infinite and the lattice spacing is assumed to be unimportant when ξ_L is large).

From (14), data for ξ_L/L for different sizes should intersect at T_{SG} and splay out below T_{SG}.

3.4 Evidence for a Finite Transition Temperature

Calculating the correlation length is an excellent technique to locate the critical point. As discussed in the previous subsection, the data should intersect at T_{SG} and splay out again at lower temperatures. This indeed occurs in the three dimensional Ising spin glass as shown in Fig. 5.

Use of the correlation length to locate the transition temperature in spin glasses was pioneered by Ballesteros et al. [21] for the $\pm J$ distribution. Prior to the work of Ballesteros et al., determination of T_{SG} generally used the "Binder ratio", a dimensionless ratio of the moments of the order parameter

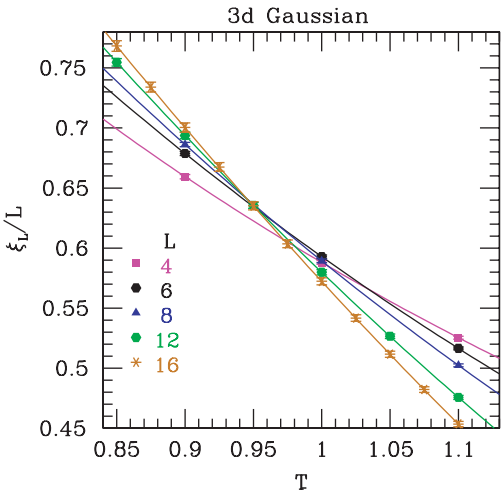


Fig. 5. Finite size scaling of the correlation length used to locate the critical point in the three-dimensional Ising spin glass with Gaussian interactions. (From Katzgraber and Young (unpublished))

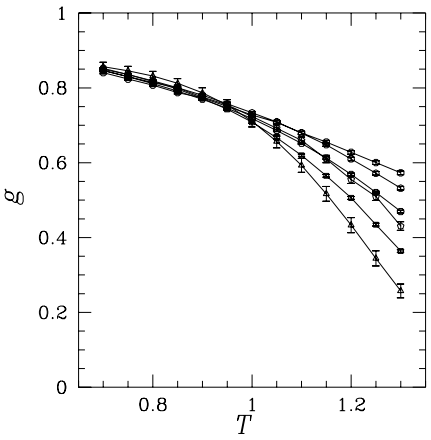


Fig. 6. Data for the Binder ratio length of the Ising spin glass with Gaussian interactions, from Marinari et al. [22]. The data merge but do not clearly splay out on the low-T side, unlike the results for the correlation length shown in Fig. 5

distribution which has a finite size scaling of the same form as in (14). However, this gives much less convincing demonstration of a transition, see Fig. 6 which shows data from Marinari et al. [22] for the Gaussian distribution.

4 Absence of a Phase Transition in a Magnetic Field

In the last part of this chapter, I will describe some recent work with Helmut Katzgraber [23] on the possible existence of a transition in a magnetic field (AT line) in a three-dimensional Ising spin glass. You will recall from the first part of the talk that an AT line is predicted by the RSB scenario for the spin glass state but not by the droplet scenario, see Fig. 7.

The Hamiltonian is now

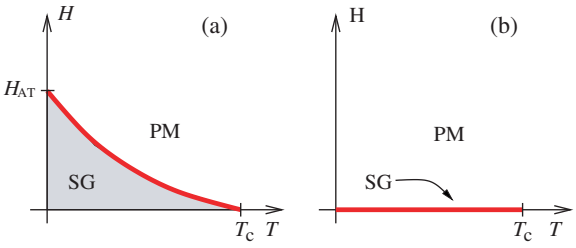


Fig. 7. Part (**a**) shows the phase diagram of a spin glass in a magnetic field expected in the RSB scenario. The line of phase transitions is called the AT line. Part (**b**) shows the analogous phase diagram in the droplet picture. There is no AT line

$$\mathcal{H} = -\sum_{\langle i,j \rangle} J_{ij} S_i S_j - \sum_i h_i S_i \ , \tag{15}$$

where h_i is the magnetic field on site i. Usually treatments of the AT line consider a uniform field, but we take a Gaussian distribution with zero mean and standard deviation H_r because, in this case, a generalization [23] of the equilibration test discussed above can be applied. For a symmetric distribution of bonds (used here), the *sign* of h_i can be "gauged away" so a uniform field is completely equivalent to a bimodal distribution of fields with $h_i = \pm H$. Our choice of a Gaussian distribution, which still has an AT line in mean-field theory, also puts disorder into the *magnitude* of the h_i.

In a magnetic field, where $\langle S_i \rangle$ is non-zero, the appropriate expression for $\chi_{\rm SG}(\mathbf{k})$ is

$$\chi_{\rm SG}(\mathbf{k}) = \frac{1}{N} \sum_{i,j} \left[\Big(\langle S_i S_j \rangle_T - \langle S_i \rangle_T \langle S_j \rangle_T \Big)^2 \right]_{\rm av} e^{i\mathbf{k}\cdot(\mathbf{R}_i - \mathbf{R}_j)} \ , \tag{16}$$

In order to evaluate the products of thermal averages without bias, 4 copies with the same bonds (at each temperature) are simulated. The expression for ξ_L is then the same, (12), as before.

Data for $H_r = 0.3$ are shown in Fig. 8. In contrast to the zero field results presented earlier in the talk, there is no sign of an intersection down to very low temperatures. Similar results have been found for even smaller values of

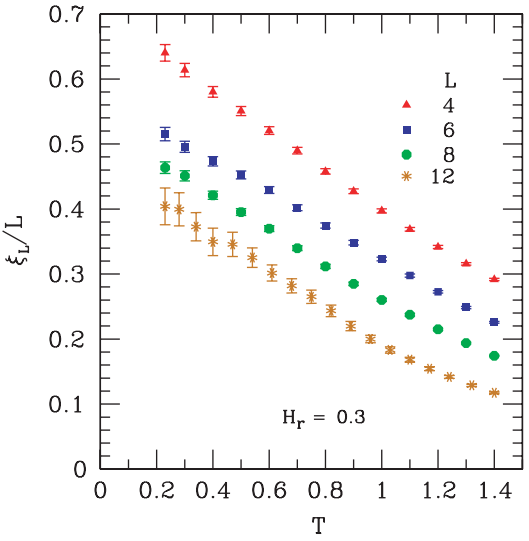


Fig. 8. Data for ξ_L/L for $H_{\rm r} = 0.3$ for different sizes. Note that in contrast to the zero-field data shown earlier there is no sign of intersections down to the lowest temperature $T = 0.23$. (From [23])

the field. This strongly suggests that there is no AT line in three-dimensional spin glasses.

This result is compatible with the droplet picture. However, other numerical results do not seem to be consistent with it. In particular the order parameter in the Parisi RSB theory is actually a probability distribution $P(q)$ with a delta function at $q = q_{EA}$ corresponding to ordering in a single valley, and a continuous part, with a non-zero weight at $q = 0$, coming when one of the two copies used to determine q according to (2) is in one valley and the other copy is in another valley. In the droplet picture, the weight at $q = 0$ vanishes with system size like $L^{-\theta}$ where θ is a positive exponent. Simulation results for $P(q)$, see e.g. [19, 20], agree much better with the RSB picture.

Hence the situation are rather confusing, and agree best with the intermediate "TNT" picture [15,16]. However, it is also possible that the system sizes that can be studied have not reached "asymptopia", where the behavior of the $L \to \infty$ limit becomes apparent. Unfortunately, there seem to be no ideas at present for algorithms which are so superior than the present ones that *significantly* bigger sizes can be studied, though *incremental* improvement can be expected.

5 Conclusions

The spin glass problem is hard. Analytical results are few and far between, and numerical studies are usually limited to quite small sizes. In fact, finding the ground state of a spin glass in three or more dimensions belongs to a hard class of optimization problems (NP-complete), for which it is believed that there is no algorithm which can find the exact ground state in a time which increases with only a power of the system size. However, the two-dimensional Ising glass is not NP-complete and there are efficient polynomial-time algorithms for determining the ground state [3]. As a result, lattices with of order a million spins can be studied in two dimensions, whereas in three dimensions the largest sizes have a few thousand spins.

In finite temperature Monte Carlo simulations, parallel tempering helps, but does not fully overcome the problem of slow dynamics. The results of simulations on modest-sized lattices below T_{SG} are not fully consistent with either the RSB or the droplet scenarios.

Acknowledgments

This work is supported by the NSF under grant No. DMR 0337049. I would also like to thank my collaborators, particularly Matteo Palassini, Helmut Katzgraber, and Lik Wee Lee.

References

1. K. Binder and A. P. Young (1986) *Rev. Mod. Phys.* **58**, p. 801
2. A. P. Young, ed. (1998) *Spin Glasses and Random Fields.* World Scientific, Singapore
3. R. G. Palmer and J. Adler (1999) *Int. J. Mod. Phys. C* **10**, p. 667
4. M. Mézard, G. Parisi, and R. Zecchina (2002) *Science.* **297**, p. 812
5. P. Nordblad and P. Svendlidh (1998) In *Spin Glasses and Random Fields.* edited by A. P. Young World Scientific, Singapore
6. R. Omari, J. J. Prejean, and J. Souletie (1983) *J. de Physique* **44**, p. 1069
7. S. F. Edwards and P. W. Anderson (1975) *J. Phys. F* **5**, p. 965
8. L. W. Lee and A. P. Young (2003) *Phys. Rev. Lett.* **90**, p. 227203 (cond-mat/0302371)
9. D. Sherrington and S. Kirkpatrick (1975) *Phys. Rev. Lett.* **35**, p. 1792
10. G. Parisi (1979) *Phys. Rev. Lett.* **43**, p. 1754
11. G. Parisi (1980) *J. Phys. A* **13**, p. 1101
12. J. R. L. de Almeida and D. J. Thouless (1978) *J. Phys. A* **11**, p. 983
13. D. S. Fisher and D. A. Huse (1986) *Phys. Rev. Lett.* **56**, p. 1601
14. D. S. Fisher and D. A. Huse (1988) *Phys. Rev. B* **38**, p. 386
15. F. Krzakala and O. C. Martin (2000) *Phys. Rev. Lett.* **85**, p. 3013 (cond-mat/0002055)
16. M. Palassini and A. P. Young (2000) *Phys. Rev. Lett.* **85**, p. 3017 (cond-mat/0002134)
17. K. Hukushima and K. Nemoto (1996) *J. Phys. Soc. Japan* **65**, p. 1604
18. E. Marinari, In *Advances in Computer Simulation*, edited by J. Kertész and I. Kondor (Springer-Verlag, 1998), p. 50 (cond-mat/9612010)
19. H. G. Katzgraber, M. Palassini, and A. P. Young (2001) *Phys. Rev. B* **63**, p. 184422 (cond-mat/0108320)
20. H. G. Katzgraber and A. P. Young (2003) *Phys. Rev. B* **67**, p. 134410 (cond-mat/0210451)
21. H. G. Ballesteros et al. (2000) *Phys. Rev. B* **62**, p. 14237 (cond-mat/0006211)
22. E. Marinari, G. Parisi, and J. J. Ruiz-Lorenzo (1998) *Phys. Rev. B* **58**, p. 14852
23. H. G. Katzgraber and A. P. Young (2004) *Phys. Rev. Lett.* **93**, p. 207203 (cond-mat/0407031)

Dipolar Fluctuations in the Bulk and at Interfaces

V. Ballenegger[1], R. Blaak[2], and J.-P. Hansen[2]

[1] Laboratoire de Physique Moléculaire, UMR CNRS 6624, Université de Franche-Comté, La Bouloie, 25030 Besançon cedex, France
[2] Department of Chemistry, Lensfield Road, Cambridge CB2 1EW, United Kingdom



Jean-Pierre Hansen

V. Ballenegger et al.: *Dipolar Fluctuations in the Bulk and at Interfaces*, Lect. Notes Phys. **704**, 45–63 (2006)
DOI 10.1007/3-540-35284-8_3
© Springer-Verlag Berlin Heidelberg 2006

1 Polar Fluids: A Closed Chapter? 47
2 Bulk Behaviour of Point and Extended Dipole Molecules 48
3 Dipolar Fluctuations in Confined Fluids 51
4 Slab Geometry 54
5 Spherical Geometry 56
6 Polarisation Effects 58
7 Summary and Outlook 60
References 61

The dielectric response of classical polar fluids is by now well understood for bulk systems, where the permittivity can be calculated by a variety of methods within the linear response regime. Near interfaces or inhomogeneities, one may attempt to describe the dielectric response of the fluid using a local dielectric tensor $\epsilon(\boldsymbol{r})$, for which an explicit expression can be derived from linear response theory. This chapter describes the limitations of this approach, exemplified by Molecular Dynamics simulations of polar fluids confined to a slit or a spherical cavity.

1 Polar Fluids: A Closed Chapter?

Many solvents are made up of highly polar molecules, water being the foremost, although somewhat atypical example (due to its hydrogen bonding capacity). Hence it is hardly surprising that a large body of theoretical work has gone into the understanding of polar liquids, and in particular of their dielectric response [1]. Although polar molecules are invariably non-spherical, much insight has been gained from integral equation theories [2] and simulations [3] of simple models involving spherical particles (e.g. hard spheres or Lennard–Jones particles) with an embedded point dipole, the Stockmayer potential being a much studied example. While this field was very active in the eighties and early nineties, it has slowed down since, giving the false impression that everything is well understood. This chapter addresses two open problems, which have only very recently received some attention:

(*a*) The point dipole limit is a valid approximation only for intermolecular distances significantly larger than their size. At shorter range, details of the molecular charge distribution become crucially important, as illustrated by the standard models for water (TIP5P, SPC/E, etc.). Hence it is of fundamental interest to investigate models with extended dipoles (constructed from two opposite charges separated by a finite distance d) and to determine how, for a given dipole moment, the extension d affects the dielectric properties and phase behaviour of dense polar fluids, compared to the point dipole limit.

(*b*) While the relation between the dielectric permittivity ϵ and dipolar fluctuations is well understood in bulk, virtually nothing is known about such fluctuations near an interface between a fluid and a dielectric medium mimicking e.g. a substrate, an electrode, or a membrane. In particular, is it legitimate to define a local permittivity $\epsilon(\boldsymbol{r})$ on a nanometric scale, and how is such a profile related to local fluctuations of the dipole moment?

This is of key importance for large-scale biomolecular simulations based on an implicit solvent assumption. The present chapter summarises some of the partial answers to these two open questions which we obtained recently [4–8].

2 Bulk Behaviour of Point and Extended Dipole Molecules

We consider spherical molecules carrying extended dipoles consisting of two opposite charges $\pm q$, displaced symmetrically by a distance $d/2$ from the centre of the molecule, such that the absolute dipole moment is $\mu = qd$. Obviously the intramolecular charge distribution $\rho(\boldsymbol{r}) = q\,[\delta(\boldsymbol{r} + \boldsymbol{d}/2) - \delta(\boldsymbol{r} - \boldsymbol{d}/2)]$ will give rise to higher order multipole moments, starting with an octopole.

Only in the limit $q \to \infty$ and $d \to 0$ for fixed μ (point dipole) will the electrostatic interaction between the molecules reduce to the familiar point dipole interaction:

$$v_{\mu,\mu}(1,2) = (\boldsymbol{\mu}_1 \cdot \boldsymbol{\nabla}_1)(\boldsymbol{\mu}_2 \cdot \boldsymbol{\nabla}_2) G(\boldsymbol{r}_1, \boldsymbol{r}_2) \tag{1}$$

where the Green's function $G(\boldsymbol{r}_1, \boldsymbol{r}_2)$ is the solution of Poisson's equation for a unit charge, subject to appropriate boundary conditions dictated by the surrounding media; for an unbounded system $G(\boldsymbol{r}_1, \boldsymbol{r}_2)$ reduces to the Coulomb potential $1/|\boldsymbol{r}_1 - \boldsymbol{r}_2|$.

In view of the fact that intramolecular charge distributions are always extended, we examined how the dipole extension, characterised by the ratio d/σ (where σ is the molecular diameter) affects the structural and dielectric properties of a dense polar fluid, for a fixed dipole moment μ. The question was investigated in some details by extensive Molecular Dynamics (MD) simulations reported in [6]. The polar molecules interact via a Lennard–Jones potential or a short-range soft-core potential $v_0(r) = 4u(\sigma/r)^{12}$, plus the long-range Coulomb interaction due to the two point charges $\pm q$ separated by a distance d inside each molecule.

The Hamiltonian of the system of N molecules is, under periodic boundary conditions, whereby the periodic replications of the basic arbitrary simulation cell of length L form an infinite sphere, which is itself embedded in an infinite region of permittivity ϵ' :

$$H = \sum_{i<j} [v_0(r_{ij}) + q_i q_j \psi(r_{ij})] - \frac{\kappa}{\sqrt{\pi}} \sum_{i=1}^{N} q_i^2 + \frac{2\pi |\boldsymbol{M}|^2}{(2\epsilon' + 1) L^3} \tag{2}$$

where $\boldsymbol{M} = \sum_{i=1}^{N} q_i \boldsymbol{r}_i$ is the total dipole moment of the sample and $\psi(r_{ij})$ is the usual Ewald sum over Coulombic interactions between two charges and their periodic images; κ is the inverse convergence length used in the Ewald summation [9].

The pair structure can be characterised either by site-site correlation functions $h_{++}(r) = h_{--}(r)$ and $h_{+-}(r)$, or by the molecular pair correlation function $h(1,2) = h(\boldsymbol{r}, \hat{\boldsymbol{\mu}}_1, \hat{\boldsymbol{\mu}}_2)$, which can be expanded in rotational invariants [10]:

$$h(1,2) = h^{000}(r) + h^{110}(r) \Phi^{110}(1,2) + h^{112}(r) \Phi^{112}(1,2) + \cdots \tag{3}$$

where

$$\Phi^{110}(1,2) = \hat{\boldsymbol{\mu}}_1 \cdot \hat{\boldsymbol{\mu}}_2 \tag{4a}$$

$$\Phi^{112}(1,2) = 3(\hat{\boldsymbol{\mu}}_1 \cdot \hat{\boldsymbol{r}})(\hat{\boldsymbol{\mu}}_2 \cdot \hat{\boldsymbol{r}}) - \hat{\boldsymbol{\mu}}_1 \cdot \hat{\boldsymbol{\mu}}_2 \tag{4b}$$

The bulk permittivity ϵ of the fluid may be determined by one of the following five routes (not mentioning applied field simulations):

(a) Via Kirkwood's fluctuation formula [11]

$$\begin{aligned} \frac{(\epsilon-1)(2\epsilon'+1)}{2\epsilon'+\epsilon} &= \frac{4\pi}{3V}\frac{\langle|\boldsymbol{M}|^2\rangle}{k_BT} \\ &= 3yg_\kappa \end{aligned} \tag{5}$$

valid for a macroscopic spherical sample of volume V embedded in a medium of permittivity ϵ'; y is the dimensionless parameter $4\pi\beta\rho\mu^2/9$ (with $\rho = N/V$ and $\beta = 1/k_BT$), while $g_\kappa = \langle|\boldsymbol{M}|^2\rangle/N\mu^2$ is the Kirkwood "g-factor", a measure of the orientational correlations between neighbouring dipoles. Onsager's celebrated result is recovered if correlations are neglected, such that $g_\kappa = 1$. The correct use of the Kirkwood's formula (5) in simulations, in conjunction with periodic boundary conditions (Ewald summations or Reaction Field), was clarified by Neumann [12]. It is obvious from (5) that the g-factor depends on the dielectric constant ϵ' around the macroscopic sample. The ϵ derived from (5) can however be shown to be independent of this boundary condition, both theoretically for an arbitrary sample shape [13,14], as well as in simulations performed in periodic boundary conditions. Kirkwood's formula (5) results from a simple linear response analysis which will be generalised to the inhomogeneous case in Sect. 3.

(b) A variant of route a is obtained by expliciting the Kirkwood "g-factor" in terms of the projection $h^{110}(r)$ of the pair distribution function (3):

$$\begin{aligned} g_\kappa &= 1 + \sum_{i\neq j}\frac{\langle\boldsymbol{\mu}_i\cdot\boldsymbol{\mu}_j\rangle}{N\mu^2} \\ &= 1 + \frac{4\pi\rho}{3}\int h^{110}_{\epsilon'}(r)r^2\mathrm{d}r \end{aligned} \tag{6}$$

where h^{110} depends sensitively on the permittivity ϵ' of the embedding medium [6,15].

(c) The projection $h^{112}(r)$ is related to ϵ by its asymptotic limit [16] which is independent of ϵ' for sufficiently large samples:

$$\lim_{r\to\infty} r^3 h^{112}(r) = \frac{(\epsilon-1)^2}{\epsilon}\frac{1}{4\pi\rho y} \tag{7}$$

Thus ϵ might also be extracted from Monte Carlo (MC) or MD data for $h^{112}(r)$, provided the sample is large enough for this projection to reach the asymptotic limit.

(d) In the case of extended (as opposed to point) dipoles, ϵ may also be calculated, in principle, from the Stillinger–Lovett perfect screening condition, valid for an ionic system [17–19].
The total charge-charge correlation function has intra and intermolecular contributions:

$$\begin{aligned} S(r) &= S_{\text{intra}}(r) + S_{\text{inter}}(r) \\ &= \left[2q^2\rho\delta(\boldsymbol{r}) - 2q^2\rho\frac{\delta(|\boldsymbol{r}|-d)}{4\pi d^2}\right] + 2q^2\rho^2\left[h_{++}(r) - h_{--}(r)\right] \end{aligned} \tag{8}$$

The second Stillinger–Lovett (or perfect screening) condition reads:

$$\frac{1-\epsilon}{\epsilon} = \frac{2\pi\beta}{3}\int r^2 S(r)\mathrm{d}\boldsymbol{r} \tag{9}$$

This route is in practice of little use for highly polar systems ($\epsilon \gg 1$), because of the unfavourable ratio $(1-\epsilon)/\epsilon$, which would require extremely accurate determinations of $S(r)$.

(e) For completeness, we mention finally Ramshaw's formula [13] which expresses the dielectric constant as an integral over the direct correlation function:

$$\frac{\epsilon-1}{\epsilon+2} = y\left[1 - \frac{\rho}{16\pi^2}\int \mathrm{d}\boldsymbol{r}\int \mathrm{d}\hat{\boldsymbol{\mu}}_1\int \mathrm{d}\hat{\boldsymbol{\mu}}_2\; c(\boldsymbol{r},\hat{\boldsymbol{\mu}}_1,\hat{\boldsymbol{\mu}}_2)\hat{\boldsymbol{\mu}}_1\cdot\hat{\boldsymbol{\mu}}_2\right]^{-1} \tag{10}$$

where the molecular direct correlation function c is related to the molecular pair correlation function h by the Ornstein–Zernike relation [2,10]:

$$\begin{aligned} h(\boldsymbol{r}_1,\hat{\boldsymbol{\mu}}_1,\boldsymbol{r}_2,\hat{\boldsymbol{\mu}}_2) =& c(\boldsymbol{r}_1,\hat{\boldsymbol{\mu}}_1,\boldsymbol{r}_2,\hat{\boldsymbol{\mu}}_2) \\ &+ \rho\int \mathrm{d}\boldsymbol{r}_3\int \mathrm{d}\hat{\boldsymbol{\mu}}_3 c(\boldsymbol{r}_1,\hat{\boldsymbol{\mu}}_1,\boldsymbol{r}_3,\hat{\boldsymbol{\mu}}_3)h(\boldsymbol{r}_3,\hat{\boldsymbol{\mu}}_3,\boldsymbol{r}_2,\hat{\boldsymbol{\mu}}_2) \end{aligned} \tag{11}$$

Route c yields accurate values of ϵ, provided a sufficiently large system is simulated for $h^{112}(r)$ to reach its asymptotic value. A similar requirement holds for route b, while route a can be used with smaller systems. An error analysis [6] shows that the choice of metallic boundary conditions at infinity ($\epsilon' = \infty$) is optimal in the sense that it minimises the uncertainty in ϵ for a given simulation time. The convergence with the number of timesteps is slow, because ϵ is related to fluctuations of the total dipole moment $\boldsymbol{M}$ around its mean $\langle\boldsymbol{M}\rangle = 0$. This is illustrated in Fig. 1 which shows that simulation times of several nanoseconds are required to arrive at a 5% accuracy.

The "sluggishness" of the convergence may be traced back to the slow decay of the autocorrelation function of the total dipole moment $\boldsymbol{M}(t)$. At liquid densities, the relaxation is essentially exponential with a relaxation time τ_M of typically 10 ps, an order of magnitude larger than the time associated

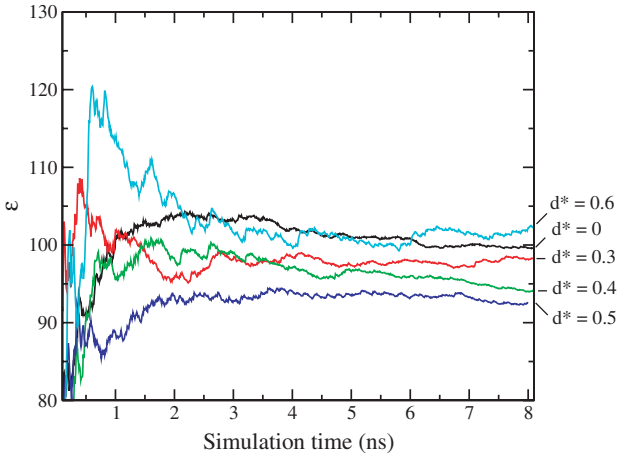


Fig. 1. Convergence of ϵ with simulation time, for dipole elongations $d^* = d/\sigma = 0$, 0.3, 0.4, 0.5 and 0.6.(After Ballenegger and Hansen [6])

with the single dipole relaxation, which follows a stretched exponential behaviour [6]. Note that τ_M increases with the elongation of d/σ, due to the enhanced tendency towards alignment (string formation) of the dipoles.

Comparing results for different elongations (but the same μ) shows that the extended dipole results begin to deviate from the point dipole data only for $d/\sigma \gtrsim 0.3$. As the ratio increases further, the dipolar molecules tend to form head-to-tail strings, and the polar fluid undergoes a transition to a hexagonal columnar phase [6], which occurs for much lower values of μ compared to the point dipole case [20, 21].

3 Dipolar Fluctuations in Confined Fluids

We now turn to the largely uncharted territory of the dielectric behaviour of confined polar fluids near dielectric interfaces. The general geometry is sketched in Fig. 2: N dipolar molecules are confined to a cavity of arbitrary shape carved out of a medium of macroscopic dielectric permittivity ϵ'. Two situations will be considered:

(a) $\epsilon' = 1$, i.e. the confining medium is non polarisable; in that case the dielectric behaviour of the fluid is modified due to the geometric confinement only.

(b) $\epsilon' > 1$, then the dielectric behaviour is affected both by geometric confinement and by the polarisation of the confining medium. Polarisation effects introduce boundary conditions at the interfaces, which can be handled either by the method of images [22], which is useful only for very simple geometries, or by a variational method based on the optimisation of

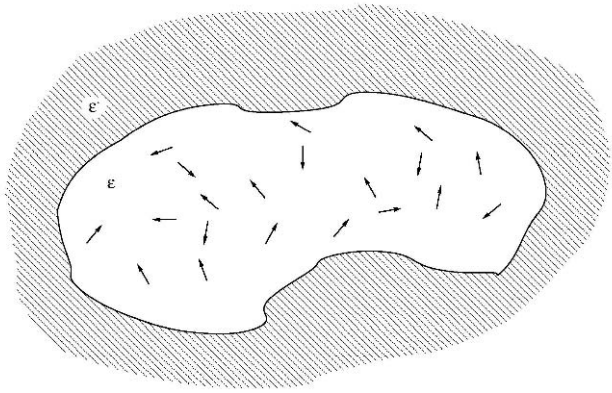


Fig. 2. Sketch of dipoles confined to a cavity embedded in a dielectric medium ϵ'

an appropriate functional with respect to the surface polarisation charge density [23–25].

The two key questions to be asked concerning the dielectric response of a polar fluid near a confining surface are the following:

(a) Is there a local, linear relationship between the induced mean polarisation density $\boldsymbol{P}(\boldsymbol{r})$ and the local (Maxwell) electric field $\boldsymbol{E}(\boldsymbol{r})$ of a form generalising the standard macroscopic expression [22], i.e.:

$$\boldsymbol{P}(\boldsymbol{r}) = \frac{\epsilon(\boldsymbol{r}) - \mathsf{I}}{4\pi} \boldsymbol{E}(\boldsymbol{r}) \tag{12}$$

where $\epsilon(\boldsymbol{r})$ is a local permittivity tensor (which reduces to a constant scalar in the bulk of an isotropic fluid)?

(b) If such a relation holds (at least for sufficiently weak applied field), what is the microscopic expression for $\epsilon(\boldsymbol{r})$, which generalises Kirkwood's relation (5)?

The limits of validity of the local relation (12) are discussed in references [5] and [16]. Roughly speaking a local relation holds provided the field does not vary appreciably over the range of the bulk pair correlation function. A general formal expression for $\epsilon(\boldsymbol{r})$ was derived in [5]. Expressions for specific, simple geometries will be given below.

The simplest geometry is a semi-infinite system confined (say to $z > 0$) by a plane ($z = 0$) and a dielectric medium ϵ' extending to negative z. Under these conditions the dielectric permittivity tensor, if it exists, is necessarily of the form:

$$\epsilon(z) = \begin{pmatrix} \epsilon_{\parallel}(z) & 0 & 0 \\ 0 & \epsilon_{\parallel}(z) & 0 \\ 0 & 0 & \epsilon_{\perp}(z) \end{pmatrix} \tag{13}$$

Exact expressions for $\epsilon_{\parallel}(z)$ and $\epsilon_{\perp}(z)$ in the asymptotic region close to the bulk limit may be derived and exhibit a $1/z^3$ variation [5]; the same is true of the anisotropic one-particle density $\rho(z,\theta)$ (where θ is the angle between a dipole and the z axis) [26]. Direct simulation of such a semi-infinite polar fluid is not possible, because periodic boundary conditions cannot be satisfied in the z-direction. Periodicity can only be achieved for a slab of polar fluid confined between two dielectric slabs, which may mimic, for instance, membranes or clay platelets.

As in the case of bulk polar fluids, the general procedure to derive a microscopic expression for the local permittivity $\epsilon(\boldsymbol{r})$ of a confined fluid is based on a linear response argument [1, 7, 27]. This proceeds in two steps: first the induced polarisation of the sample is related to the cavity field $\boldsymbol{E}^c$, i.e. the field created inside the *empty* cavity by an external field $\boldsymbol{E}'$. The second step is to relate the Maxwell field inside the *filled* cavity to the cavity field; $\epsilon(\boldsymbol{r})$ then follows from the definition (12). The first step may be formally carried out for a cavity of arbitrary shape. Let $\boldsymbol{m}(\boldsymbol{r})$ be the microscopic polarisation density:

$$\boldsymbol{m}(\boldsymbol{r}) = \sum_{i=1}^{N} \boldsymbol{\mu}_i \delta(\boldsymbol{r} - \boldsymbol{r}_i) \tag{14}$$

where $\boldsymbol{r}_i$ and $\boldsymbol{\mu}_i$ are the centre of mass position and dipole moment of the ith molecule ($1 \leq i \leq N$). The total dipole moment of the sample is:

$$\boldsymbol{M} = \int_{\text{cavity}} \boldsymbol{m}(\boldsymbol{r}) \mathrm{d}\boldsymbol{r} \tag{15}$$

The mean local polarisation density is the statistical average of (14). Let $\boldsymbol{E}'$ be a uniform externally applied field in the confining dielectric, far from the cavity. The induced polarisation density is:

$$\begin{aligned} \Delta\boldsymbol{P}(\boldsymbol{r}) &= \boldsymbol{P}(\boldsymbol{r}) - \boldsymbol{P}_0(\boldsymbol{r}) \\ &= \langle \boldsymbol{m}(\boldsymbol{r}) \rangle_{\boldsymbol{E}'} - \langle \boldsymbol{m}(\boldsymbol{r}) \rangle \end{aligned} \tag{16}$$

where the two terms are the polarisations in the presence and absence ($\boldsymbol{P}_0(\boldsymbol{r})$) of external field. In many situations symmetry implies that $\boldsymbol{P}_0(\boldsymbol{r}) = 0$. The general relation between $\Delta\boldsymbol{P}$ and the corresponding difference in Maxwell fields $\Delta\boldsymbol{E}(\boldsymbol{r}) = \boldsymbol{E}(\boldsymbol{r}) - \boldsymbol{E}_0(\boldsymbol{r})$ is:

$$\Delta\boldsymbol{P}(\boldsymbol{r}) = \frac{1}{4\pi} \int_{\text{cavity}} \chi(\boldsymbol{r}, \boldsymbol{r}') \cdot \Delta\boldsymbol{E}(\boldsymbol{r}') \mathrm{d}\boldsymbol{r}' \tag{17}$$

which reduces to (12) if a local relation $\chi(\boldsymbol{r}, \boldsymbol{r}') = \chi(\boldsymbol{r})\delta(\boldsymbol{r} - \boldsymbol{r}') = (\epsilon(\boldsymbol{r}) - \mathsf{I})\delta(\boldsymbol{r} - \boldsymbol{r}')$ holds and if $\boldsymbol{E}_0(\boldsymbol{r}) = 0$ by symmetry. Expliciting the statistical average $\langle \boldsymbol{m}(\boldsymbol{r}) \rangle_{\boldsymbol{E}'}$ in (16) we arrive at:

$$\Delta \boldsymbol{P}(\boldsymbol{r}) = \frac{\int [\boldsymbol{m}(\boldsymbol{r}) - \langle \boldsymbol{m}(\boldsymbol{r}) \rangle] \exp\{-\beta [U_{\epsilon'}(1,\ldots,N) - \boldsymbol{M}\cdot\boldsymbol{E}^c]\}\, \mathrm{d}1 \ldots \mathrm{d}N}{\int \exp\{-\beta [U_{\epsilon'}(1,\ldots,N) - \boldsymbol{M}\cdot\boldsymbol{E}^c]\}\, \mathrm{d}1 \ldots \mathrm{d}N} \tag{18}$$

where $\mathrm{d}i = \mathrm{d}\boldsymbol{r}_i \mathrm{d}\hat{\boldsymbol{\mu}}_i$, $U_{\epsilon'}$ is the total interaction energy of N dipolar molecules within the cavity embedded in the dielectric ϵ', while the total dipole moment is coupled to the cavity (rather than the external) field.

Linearisation of the Boltzmann factors with respect to $\boldsymbol{E}^c$ immediately leads to the linear response result:

$$\Delta \boldsymbol{P}_\alpha(\boldsymbol{r}) = \beta \sum_{\gamma=x,y,z} [\langle m_\alpha(\boldsymbol{r}) M_\gamma \rangle - \langle m_\alpha(\boldsymbol{r}) \rangle \langle M_\gamma \rangle]\, E_\gamma \tag{19}$$

where the statistical averages are to be taken over the unperturbed system (i.e. for $\boldsymbol{E}' = 0$).

The next task is to relate $\boldsymbol{E}^c$ to $\boldsymbol{E}$; this depends on the geometry of the cavity and the permittivity ϵ'. In Sect. 4 we consider the slab geometry, while the case of a spherical cavity will be examined in Sects. 5 and 6.

4 Slab Geometry

The "cavity" reduces to an infinite slab in the x and y directions, confined between two semi-infinite dielectric media ϵ'. The interfaces are planes orthogonal to z. The local dielectric tensor is of the form (13). By symmetry, the component of $\boldsymbol{P}$ parallel to the planes vanishes in the absence of an external field, so that (12) splits into:

$$\boldsymbol{P}_\parallel(z) = \frac{\epsilon_\parallel(z) - 1}{4\pi} \boldsymbol{E}_\parallel(z) \tag{20a}$$

$$P_\perp(z) = \frac{\epsilon_\perp(z) - 1}{4\pi} E_\perp(z) \tag{20b}$$

where all vectors with the subscript $\parallel$ are 2d vectors in the x-y plane.

The standard electrostatic boundary conditions relate the cavity to the external field, i.e.:

$$\begin{aligned} \boldsymbol{E}^c_\parallel &= \boldsymbol{E}'_\parallel \\ \boldsymbol{E}^c_\perp &= \epsilon' \boldsymbol{E}'_\perp \end{aligned} \tag{21}$$

Maxwell's equation $\boldsymbol{\nabla} \wedge \boldsymbol{E}(z) = 0$ shows that $\boldsymbol{E}_\parallel$ is independent of z and hence $\boldsymbol{E}_\parallel = \boldsymbol{E}'_\parallel$ everywhere. Remembering (20a), we conclude that:

$$\boldsymbol{P}_\parallel(z) = \frac{\epsilon_\parallel(z) - 1}{4\pi} \boldsymbol{E}^c_\parallel \tag{22}$$

Equations (19) and (22), together with the isotropy in the x-y plane then imply the following generalised Kirkwood relation for $\epsilon_\parallel(z)$:

$$\epsilon_{\parallel}(z) = 1 + 2\pi\beta \left[\langle \boldsymbol{m}_{\parallel}(z) \cdot \boldsymbol{M}_{\parallel} \rangle - \langle \boldsymbol{m}_{\parallel}(z) \rangle \cdot \langle \boldsymbol{M}_{\parallel} \rangle \right] \tag{23}$$

Note that the dipolar density at z must be correlated not with itself (as one might naively have guessed) but with the total dipole moment of the slab.

A similar calculation, now using the Maxwell relation $\boldsymbol{\nabla} \cdot \boldsymbol{D}(z) = 0$ (with $\boldsymbol{D} = \boldsymbol{E} + 4\pi\boldsymbol{P}$), leads to the following expression for $\epsilon_{\perp}(z)$ [7]:

$$\frac{\epsilon_{\perp}(z) - 1}{\epsilon_{\perp}(z)} = 1 + 2\pi\beta \left[\langle m_{\perp}(z) M_{\perp} \rangle - \langle m_{\perp}(z) \rangle \langle M_{\perp} \rangle \right] \tag{24}$$

One immediately notes that the unfavourable ratio on the left hand side will make it very difficult to extract accurate values of $\epsilon_{\perp}(z)$ from simulation estimates of the fluctuations on the right hand side.

A first attempt to extract $\epsilon_{\parallel}(z)$ and $\epsilon_{\perp}(z)$ from MD simulations was made in [7], for a system of "soft spheres" (short-range pair interactions $v(r) = 4u(\sigma/r)^{12}$) with extended dipoles ($d/\sigma = 1/3$), confined to a slab of width L, surrounded on both sides by vacuum ($\epsilon' = 1$). The electrostatic interaction of a periodic array of such slabs can be handled by 3d Ewald summations with a dipole layer correction term [28–30]. $\epsilon_{\parallel}(z)$ and $\epsilon_{\perp}(z)$ can be computed from the fluctuation formulae (23) and (24), or by applying a uniform external field $\boldsymbol{E}'$ parallel (for $\epsilon_{\parallel}$) or perpendicular (for $\epsilon_{\perp}$) to the slab. In the parallel case a measurement of $\boldsymbol{P}_{\parallel}(z)$ then directly determines $\epsilon_{\parallel}(z)$ via (22), since $\boldsymbol{E}^c_{\parallel} = \boldsymbol{E}'_{\parallel}$. $\epsilon_{\perp}(z)$ may also be estimated in a similar way [7]. Results for $\epsilon_{\parallel}(z)$ and two different dipole moments are plotted in Fig 3. The two

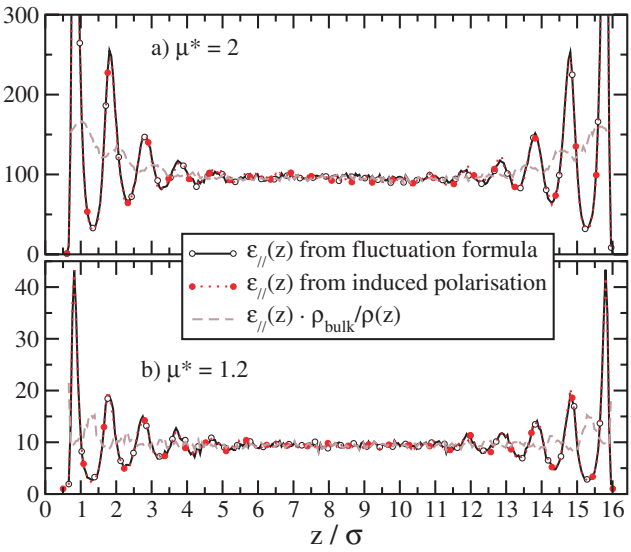


Fig. 3. Parallel component of the permittivity tensor from fluctuation formula (23) and from the response to an external field $E' = 0.1$ V/nm along the x axis. (Reused with permission from [7]. Copyright 2005, American Institute of Physics)

estimates of $\epsilon_{\parallel}(z)$ based on (22) and (23) are in excellent agreement. The pronounced oscillations of $\epsilon_{\parallel}(z)$ near the two interfaces are intimately correlated with the oscillations in the density profile $\rho(z)$, which reflect the layering of the dipolar spheres near the walls, as is clear form the plots of $\epsilon_{\parallel}(z)\rho_{\text{bulk}}/\rho(z)$, where ρ_{bulk} is the bulk density of the fluid, reached half-way between the two plates where $\epsilon(z)$ tends to its bulk value. The behaviour of the "envelope" $\epsilon_{\parallel}(z)\rho_{\text{bulk}}/\rho(z)$ near the interfaces suggests that on a suitably coarse-grained scale, $\epsilon(z)$ tends to increase over its bulk value on approaching the dielectric walls. This behaviour may change qualitatively when the dielectric media on both sides of the slab are polarisable ($\epsilon' > 1$) or carry a surface charge; these issues will be addressed in future simulation work.

As anticipated, $\epsilon_{\perp}(z)$ fails to converge to physically acceptable values, even for very long (several ns) MD runs, except in the bulk region, far from the interfaces, where it tends to the same limit as $\epsilon_{\parallel}(z)$.

An unexpected result of the simulations in [7] is the dramatic "overscreening" of an externally applied field by the local polarisation density in the immediate vicinity of the interfaces, where the ratio $E(z)/E'$ can become very large and negative (typically $\simeq -2$).

The general behaviour described in this section is not specific to dipolar soft spheres, but remains at least qualitatively very similar for SPC/E water confined to a slit [7].

5 Spherical Geometry

We next consider a dipolar fluid confined to a spherical cavity of radius R carved out of a dielectric medium ϵ'. The cavity field is now $\boldsymbol{E}^c = 3\epsilon'\boldsymbol{E}'/(2\epsilon'+1)$ [22], and adaptation of the linear response argument of Sect. 3 to this isotropic case leads to

$$\frac{(\epsilon-1)(2\epsilon'+1)}{2\epsilon'+\epsilon} = \frac{4\pi\beta}{3V}\left[\langle \boldsymbol{m}(\boldsymbol{r})\cdot\boldsymbol{M}\rangle - \langle \boldsymbol{m}(\boldsymbol{r})\rangle\cdot\langle \boldsymbol{M}\rangle\right] \tag{25}$$

where $\boldsymbol{r}$ can be any point in the bulk of the system, except in the vicinity of the spherical interface. Away from the interface, $\boldsymbol{m}(\boldsymbol{r})$ may hence be replaced by $\boldsymbol{M}/V$ so that (25) leads back to Kirkwood's formula (5) (with $\langle \boldsymbol{M}\rangle = 0$, which in simulations is only achieved for sufficiently long runs).

We now consider the case of a radial, external field due to an "external" charge q placed at the centre of the cavity filled with polar molecules, so that $\boldsymbol{E}'(\boldsymbol{r}) = q\hat{\boldsymbol{r}}/r^2$. In that case the permittivity depends only on the radial distance from the centre, and the linear response argument leads to the microscopic expression [7]:

$$\frac{\epsilon(r)-1}{\epsilon(r)} = 4\pi\beta\int_{\text{cavity}}\langle m(\boldsymbol{r})m(\boldsymbol{r}')\rangle\left(\frac{r}{r'}\right)^2 \mathrm{d}\boldsymbol{r}' \tag{26}$$

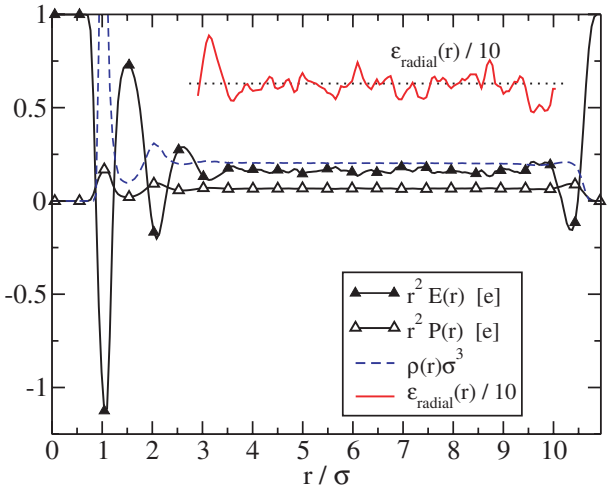


Fig. 4. Radial electric field, polarisation density, molecular density, and permittivity profiles for a spherical droplet of polar fluid ($\mu^* = 2$, $T^* = 1.35$, $\rho^* = 0.2$) when an ion of unit electronic charge (reduced charge $q^* = 28.7$) is present at the origin. The *dotted line* indicates the bulk dielectric constant $\epsilon = 6.3$ (divided by 10). (Reused with permission from [7]. Copyright 2005, American Institute of Physics)

where $m = \boldsymbol{m} \cdot \hat{\boldsymbol{r}}$ is the radial projection of the microscopic polarisation density (14). Note that, contrary to (25), this relation does not depend explicitly on ϵ' (but of course implicitly through the Boltzmann weight in the statistical average).

Equation (26) is not very convenient for computational purposes, because of the integration on the right hand side. The local and applied fields are related by the intuitively satisfactory relation [7]

$$E(r) = \frac{E'(r)}{\epsilon(r)} \tag{27}$$

which is a direct consequence of the local assumption (12). The profile $E(r) = q(r)/r^2$, where $q(r) = q - 4\pi r^2 P(r)$ is the charge inside a sphere of radius r and $P(r)$ is the radial component of the polarisation density, which may be measured via definition (16) (with $P_0 = 0$ for linear dipoles). Examples of the various profiles, when an ion of charge $q^* = q\mu/(\sigma^2 u) = 28.7$ is placed at the origin, are shown in Fig. 4 in the case where $\epsilon' = 1$. The resulting $\epsilon(r)$ is compatible with the bulk value away from the centre and the interface, but is ill-defined as the latter are approached. Note that the "overscreening" effect, observed in the slab geometry, is also very marked near the central charge.

6 Polarisation Effects

So far we have carefully avoided the complications due to the polarisation charge induced on the confining surfaces by the dipoles of the confined polar fluid, by assuming $\epsilon' = 1$ in all simulations. For $\epsilon' > 1$ electrostatic boundary conditions must be satisfied at the surface separating the confining medium from the atomistically resolved polar fluid, where molecules evolve in vacuo ($\epsilon = 1$). Let $\boldsymbol{n}$ be the normal to the surface carrying a surface charge density σ (only the case $\sigma = 0$ will be considered throughout). The electric fields $\boldsymbol{E}'$ and $\boldsymbol{E}$ on either side of the surface obey the conditions [22]

$$(\epsilon' \boldsymbol{E}' - \epsilon \boldsymbol{E}) \cdot \boldsymbol{n} = 4\pi\sigma \tag{28a}$$

$$(\boldsymbol{E}' - \boldsymbol{E}) \wedge \boldsymbol{n} = 0 \tag{28b}$$

For sufficiently simple geometries, the solution to Maxwell's equations $\boldsymbol{\nabla} \wedge \boldsymbol{E} = 0$ and $\boldsymbol{\nabla} \cdot (\epsilon \boldsymbol{E}) = 0$, subject to the boundary conditions (28) may be conveniently obtained by the introduction of image charges. In the simplest case of a single planar surface separating the two media as depicted in Fig. 5, the electrostatic potential $\Phi(\boldsymbol{r})$ inside the medium ϵ due to a single charge q is the sum of the potentials due to that charge and a single image charge $q' = (\epsilon - \epsilon')q/(\epsilon + \epsilon')$, positioned at the mirror location with respect to the planar surface, in the absence of the dielectric discontinuity. In the case of the slab geometry used in simulations, an infinite array of images is required to satisfy the boundary conditions. In the case of a spherical cavity the problem turns out to be highly non-trivial [8], except in the case where the confining medium is a metal ($\epsilon' = \infty$) [22]. Referring to Fig. 6, the electrostatic potential at $\boldsymbol{r}$ due to a simple charge q placed at $\boldsymbol{d}$ is [8]:

$$\begin{aligned} \Phi(\boldsymbol{r}) = & \frac{q}{4\pi\epsilon} \left\{ \frac{1}{|\boldsymbol{r} - \boldsymbol{d}|} + (1 - 2\kappa) \frac{R/d}{|\boldsymbol{r} - \boldsymbol{D}|} \right\} \\ & + \frac{q(1 - 2\kappa)}{4\pi\epsilon' R} F_1 \left(\kappa, \frac{1}{2}, \frac{1}{2}, 1 + \kappa; x\mathrm{e}^{\mathrm{i}\theta}, x\mathrm{e}^{-\mathrm{i}\theta} \right) \end{aligned} \tag{29}$$

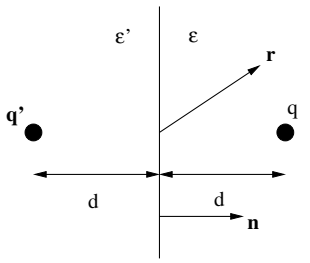


Fig. 5. The electrical potential field at $\boldsymbol{r}$ due to a single charge q in the case of a planar interface separating the media ϵ and ϵ', is effectively described by the combined contributions of the charge q and its image q'

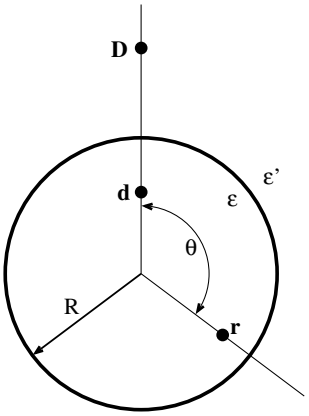


Fig. 6. Schematic representation of a charge at position $\boldsymbol{d}$ inside a spherical cavity of radius R and permittivity ϵ embedded in an infinite medium ϵ'

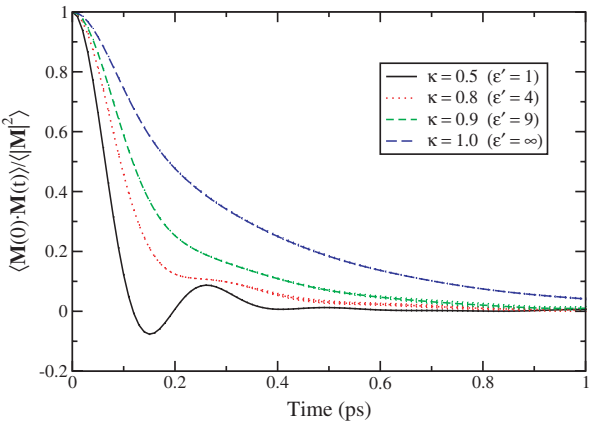


Fig. 7. Time correlation function $\langle \boldsymbol{M}(0) \cdot \boldsymbol{M}(t) \rangle$ of the total dipole moment of a polar fluid inside a spherical cavity surrounded by a dielectric medium with $\epsilon' = 1$, 4, 9, and ∞ ($N = 1000$, $R = 3$ nm)

where F_1 is a hyper-geometric function in the two complex variables $x\mathrm{e}^{\mathrm{i}\theta}$ and $x\mathrm{e}^{-\mathrm{i}\theta}$, with $x = rd/R^2$, $\kappa = \epsilon'/(\epsilon + \epsilon')$, and $\boldsymbol{D} = (R/d)^2\boldsymbol{d}$. The classic result for a cavity inside a metal is recovered when $\epsilon \to \infty$ or $\kappa \to 1$.

The dielectric behaviour of a polar fluid trapped inside a spherical cavity is very sensitive to the dielectric permittivity ϵ' of the confining medium. This is illustrated in Fig. 7, which shows the correlation function of the total dipole moment $\boldsymbol{M}(t)$ of $N = 1000$ dipolar molecules with $\mu^* = 2$ trapped inside a cavity of radius $R = 3$ nm. The oscillatory relaxation observed for $\epsilon' = 1$ (vacuum) is seen to go over into a very slow relaxation when $\epsilon' = \infty$. The slowing down of the relaxation as ϵ' increases agrees with the increase

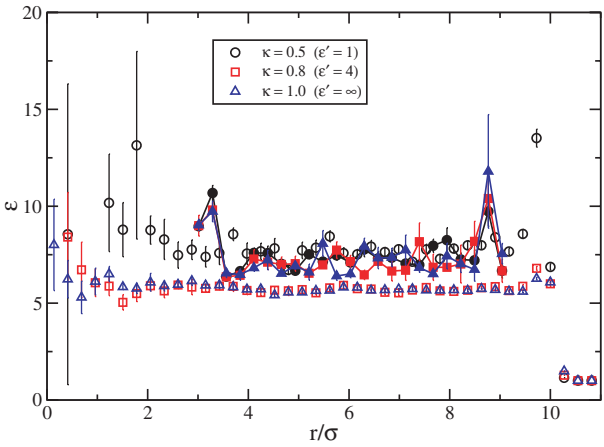


Fig. 8. Dielectric profiles $\epsilon(r)$ inside a spherical cavity, with a radius of 4 nm and 1000 dipoles, surrounded by a dielectric medium with $\epsilon' = 1$, 4, and ∞ obtained by the fluctuation formula (25) (*open symbols*) and (non-) linear response (27) (*filled symbols*)

of the Debye relaxation time in bulk dielectrics, predicted by Neumann and Steinhauser [31]. Permittivity profiles, $\epsilon(r)$, as estimated, for three values of ϵ', from the fluctuation formula (25), and from the (non-linear) response to a central charge (27) are shown in Fig. 8. In the case $\epsilon' = 1$, the linear and non-linear responses are seen to be rather close. For $\epsilon' = 4$ and $\epsilon' = \infty$, the linear profiles are significantly lower than their $\epsilon' = 1$ counterpart, while the non-linear response profiles appear to be less sensitive to the value of ϵ'.

7 Summary and Outlook

The main message to be taken away from this chapter is that a microscopic computation of the dielectric response of a polar fluid, and in particular of the dielectric permittivity in the bulk and at interfaces, is still an arduous and non-trivial task, despite several decades of efforts due to the following points:

(*a*) The dielectric behaviour of fluids of molecules with extended dipoles closely matches that of point dipoles for extensions $d/\sigma \lesssim 0.3$. Dramatic differences appear for $d/\sigma \gtrsim 0.5$, due to a stronger tendency towards head-to-tail string formation.

(*b*) Kirkwood's fluctuation formula (5) remains the most efficient route to estimating the bulk permittivity ϵ by computer simulations, with "metallic" boundary conditions ($\epsilon' = \infty$) at infinity. It must be emphasised, however, that MD trajectories of at least a few ns are required to achieve an accuracy of the order of 5%. Alternative estimates based on the asymptotic

behaviour of partial correlation functions require larger system sizes and tend to be less accurate.

(*c*) Assuming a local relation between polarisation and electric field profiles, one can derive generalised fluctuation formulae for the components of a local permittivity tensor $\epsilon(\boldsymbol{r})$. These formulae depend on the geometry of the problem, i.e. on the shape of the confining surface(s), and on the dielectric permittivity ϵ' of the confining medium, which is treated as a dielectric continuum. Whenever this medium is polarisable ($\epsilon' > 1$), the dielectric boundary conditions at the confining surface must be accounted for by the method of images (which is tractable for the simplest geometries only), or by a recently developed variational procedure for the calculations of the surface polarisation charge density.

We have shown that in the slab geometry the longitudinal permittivity $\epsilon_{\parallel}(z)$ appears to be well defined, and strongly influenced by the layering of densely packed polar molecules near the wall. The transverse component $\epsilon_{\perp}(z)$ is, however, ill-defined and may take on unphysical (negative) values near the walls, thus pointing to the break-down of the assumption of locality.

Preliminary MD data for a polar fluid inside a spherical cavity point to a significant dependence of the dielectric response of the sample to the permittivity ϵ' of the confining medium. Although spatially varying local permittivities are being used in implicit solvent simulation of biomolecular assemblies, the conclusion to be drawn from the preliminary results presented in this chapter is that the use of such permittivity profiles on nanoscales is dubious for the least. Moreover the simultaneous use of a fully molecular description of the confined polar fluid, and of a continuum representation of the confining medium (e.g. solid substrates, colloids or membranes) is obviously inconsistent, and introduces artificially sharp dielectric discontinuities at the surfaces. A more satisfactory, but also more computer-intensive model of polar fluids at interfaces would be to use an atomistic representation of the confining medium. We are presently replacing continuous dielectric slabs by polarisable atoms on a lattice. The polarisation of substrate atoms by the charge distribution on the molecules of the polar fluid can be handed by well established, self-consistent methods [32]. The dielectric discontinuity is thus smoothed out over atomistic scales, and a comparison will be made between the dielectric response of a polar fluid near a continuous or atomically resolved substrate.

References

1. P. A. Madden and D. Kivelson (1984) A consistent molecular treatment of dielectric phenomena. *Adv. Chem. Phys.* **56**, p. 467.
2. J. P. Hansen and I. R. McDonald (1986) *Theory of Simple Liquids.* Academic Press, London, 2nd ed. (Chap. 12)
3. P. G. Kusalik, M. E. Mandy, and I. M. Svishchev (1994) The dielectric constant of polar fluids and the distribution of the total dipolemoment. *J. Chem. Phys.* **100**, p. 7654 (and references therein)

4. R. Finken, V. Ballenegger, and J.-P. Hansen (2003) Onsager model for a variable dielectric permittivity near an interface. *Mol. Phys.* **101**, p. 2559
5. V. Ballenegger and J.-P. Hansen (2003) Local dielectric permittivity near an interface. *Europhys. Lett.* **63**, p. 381
6. V. Ballenegger and J.-P. Hansen (2004) Structure and dielectric properties of polar fluids with extended dipoles: results from numerical simulations. *Mol. Phys.* **102**, p. 599
7. V. Ballenegger and J.-P. Hansen (2005) Dielectric permittivity profiles of confined polar fluids. *J. Chem. Phys.* **122**, p. 114711
8. R. Blaak and J.-P. Hansen (2006) Dielectric response of a polar fluid trapped in a spherical nanocavity. *J. Chem. Phys.* **124**, p. 144714
9. J.-P. Hansen (1986) Molecular-Dynamics Simulation of Coulomb Systems in Two and Three Dimensions. In *Molecular Dynamics Simulation of Statistical Mechanical Systems*. Edited by G. Ciccotti and W. G. Hoover, North-Holland, Amsterdam, Proceedings of the International School of Physics "Enrico Fermi," Course XCVIII, Varenna, 1985, pp. 89–129.
10. C. G. Gray and K. E. Gubbins (1984) *Theory of Molecular Fluids*. vol. 1, Clarendon Press, Oxford
11. J. G. Kirkwood (1939) The dielectric polarization of polar liquids. *J. Chem. Phys.* **7**, p. 911
12. M. Neumann (1983) Dipole moment fluctuation formulas in computer simulations of polar systems. *Mol. Phys.* **50**, p. 841
13. J. D. Ramshaw (1977) Existence of the dielectric constant in rigid dipole fluids: The functional derivative approach. *J. Chem. Phys.* **66**, p. 3134
14. A. Alastuey and V. Ballenegger (2000) Statistical mechanics of dipolar fluids: dielectric constant and sample shape. *Physica A* **279**, p. 268
15. J. M. Caillol (1992) Asymptotic behavior of the pair-correlation function of a polar liquid. *J. Chem. Phys.* **96**, p. 7039
16. G. Nienhuis and J. M. Deutch (1971) Structure of dielectric fluids. I. the two particle distribution function of polar fluids. *J. Chem. Phys.* **55**, p. 4213
17. F. H. Stillinger, Jr. and R. Lovett (1968) Ion-pair theory of concentrated electrolytes. I. basic concepts. *J. Chem. Phys.* **48**, p. 3858
18. R. Lovett and F. H. Stillinger, Jr. (1968) Ion-pair theory of concentrated electrolytes. II. approximate dielectric response calculation. *J. Chem. Phys.* **48**, p. 3869
19. Ph. A. Martin (1988) Sum rules in charged fluids. *Rev. Mod. Phys.* **60**, p. 1075
20. D. Wei and G. N. Patey (1992) Orientational order in simple dipolar liquids: Computer simulation of a ferroelectric nematic phase. *Phys. Rev. Lett.* **68**, p. 2043
21. J. J. Weis, D. Levesque, and G. J. Zarragoicoechea (1992) Orientational order in simple dipolar liquid-crystal models. *Phys. Rev. Lett.* **69**, p. 913
22. J. D. Jackson (1999) *Classical Electrodynamics*. Wiley, New York, 3rd ed.
23. R. Allen, J.-P. Hansen, and S. Melchionna (2001) Electrostatic potential inside ionic solutions confined by dielectrics: a variational approach. *Phys. Chem. Chem. Phys.* **3**, p. 4177
24. R. Allen and J.-P. Hansen (2003) Electrostatic interactions of charges and dipoles near a polarizable membrane. *Mol. Phys.* **101**, p. 1575
25. P. Attard (2003) Variational formulation for the electrostatic potential in dielectric continua. *J. Chem. Phys.* **119**, p. 1365

26. J. P. Badiali (1989) Structure of a polar fluid near a wall, exact asymptotic behavior of the profile, relation with the electrostriction phenomena and the kerr effect. *J. Chem. Phys.* **90**, p. 4401
27. H. A. Stern and S. E. Feller (2003) Calculation of the dielectric permittivity profile for a nonuniform system: Application to a lipid bilayer simulation. *J. Chem. Phys.* **118**, p. 3401
28. I.-C. Yeh and M. L. Berkowitz (1999) Ewald summation for systems with slab geometry. *J. Chem. Phys.* **111**, p. 3155
29. A. Arnold, J. de Joannis, and C. Holm (2002) Electrostatics in periodic slab geometries. *J. Chem. Phys.* **117**, p. 2496
30. J. de Joannis, A. Arnold, and C. Holm (2002) Electrostatics in periodic slab geometries. *J. Chem. Phys.* **117**, p. 2503
31. M. Neumann and O. Steinhauser (1983) On the calculation of the frequency-dependent dielectric constant in computer simulations. *Chem. Phys. Lett.* **102**, p. 508
32. G. Lamoureux and B. Roux (2003) Modeling induced polarization with classical drude oscillators: Theory and molecular dynamics simulation algorithm. *J. Chem. Phys.* **119**, p. 3025 (and references therein)

Theory and Simulation of Friction and Lubrication

M.H. Müser

Department of Applied Mathematics, University of Western Ontario, London, ON, Canada N6A 5B7
mmuser@uwo.ca



Martin H. Müser

M.H. Müser: *Theory and Simulation of Friction and Lubrication*, Lect. Notes Phys. **704**, 65–104 (2006)
DOI 10.1007/3-540-35284-8_4
© Springer-Verlag Berlin Heidelberg 2006

1 Introduction 67

2 Theoretical Background 68

2.1 Friction Mechanisms 69
2.2 Velocity-Dependence of Friction 72
2.3 Load-Dependence of Friction and Contact Mechanics 74
2.4 Role of Interfacial Symmetry 75
2.5 Common Toy Models 77

3 Computational Aspects 78

3.1 Imposing Load and Shear 78
3.2 Including Surface Roughness and Elastic Deformations 80
3.3 Imposing Constant Temperature 84
3.4 Determination of Bulk Viscosities 89

4 Selected Case Studies 95

4.1 Superlubricity and the Role of Roughness at the Nanometer Scale 95
4.2 Physics and Chemistry of Lubricant Additives 99

References 101

Molecular dynamics (MD) and related simulation techniques have proven indispensable in unraveling the microscopic origins of many tribological phenomena such as friction, lubrication, and wear. This chapter is meant to serve as a guide for conducting MD simulations to further deepen our understanding of the processes that occur when two surfaces are in relative sliding motion. Some of the key mechanisms leading to friction will be discussed first. Knowledge of these mechanisms is imperative to both set up and interpret the results of simulations. However, the focus of this chapter will be on technical aspects such as how to construct realistic surface profiles and how to impose load, shear, and temperature during simulations. Finally, a few selected MD studies will be presented.

1 Introduction

Atomistic simulations of friction between solids have received growing attention in the last decade. This increase of interest has been spurred by the miniaturization of mechanical devices, the peculiar behaviour of condensed matter at the nanoscale, and advances in simulating ever more accurately chemically complex lubricants and surfaces [1–3]. Computer simulations have contributed significantly to the identification of dissipation mechanisms and, in some cases, have overthrown previously established explanations of the origin of friction. Despite impressive progress in the field, many open questions of scientific and technological interest persist. To name a few, the microscopic mechanisms leading to sliding-induced wear remain elusive, it is unclear how the interactions between lubricant additives affect lubricant performance and even the friction mechanisms in many relatively well-defined, nano-scale systems have not yet been convincingly identified. This chapter will address these and related issues and point to some unsolved questions. However, the main objective will be to provide guidelines as to how to conduct tribological simulations.[1]

Sliding surfaces are generally in a non-equilibrium state. One of the difficulties in simulating materials far away from equilibrium is the lack of a general principle such as minimization of free energy. Along these lines, the equivalence of ensembles in the thermodynamic limit does not apply to many tribological systems. This can be illustrated as follows. In a first simulation of two sliding surfaces, the kinetic friction force, F_{k}, is determined at constant load and constant sliding velocity. If a consecutive simulation is run in which similar parameters are employed except that now the separation is constrained to d, then one may well obtain completely different values for F_{k} from the two simulations. This example demonstrates that implementing boundary conditions properly is crucial in non-equilibrium simulations if we want to make

[1] Tribology is the science of surfaces in relative motion. It is sometimes also defined as the science of friction, lubrication, and wear.

reliable predictions. It also illustrates one of many pitfalls that can diminish considerably the value of a largely well-designed tribological simulation. These pitfalls often result from convenience where the unrealistic treatment, in this case constant separation, is easier to implement than the experimental condition, such as constant normal load.

Another important trap to avoid is the use of any artificial symmetries that do not exist in experiments. Unfortunately, it is almost common practice to simulate two identical, perfectly aligned surfaces, which are termed commensurate, although such interfaces are known to behave in a manner that is qualitatively different than that observed when two dislike, or misoriented surfaces, called incommensurate, rub against each other. Neglecting surface curvature or not allowing lubricant to become squeezed out of the contact can also lead to behaviour that would not be found in a regular laboratory experiment. Additionally, temperature control may be a more sensitive issue for systems far from equilibrium than in equilibrium simulations. By imposing shear, we constantly pump energy into the system, which needs to be removed. Thermostating naively may induce unrealistic velocity profiles in the sheared lubricant or lead to other undesired artifacts.

Before discussing the technical aspects of atomistic friction simulations, this chapter will give a small overview of the theoretical background of friction at small velocities. Without such a background, it is difficult to ask meaningful questions and to interpret the outcome of a simulation. After all, our goal goes beyond simply reproducing experimental results. It is often helpful to classify the processes into categories such as linear response, out-of-equilibrium steady state, and strongly irreversible. Moreover, a good understanding of the theoretical background will aid in determining which aspects of the simulations deserve particular focus and which details are essentially irrelevant. In some situations, for example when simulating boundary lubricants exposed to large pressures, the results are typically insensitive to the precise choice of the thermostat, while the opposite holds if the two sliding surfaces are separated by a high Reynold number fluid, or if polymers are grafted onto each surface.

In Sect. 2, some theoretical aspects of friction between solids will be explained. Section 3 contains an overview of algorithms that have been used in the simulation of tribological phenomena, and some selected case studies will be presented in Sect. 4.

2 Theoretical Background

Every-day experience tells us that a finite threshold force, namely the static friction force, F_s, has to be overcome whenever we want to initiate lateral motion of one solid body relative to another. Conversely, when attempting to drag a solid through a fluid medium, there is no such threshold. Instead, one only needs to counteract friction forces linear in the (final) sliding velocity v_0.

It came as a surprise when Hirano and Shinjo suggested that static friction between solids in ultra-high vacuum may essentially disappear as well [4, 5]. While their suggestion contradicts our intuition, which is based on everyday experience, it does not necessarily contradict Newtonian mechanics. If the slider and substrate have homogeneous surfaces and wear and plastic deformation are negligible, then one may expect the same (free) energy at the beginning of the sliding process as at its end, because of translational invariance. In such a case, no work would have to be done on the system implying the possibility of ultra-low friction, or, in the words by Hirano and Shinjo, superlubricity.

The microscopic justification for the possibility of the virtual absence of lateral forces between solids can be supported by the following argument: There are as many bumps (or atoms) in the substrate pushing the slider to the right as there are surface irregularities pushing it to the left. Hence, statistically speaking, there is the possibility of an almost perfect annihilation of lateral forces. Allowing for long-range, *elastic* deformations does not alter the almost systematic annihilation of lateral forces, unless the systems are extremely soft or extremely rough. Although this is a subject of current research, it appears that this statement remains valid even if one considers surface profiles that would be characteristic of engineering surfaces such as pistons and cylinders in car engines.

In the following, we will discuss why surfaces are usually not superlubric. In general, this is due to energy dissipation through various mechanisms during sliding. The knowledge of the relevant dissipation mechanisms is important when setting up tribological simulations, as it allows one to estimate the validity of the simulations for given laboratory conditions.

2.1 Friction Mechanisms

It has long been realized that solid friction is intimately connected to hysteresis and to what one may call plugging motion. This is best illustrated in the model proposed independently by Prandtl [6] and by Tomlinson [7]. In their model, a surface atom of mass m is coupled to its lattice site via a harmonic spring of stiffness k. The lattice site, which moves at constant velocity v_0, is assumed to be located in the origin at time $t = 0$. Besides the interaction with its lattice site, the atom experiences a coupling $V_0 \cos(2\pi x/a)$ to the substrate, where V_0 has the unit of energy and reflects the strength of the coupling, a is the lattice constant of the substrate, and x is the current position of the surface atom. Introducing a viscous damping proportional to velocity $\dot{x}$ and damping coefficient γ,[2] the surface atom's equation of motion reads

[2] The damping term can be motivated by the following microscopic picture. When the atom moves very slowly, some lattice vibrations will be induced in the substrate so that energy – or heat – flows away from the interface. The damping term, which may be position dependent, can be calculated in principle from generally

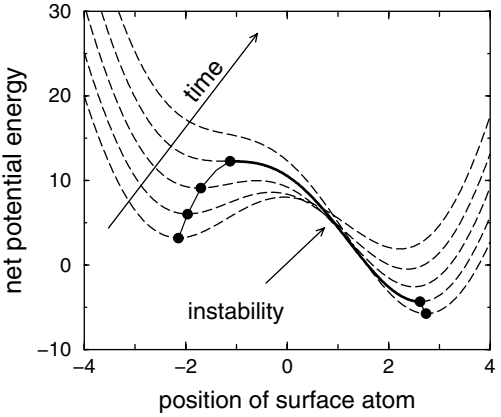


Fig. 1. Illustration of an instability in the Prandtl-Tomlinson model. The sum of the substrate potential and the elastic energy of the spring is shown at various instances in time. The energy difference between the initial and the final point of the thick line will be the dissipated energy when temperature and sliding velocities are very small

$$m\ddot{x} + \gamma\dot{x} = k(v_0 t - x) + \frac{2\pi}{a} V_0 \sin(2\pi x/a) \; . \tag{1}$$

If k is very large, i.e., k is greater than the maximum curvature of the potential, $V''_{\max} = (2\pi/a)^2 V_0$, then there is always a unique equilibrium position $x_{\rm eq} \approx v_0 t$ for the atom and the atom will always be close to $x_{\rm eq}$. Consequently, the friction will be linear in v_0 at small values of v_0.

Things become more interesting once $V''_{\max}$ exceeds k. Now there can and will be more than one stable position at certain instances of time as one can see in Fig. 1. The time dependence of the combined substrate and spring potential reveals that mechanically stable positions disappear at certain instances in time due to the motion of the spring. Consequently, an atom cannot find a mechanically stable position at a time $t + \delta t$ in the vicinity of a position that was stable a small moment δt ago. At times slightly larger than t, the position of the atom becomes *unstable* and, hence, it must move forward quickly towards the next potential energy minimum. After sufficiently many oscillations around the new mechanical equilibrium, most of the potential energy difference between the new and the old equilibrium position is dissipated into the damping term. As a consequence, for sufficiently small v_0, the dissipated energy per sliding distance is rather independent of v_0 and (in the present example of a bistable system) similarly independent of γ.

Despite its merits, which are further discussed in Sect. 2.5, one should not take the Prandtl-Tomlinson model too literally. There simply is no reason why the inter-bulk coupling, reflected by $V''_{\max}$, should be stronger than the

valid statistical mechanics arguments by integrating out the substrate's lattice vibrations. A more detailed justification is beyond the scope of this chapter.

intra-bulk coupling k. But even if it were, one would have to expect more dramatic processes than elastic instabilities, such as cold welding and plastic deformation, so that the assumption of an elastic coupling in the slider would break down completely. One could certainly argue that similar instabilities may occur at larger length scales, involving collective degrees of freedom [8]. However, it appears that elastic instabilities do not contribute considerably to dissipation [9]. A notable exception to this rule is rubber, for which sliding friction is related to internal friction rather than to dissipation taking place at the interface [10].

A traditional explanation of solid friction, mainly employed in engineering sciences, is based on plastic deformation [11]. It is assumed that plastic flow occurs at most microscopic points of contact, so that the normal, local pressures correspond to the hardness, σ_{h}, of the softer of the two opposed materials. The (maximum) shear pressure is given by the yield strength, σ_{y}, of the same material. The net load, L, and the net shear force, F_{s}, follow by integrating σ_{h} and σ_{y} over the *real* area of contact, A_{real}, respectively, i.e., $L = \sigma_{\mathrm{h}} A_{\mathrm{real}}$ and $F_{\mathrm{s}} = \sigma_{\mathrm{y}} A_{\mathrm{real}}$.[3] Hence, the plastic deformation scenario results in the following (static) friction coefficient

$$\mu_{\mathrm{s}} = \sigma_{\mathrm{y}} / \sigma_{\mathrm{h}} \ , \tag{2}$$

where μ_{s} is defined as the ratio of F_{s} and L. Although this explanation for a linear relationship between friction and load has been used extensively in the literature, Bowden and Tabor, who suggested this idea, were aware of the limitations of their model and only meant to apply it to contacts between (bare) metals [11]. There are two important objections to the claim that plastic deformation is generally a dominant friction mechanism. Usually, friction between two solids does not (only) depend on the mechanical properties of the softer of the two opposing materials, but on both materials and the lubricant in the contact. Moreover, theoretical calculations of typical surface profiles have shown that plastic flow should occur at only a very small fraction of the the total number of contact points [12].

So far, we have not considered lubricants added intentionally, such as oils, or unintentionally in the form of contaminants, such as short, airborne hydrocarbons. However, such adsorbed molecules alter dramatically the behaviour of sliding contacts, as long as they do not become squeezed out the microscopic points of contact [13, 14]. From an engineering point of view, such molecules keep the two opposed surfaces from making intimate mechanical contact, thereby reducing plastic deformation and wear. However, they also keep surfaces from becoming superlubric. The last one or two layers of lubricant do not become squeezed out of the contact and solidify due to the typically large pressures at the microscopic scale. In this regime, one gener-

[3] Intimate mechanical contact between macroscopic solids occurs at isolated points only, typically at a small fraction of the apparent area of contact. The net area of this intimate contact is called the real area of contact A_{real}.

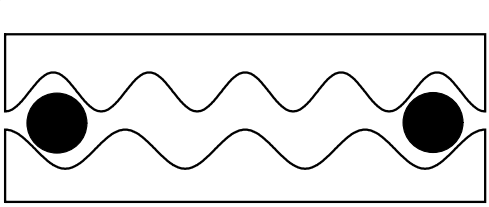


Fig. 2. Schematic representation of the way how adsorbed atoms can lock two non-matching solids. From [14]

ally talks about boundary lubrication. As the interactions between lubricant particles is relatively weak, the adsorbed atoms and molecules will predominantly try to satisfy the interactions with the confining walls. This can lock the surfaces geometrically, as illustrated schematically in Fig. 2. When sliding the top wall relative to the substrate, an energy barrier has to be overcome, generating a static friction force.

There are many other mechanisms leading to dissipation, although they may be less universal than those related to boundary lubricant-induced, geometric frustration. Chemical changes in lubricant molecules, reversible or irreversible, produce heat. Examples are configurational changes in hydrogen-terminated diamond surfaces [16] or terminal groups of alkane chains through isomerization [15] and sliding- and pressure-induced changes in the coordination numbers of surface or lubricant atoms [17, 18]. Although the microscopic details differ significantly, all these examples exhibit a molecular hysteresis similar to the one described in the context of the Prandtl-Tomlinson model. There are also many strongly irreversible tribological phenomena, such as cold-welding, scraping, cutting, or uncontrolled, catastrophic wear. Characterizing them is often tedious, because many of these strongly irreversible processes are system specific and lack a steady state. For these reasons, we will focus mainly on non-equilibrium, steady-state type situations.

It is yet important to realize that simulating strongly irreversible processes often requires more care than those with a well-defined steady state. Most experimental systems are open, while simulations employ confining walls and periodic boundary conditions parallel to the interface. Any debris generated will remain in the interface in the simulations, unless special precaution is taken. One means of simulating open systems is to incorporate lubricant reservoirs, however, this leads to a significant increase in computational effort.

2.2 Velocity-Dependence of Friction

Solid friction is typically relatively independent of the sliding velocity v_0. This finding, also known as Coulomb's law of friction, [19] can be rationalized nicely in the Prandtl-Tomlinson model. A given number of instabilities occurs per sliding distance, Δx. Each instability will produce a similar amount of heat, ΔQ. In steady state, one may therefore associate the kinetic friction force, F_{k}, with the quotient

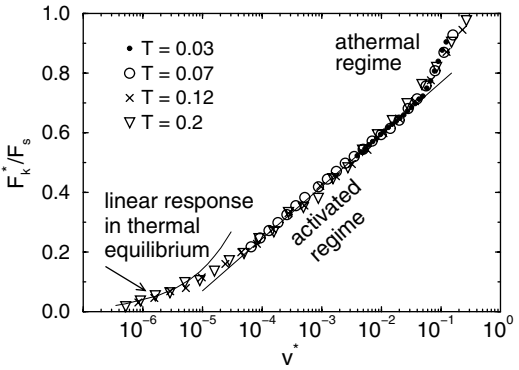


Fig. 3. Typical velocity relationship of kinetic friction for a sliding contact in which friction is due to adsorbed layers confined between two incommensurate walls. The kinetic friction F_k is normalized by the static friction F_s. At extremely small velocities v^*, the confined layer is close to thermal equilibrium and consequently F_k is linear in v^*, as to be expected from linear response theory. In an intermediate velocity regime, the velocity dependence of F_k is logarithmic. Instabilities or "pops" of the atoms can be thermally activated. At large velocities, the surface moves too quickly for thermal effects to play a role. Time-temperature superposition could be applied. All data were scaled to one reference temperature. From [20]

$$F_\mathrm{k} = \frac{\Delta Q}{\Delta x} \,. \tag{3}$$

Once temperature comes into play, the jumps of atoms may be invoked prematurely via thermal fluctuations. Consequently, the spring pulling the surface atom will be less stretched on average. This will decrease the average friction force and render F_k rate or velocity dependent, typically in the following form

$$F_\mathrm{k} \approx F_\mathrm{k}(v_\mathrm{ref}) + c\left(\ln\frac{v_0}{v_\mathrm{ref}}\right)^\gamma \,, \tag{4}$$

where c is a constant, v_ref a suitable reference velocity, and γ an exponent in the order of unity. Of course, this equation will only be valid over a limited velocity range. In many cases, F_k becomes linear in v_0 at very small values of v_0, i.e., when one enters the linear response regime, in which the system is always close to thermal equilibrium.

An example for the velocity dependence of friction is given in Fig. 3 for a boundary lubricant confined between two incommensurate surfaces [20]. For the given choice of normal pressure and temperature, one finds four decades in sliding velocity, for which (4) provides a reasonably accurate description. In the present case, c is positive and the exponent γ is unity. Neither of the two statements is universal. For example, the Prandtl-Tomlinson model can best be described with $\gamma = 2/3$ in certain regimes, [21,22] while confined boundary lubricants are best fit with $\gamma = 1$ [20,23]. Moreover, the constant c can become negative, in particular when junction growth is important. The local contact

areas can grow with time due to slow plastic flow of the opposed solids or due to adhesive interactions mediated by water capillaries, which increase in time [8, 24, 25].

2.3 Load-Dependence of Friction and Contact Mechanics

Many macroscopic systems show an almost linear relationship between (static) friction, F_s, and load, L,

$$F_\mathrm{s} = \mu_\mathrm{s} L \ , \tag{5}$$

where the (static) friction coefficient, μ_s, does not depend on the apparent area of contact. The origin of this linear dependence, which is also called Amontons' law, is subject to controversy.

One explanation of Amontons' law is based on microscopic arguments. It had been argued by Bowden and Tabor [11] that the following constitutive relation for shear stress σ_s and normal stress σ_n holds microscopically

$$\sigma_\mathrm{s} = \sigma_0 + \alpha \sigma_\mathrm{n} \tag{6}$$

for many systems, where σ_0 and α are constant. From this one obtains $F_\mathrm{s} = \sigma_0 A_\mathrm{real} + \alpha L$, so that one may associate α as the (differential) friction coefficient provided that σ_0 is sufficiently small. Simulations of boundary lubricants for systems with flat surfaces suggest that (6) may often be reasonably accurate up to pressures close to the yield strength of solids and that the term related to σ_0 are indeed often negligible [13, 26].

The reason for the linearity of shear and normal pressures can be rationalized qualitatively by considering Fig. 2. In order for the top wall to move to the right it must move up a slope, which is dependent upon how the adsorbed atom is interlocked between the substrate and the slider.

Of course, this argument is highly qualitative, because it assumes implicily that non-bonded atoms behave similar to hard disks in areas of high pressure. Moreover, this argument must be modified if curved surfaces are considered [27]. However, it appears to be a reasonable approximation for many systems.

Another scenario leading to Amontons' law is related to the macroscopic contact mechanics. Even highly polished surfaces are rough on many different length scales. A way of characterizing roughness is to average or measure the height difference auto-correlation function $C_2(\Delta r)$

$$C_2(\Delta r) = \left\langle [h(\mathbf{r}) - h(\mathbf{r} + \Delta \mathbf{r})]^2 \right\rangle \ , \tag{7}$$

over one or several statistically identical samples, where $h(\mathbf{r})$ is the height of a sample's surface at the position $\mathbf{r} = (x, y)$. Thus, $C_2(\Delta r)$ states what variation in height we expect if we move away a distance Δr from our current position. For many real surfaces, power law behaviour according to

$$C_2(\Delta r) \propto \Delta r^{2H} \ , \tag{8}$$

is found, where H is called the Hurst roughness exponent. $H = 1/2$ would correspond to a random walk in the height as we move laterally over the surface. Surfaces satisfying 8 are called self-similar. Ways of constructing self-similar surfaces for simulations will be described in Sect. 3.2.

When two macroscopic solids with fractal surfaces are brought into contact, only a small fraction of the surfaces, the so-called real area of contact, A_{real}, will be in microscopic, mechanical contact. It can be shown that the pressure distribution averaged over these real contacts is surprisingly independent of the externally imposed load, L, provided that the surfaces are not too adhesive or too compliant [12, 28, 29]. This implies that Amontons' law can also result from macroscopic contact mechanics irrespective of the local relation between normal and shear pressure. However, when conditions are less ideal and adhesion and plastic deformation are starting to play a role, the independence of the pressure distribution on the net load is not valid any longer [30]. Hence, different scenarios can lead to the observation of Amontons' law depending on the details of the system of interest.

2.4 Role of Interfacial Symmetry

Imagine two egg cartons placed on top of each other. If you try to move the top carton by applying a lateral force, you will have to pull harder if the cartons are oriented than if they are brought out of registry. If the two cartons are separated by eggs, there still will be a tremendous influence of the orientation on the required lateral force to slide the cartons. The same notion holds for surfaces that are separated by confined atoms and molecules. Commensurate surfaces, i.e., those that are identical and perfectly aligned, will have the tendency to have much larger (static) friction than those that are misoriented or dislike. Therefore, whenever we impose symmetries into our systems, we risk observing behaviour that is inconsistent with observed when these symmetries are absent. Since opposing surfaces are essentially always dislike, unless they are prepared specifically, it will be important to avoid symmetries in simulations as much as possible.

It may come as a surprise to some that two commensurate surfaces can withstand finite shear strength even if they are separated by a fluid [31]. But one has to keep in mind that breaking translational invariance automatically induces a potential of mean force $\mathcal{F}$. This is why metals, which are mainly glued together through "fluid-like" conduction electrons, have finite shear moduli. Due to the symmetry breaking, commensurate walls can be pinned even by an ideal gas embedded between them [32]. The reason is that $\mathcal{F}$ scales linearly with the area of contact. In the thermodynamic limit the energy barrier for the slider to move by one lattice constant becomes infinitely high so that the motion cannot be thermally activated and, hence, static friction becomes finite. No such argument applies when the surfaces do not share a common period.

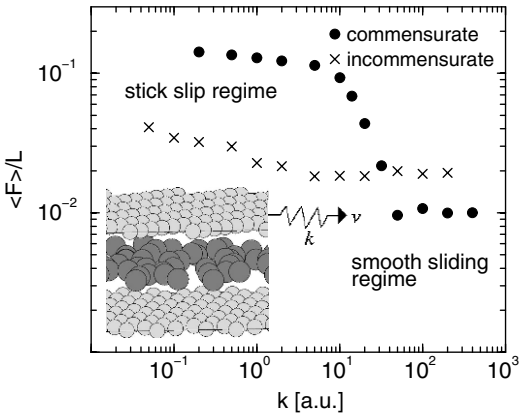


Fig. 4. Kinetic friction F_k as a function of the stiffness k of the spring pulling the upper wall at constant, small velocity. The inset shows a part of the simulated system. At large values of k, the slider moves at the same velocity as the spring and the smooth sliding kinetic friction is probed. At small values of k, the system manages to lock into a potential energy minimum, similar to what happens in the Prandtl-Tomlinson model. The surface then undergoes plugging or "stick-slip" motion as a whole. In that regime, the measured friction approaches the value for static friction. Commensurability affects the measured values for F_k in both regimes sensitively. From [20]

Not only static friction, F_s, but also kinetic friction, F_k, is affected by commensurability. If two crystalline surfaces are separated by one atomic layer only, F_k may actually be reduced due to commensurability, although static friction is increased [20]. The strikingly different behaviour for commensurate and incommensurate systems is demonstrated in Fig. 4.

Unfortunately, it can be difficult to make two surfaces incommensurate in simulations; particularly when two identical, crystalline surfaces are slid against each other. The reason is that only a limited number of geometries conform to the periodic boundary conditions in the lateral direction. Each geometry needs to be analyzed separately and there is little general guidance one can give. For surfaces with trigonal symmetry, such as [111] surfaces of face-centered cubic crystals, it is often convenient to rotate the top wall by 90°. This rotation does not map the trigonal lattice onto itself. The numbers of unit cells in x and y direction should be chosen such that they need to be strained only marginally to form an interface with a square geometry. The top view of some incommensurate structures between trigonal surfaces is shown in Fig. 5. In most cases, the measured friction between incommensurate walls is relatively insensitive to how incommensurability is achieved, as long as the roughness of the two opposing walls remains constant [14].

Typical surfaces are usually not crystalline but have amorphous layers on top. These amorphous walls are much rougher at the atomic scale than

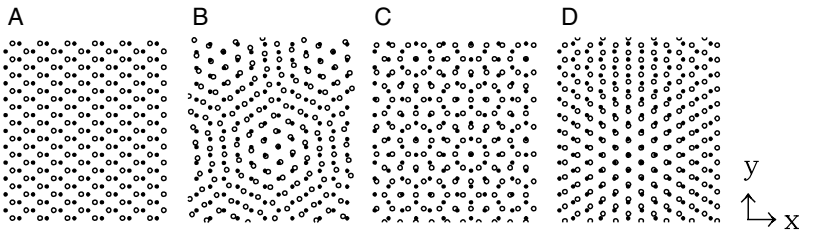


Fig. 5. Projections of atoms from the bottom (*solid circles*) and top (*open circles*) surfaces into the plane of the walls. (A through C) The two walls have the same structure and lattice constant, but the top wall has been rotated by 0°, 11.6°, or 90°, respectively. (D) The walls are aligned, but the lattice constant of the top wall has been reduced by 12/13. The atoms can only achieve perfect registry in the commensurate case (A). From [13]

the model crystalline surfaces, which one prefers to use for computational convenience and for fundamental research. The additional roughness at the microscopic level due to disorder increases the friction between surfaces considerably even when they are separated by a boundary lubricant [14]. However, no studies have been done to explore the effect of roughness on boundary-lubricated systems systematically and only a few attempts have been made to investigate dissipation mechanisms in the amorphous layers under sliding conditions from an atomistic point of view.

2.5 Common Toy Models

The Prandtl-Tomlinson model introduced in Sect. 2.1 is the most commonly used toy model for simulations of frictional phenomena. It has a tremendous didactic value, because it shows nicely the important role of instabilities. Moreover, it is used frequently to describe quite accurately the dynamics of an atomic force microscope tip that is dragged over a periodic substrate [33]. It is worthwhile to run a few simulations of the Prandtl-Tomlinson model and to explore its rich behaviour. In particular, interesting dynamics occurs when k is sufficiently small such that the surface atom is not only bistable but multistable and γ is so small that the motion is underdamped. In such a situation, the atoms do not necessarily become arrested in the next available, mechanically stable site after after depinning and interesting non-linear dynamics can occur, such as non-monotonic friction-velocity dependence.

Another frequently-used model system is the Frenkel Kontorova (FK) model, in which a linear harmonic chain is embedded in an external potential. For a review we refer to [34]. The potential energy in the FK model reads

$$V = \sum_i \left\{ \frac{1}{2} k (x_{i+1} - b - x_i)^2 - V_0 \cos(2\pi x_i / a) \right\} , \tag{9}$$

where a and b are the lattice constant of substrate and slider respectively, and k is the strength of the spring connecting two adjacent atoms in the slider. As in the Prandtl Tomlinson model, finite friction is found when atoms can find more than one mechanically meta-stable position and become unstable during sliding. Experience indicates that it is not possible to reproduce tribological experimental results with the FK model despite the increase in complexity with respect to the model of Prandtl and Tomlinson. This calls into question the use of the FK model when interpreting experimental results. In particular, when parametrized realistically and generalized to higher dimensions, it is found that most incommensurate interfaces between crystals should be superlubric within the approximations of the FK model [5]. Otherwise, when instabilities do occur, the FK model can only describe the early time behaviour of flat sliding interfaces [35]. Conversely, in other contexts, such as the motion of charge density waves, pinning and dissipation may be realistically described by the FK model.

3 Computational Aspects

Simulating solids in relative motion may require considerations additional to those needed for equilibrium simulations of bulk phases. Shear and load need to be imposed in a way that mimics experimental setups. Surfaces have to be defined and often it is important to include their deformation in the simulation accurately, which we want to do at a small computational expense. Heat needs to be removed, which requires us to know the properties of thermostats. This chapter will be concerned with these and related aspects.

3.1 Imposing Load and Shear

In many simulations of tribological phenomena, two opposed solids are separated by a lubricating film. A sketch is shown in Fig. 6. It is natural to subdivide the system into a substrate, a slider, and the remaining system. Sometimes, one may only be interested in the bulk properties of a lubricant under shear or under extreme pressure conditions, in which case, there is no need to introduce surfaces, see the discussion in Sect. 3.4. Otherwise, when walls are included, it is certainly desirable to keep the interface as unperturbed as possible from any external mechanical forces. This is why it is good practice to only couple the outermost layers of the substrate and slider to constraints, external forces, and thermostats. In the following, the term bottom layer will be used to specify the outermost layer of the substrate and top layer will stand for the outermost layer of the slider, although one may certainly choose more than one single layer to be part of it. All explicitly simulated atoms that are not part of the outermost layers will be referred to as the embedded system, even if they belong to the confining walls.

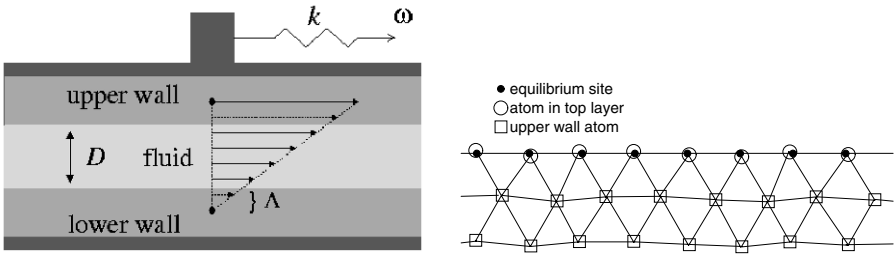


Fig. 6. *Left*: Schematic graph of the set-up for the simulation of rubbing surfaces. Upper and lower walls are separated by a fluid or a boundary lubricant of thickness D. The outermost layers of the walls, represented by dark color, are often treated as rigid unit. The bottom most layer is fixed in laboratory system and the upper most layer is driven externally, for instance by a spring of stiffness k. Also shown is a typical, linear velocity profile for a confined fluid with finite velocities at the boundary. The length at which the fluid's drift velocity would extrapolate to the wall's velocity is called the slip length Λ. *Right*: The top wall atoms in the rigid top layer are set onto their equilibrium sites or coupled elastically to them. The remaining top wall atoms interact through interatomic potentials, which certainly may be chosen to be elastic

By convention, we will keep the center of mass of the bottom layer fixed and couple the top layer to an external driving device. There are three commonly used modes under which top layers are driven:

1. Predefined trajectory, e.g., $X = X(t)$
2. Predefined force, e.g., $F = F(t)$
3. Pulling with a spring, e.g., $F_x = -k[X - X_0(t)]$, where F_x would be the force acting on the top layer in x direction, k would reflect the (effective) stiffness of the driving device, and $X_0(t)$ denotes the position of the driving device as a function of time.

The typical choices for the predefined trajectories or forces are constant velocity, including zero velocity, constant separation, and constant forces and/or oscillatory velocities and forces. It is certainly possible to drive different Cartesian coordinates with different modes, e.g., to employ a constant force or load perpendicular to the interface and to use a predefined velocity, constant or oscillatory, parallel to a direction that has no component normal to the interface. Mimicking experiments done with a tribometer would typically best be done in constant velocity or constant force modes, whereas rheometers usually employ oscillatory motion in lateral direction.

Note that pulling a point particle over a periodic potential in mode (3) resembles the Prandtl-Tomlinson model discussed in the previous section. As is the case for the Prandtl Tomlinson model, the result for the (kinetic) friction force can depend sensitively on the stiffness of the driving spring, see also Fig. 4. Weak springs tend to produce higher friction than soft springs. This can be important to keep in mind when comparing simulations to experiments.

The measured friction is not only a function of the interface but also of how the interface is driven.

It is often beneficial to define a coordinate $\mathbf{R}_{\rm tl}$ that describes the center of mass of the top layer. There are three common ways how to set up the top layer. (i) To confine the position of top layer atoms $\mathbf{r}_n$ to (lattice) sites $\mathbf{r}_{n,0}$, which are connected rigidly to the top layer. (ii) To couple the top layer atoms elastically to sites $\mathbf{r}_{n,0}$ fixed relative to the top layer, e.g., with springs of stiffness k. (iii) To employ an effective potential, such as a Steele potential, $V_{\rm S}$, [36] between embedded (em) atoms and top layer. There are specific advantages and disadvantages associated with each method. Approach (i) may be the one that is most easily coded, (ii) allows one to thermostat effectively the outermost layer, while (iii) is probably cheapest in terms of CPU time.

Depending on the choice of the interaction between the top layer and the embedded system, the force on the top wall, $\mathbf{F}_{\rm tl}$, needs to be evaluated differently.

$$\mathbf{F}_{\rm tl} = \mathbf{F}_{\rm ext} + \begin{cases} \sum_{n\in{\rm tl}} \mathbf{f}_{\rm n} & ({\rm i}) \\ \sum_{n\in{\rm tl}} -k\,(\mathbf{r}_n - \mathbf{r}_{n,0}) & ({\rm ii}) \\ \sum_{n\in{\rm em}} -\nabla_n V_{\rm S}(\mathbf{r}_n) & ({\rm iii}) \end{cases}, \tag{10}$$

where $\mathbf{f}_{\rm n}$ in line (i) denotes the force on atom n and $\nabla_n V_{\rm S}$ is the gradient of the surface potential with respect to an embedded atom's position. $\mathbf{F}_{\rm tl}$ will be used to calculate the acceleration of the top layer, resulting in a displacement $\Delta\mathbf{R}_{\rm tl}$. This displacement needs to be added to the sites $\mathbf{r}_{n,0}$ contained in the top layer in cases (i) and (ii).

It is certainly possible to choose the mass $M_{\rm tl}$ of the top layer arbitrarily. For example, one may have $M_{\rm tl}$ incorporate some mass of the top wall that is not explicitly included in the simulation. However, when doing this, one needs to be aware of two effects. First, the time scale gap between the fast atomic motion and the slow collective motion of the confining wall. However, having the top wall move on shorter time scales than in real systems may help to overcome the time scale gap between simulations and experiments. Second, the measured friction may depend on the mass of the top wall when it is pulled with a spring. Large masses favor smooth sliding over stick-slip motion and hence reduce the measured friction [37,38]. In general, it turns out to be very difficult to deduce information on the embedded system from tribological or rheological experiments without considering carefully the properties and the driving of the external system [39].

3.2 Including Surface Roughness and Elastic Deformations

Surface Geometries

Crystalline surfaces are often employed both experimentally and in simulations, when studying friction from a fundamental point of view. It is a relatively straightforward procedure to set up crystalline surfaces, which is why

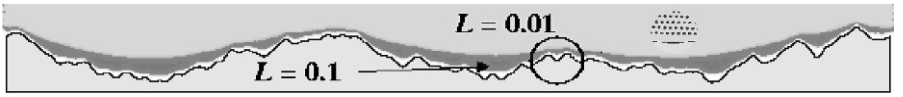


Fig. 7. Flat elastic manifold pressed against a self-affine rigid surface for different loads L per atom in top wall. The inset shows some atomic-scale details of the contact

we will not comment on them in any more detail. However, as mentioned in Sect. 2.3, many (engineering) surfaces are self similar over several length scales. A profile of a self-similar surface geometry is shown in Fig. 7 together with a flat elastic object pressed onto the rough substrate. In the recent past, there has been an intensified interest in modeling more realistically surface profiles, however, so far the research has focused on contact mechanics rather than on sliding motion between fractal surfaces [12, 28, 29].

There are various ways of constructing self-affine surfaces [40]. Some of them do not allow one to produce different realizations of surface profiles, for example by making use of the Weierstrass function. Such methods should be avoided in the present context, because it would be hard to make statistically meaningful statements without averaging over a set of statistically independent realizations. An appropriate method through which to construct self-similar surfaces is to use a representation of the height profile $h(x)$ via its Fourier transforms $\tilde{h}(q)$.

In reciprocal space, the self-affine surfaces described in (7) and (8) are typically characterized by the spectrum $\tilde{S}(q)$ defined as

$$\tilde{S}(q) = \langle \tilde{h}(q)\tilde{h}^*(q) \rangle \,, \tag{11}$$

with

$$\begin{aligned} \langle \tilde{h}(q) \rangle &= 0 \\ \langle \tilde{h}(q)\tilde{h}^*(q') \rangle &\propto q^{-2H-d}\delta(q-q') \,, \end{aligned} \tag{12}$$

where d is the number of independent coordinates on which the height depends, i.e., $d = 1$ if $h = h(x)$, and $d = 2$ if $h = h(x, y)$. $\tilde{S}(q)$ is the Fourier transform of the height autocorrelation function $S(\Delta x) = \langle h(x)h(x + \Delta x) \rangle$. The height-difference autocorrelation function $C_2(\Delta x)$ and $S(\Delta x)$ are related through $2S(\Delta x) = C_2(0) - C_2(\Delta x)$.

In principle, to fully characterize the stochastic properties of surfaces, higher-order correlations of the height function have to be incorporated. However, doing this is tedious computationally and, in most cases, will probably not change significantly the results of the simulation.

One approach to the generation of height profiles is to draw (Gaussian) random numbers for the real and complex parts of $\tilde{h}(q)$ with a mean zero and defined variance and to divide the random number by a term proportional to $q^{H+d/2}$, so that (12) is satisfied. Furthermore, $\tilde{h}(-q)$ must be chosen to be the complex conjugate to ensure that $h(x)$ is a real-valued function.

Alternatively, one may simply write $h(x)$ as a sum over terms $h(q)\cos(qx+\varphi_q)$. In this case, one needs to draw one (Gaussian) random number with the proper second moment of $h(q)$ with zero mean and one random number for each phase φ_q, which is uniformly distributed between 0 and 2π and filter the absolute value of $h(x)$ in the same way as described in the previous paragraph. There are also other techniques which allow one to generate fractal surfaces. One of them is the so-called midpoint algorithm, described in [30].

Multiscale Approaches

It is certainly desirable to simulate as many layers of the confining walls as possible, in order to closely reproduce experimental situations. However, from a computational point of view, one would like to simulate as few degrees of freedom as possible. Unless conditions are special, all processes far away from the interface can be described quite accurately within elastic theory or other methods that allow for a description of plastic deformations, such as finite elements. The advantage of these continuum-theory based methods is that it is possible to coarse-grain the system increasingly as one moves away from the interface, thereby reducing the computational effort. New methodological developments even allow one to couple atomistic simulations to continuum-theory descriptions [41, 42]. It would be out of the scope of this chapter to provide a detailed description, however, the coordination discretisation scheme shown in Fig. 8 alludes to how one must proceed when incorporating different mesh sizes into a simulation.

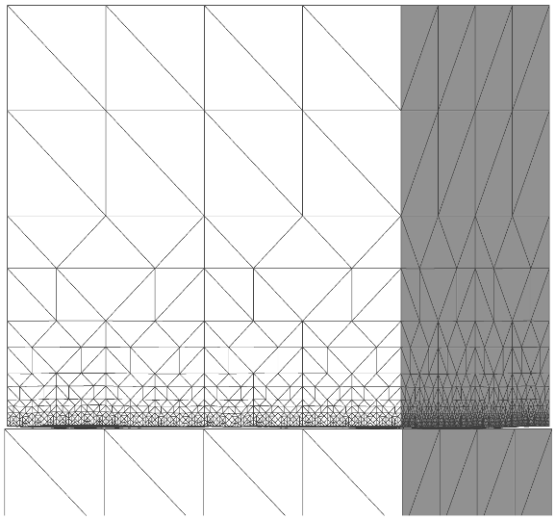


Fig. 8. Representation of a finite-element mesh for the simulation between a fractal, elastic object and a flat substrate. From [12]

While quasi-static processes can be modeled quite well with continuum-mechanics based models employing varying mesh size, this is not the case for dynamic processes. Whenever there is a region where the coarse-grain level is changed, one risks to introduce artificial dynamics. In particular, the transmission of sound waves and energy density is suppressed whenever the mesh size changes. It is not possible to have the proper momentum and energy transfer across the boundary when employing a Hamiltonian-based description. It is important, however, to realize that the computational effort required to simulate a three-dimensional system of linear dimension L only scales with $L^2 \ln L$ using coarse-grained models, as opposed to the L^3 scaling for brute-force methods. It is often advisable to sacrifice realistic dynamics rather than system size.

An alternative method would be to integrate out the elastic degrees of freedom located above the layer that we chose to be the top layer in our simulation [43]. The elimination of the degrees of freedom can be done within the context of Kubo theory, or more precise Zwanzig formalism, leading to effective (potentially time-dependent) interactions between the atoms in the top layer [44, 45]. These effective interactions include those mediated by the degrees of freedom, which have been integrated out. For periodic solids, a description in reciprocal space decouples different wave vectors, q, at least as far as the *static* properties are concerned. This in turn implies that the computational effort also remains in the order of $L^2 \ln L$, provided that use is made of the fast Fourier transform for the transformation between real and reciprocal space. The description is exact for purely harmonic solids, so that one can mimic the static contact mechanics between a purely elastic lattice and a substrate with one single layer only.

There is even the possibility of including *dynamical* effects in terms of time-dependent friction terms (plus random forces at finite temperatures) [44, 45]. However, it may not be advisable to take advantage of this possibility, as the simulation would become increasingly slow with increasing number of time steps. Moreover, the simulation will slow down considerably in higher dimensions due to the non-orthogonality of the dynamical coupling in reciprocal space.

To be specific regarding the formalism, let $\tilde{u}_{qi\alpha}(t)$ denote the α component of the displacement field associated with wave vector q and eigenmode i at time t.[4] In the absence of external forces, which can simply be added to the equation, the equation of motion for the coordinates that are not thermostatted explicitly, $\tilde{u}_{qi\alpha}$, would read:

$$M\ddot{\tilde{u}}_{qi\alpha}(t) = -G_{qi}u_{qi\alpha}(t) + \int_{-\infty}^{t} dt' \sum_{q',j} \sum_{\beta=1}^{3} \gamma_{q\,i\alpha}^{q'\,j\beta}(t-t')\dot{u}_{q'\,j\beta}(t') + \Gamma_{q\,i\alpha}(t) \ , \quad (13)$$

[4] At this point, all interactions within the top plate and above it have already be integrated out. The $\tilde{u}_{qi\alpha}$ are only eigenmodes within this reduced or effective description but not of the full semi-infinite solid.

where the G_{qi} are the (static) Green's functions, or effective spring constants, associated with eigenmode i and wavelength q. The knowledge of these functions enables us to work out the static contact mechanics. The time-dependent damping coefficients $\gamma_{q\,i\alpha}^{q'\,j\beta}(t-t')$ in (13) reflect the dynamical coupling between various eigenmodes. There is no reason why this coupling should be diagonal in any of its indices and thus, including the terms related to dynamics, increases memory requirements and slows down the speed of the calculation tremendously, i.e., beyond the expense of approximating a semi-infinite solid by a discrete, elastic lattice of size L^3. Included in (13) are random forces $\varGamma_{q\,i\alpha}(t)$, which must be used at finite temperature to counterbalance the time-dependent damping term. The random and damping terms have to be chosen such that they satisfy the fluctuation-dissipation theorem [46].

3.3 Imposing Constant Temperature

The external driving imposed on solids leads to dissipation of energy or heat. In experiments, this heat diffuses away from the interface into the bulk and eventually into the experimental apparatus. In simulations, system sizes are rather limited which makes it necessary to remove heat artificially if one wants to control temperature. Ideally, this is done by thermostating the outermost layers only. Sometimes, however, there are no confining walls, for example in simulations of bulks fluids, or for some reason, the confining walls are better kept rigid. In these cases, the thermostat needs to be applied to the sheared system directly.

There are numerous ways of thermostatting systems that are in equilibrium, each with its specific advantages and disadvantages. For situations in which the system is far from equilibrium, stochastic thermostats have proven particularly beneficial, Langevin thermostats being the prototype, [47] and dissipative particle dynamics (DPD) being a modern variation thereof [48]. While stochastic thermostats can be motivated in principle from linear response theory, i.e., there are rigorous schemes for the derivation of the damping terms and the fluctuation terms contained in stochastic thermostats [46]. We will not provide these arguments here and instead focus on their implementation and properties.

Langevin Thermostat

In the Langevin description, one assumes that the degrees of freedom that are not explicitly taken into account, exert, on average, a damping force linear in velocity, $\gamma_i \dot{\mathbf{r}}_i$, as well as additional random forces $\varGamma_i(t)$. This leads to the following equation of motion for particle number i:

$$m_i \ddot{\mathbf{r}}_i + \gamma_i(\dot{\mathbf{r}}_i - \langle \mathbf{v}_i \rangle) = -\nabla_i V + \varGamma_i(t)\ , \tag{14}$$

where the damping coefficient γ_i and the α component of the random forces $\varGamma_{i\alpha}(t)$ acting on particle i should obey

$$\begin{aligned} \langle \Gamma_{i\alpha}(t) \rangle &= 0 \\ \langle \Gamma_{i\alpha}(t) \Gamma_{j\beta}(t') \rangle &= 2\gamma_i k_B T \, \delta(t - t') \, \delta_{ij} \, \delta_{\alpha\beta} \\ &\to 2\gamma_i k_B T \, \frac{1}{\Delta t} \, \delta_{t,t'} \, \delta_{ij} \, \delta_{\alpha\beta} \, , \end{aligned} \tag{15}$$

in order to satisfy the fluctuation-dissipation theorem. In (14), V denotes the total potential energy and $\langle \mathbf{v}_i \rangle$ the expected drift velocity, [49] e.g., $\langle \mathbf{v}_i \rangle = 0$ in the bottom layer and $\langle \mathbf{v}_i \rangle = \mathbf{v}_{\mathrm{tl}}$ if atom i belongs to the to top layer. The last line in (15) refers to the discrete time description used in molecular dynamics in which Δt is the time step. When using predictor-corrector methods (velocity Verlet is a second-order Gear predictor corrector method), it is necessary to keep in mind that random terms cannot be predicted. Therefore, one should only apply the predictor-corrector schemes to the deterministic parts of the equation of motion. In those cases, where very high damping is employed, time steps can be kept large when employing efficient integration schemes [52]. In general, however, one should keep thermostating sufficiently weak as to avoid externally imposed overdamping.

It also needs to be emphasized that there is no need to chose the random forces from a Gaussian distribution, unless one is interested in short-time dynamics. It is much faster to generate uniformly distributed random numbers for the $\Gamma_i(t)$'s on an interval $[-\sqrt{3}\sigma, \sqrt{3}\sigma]$, where σ is the standard deviation of the Gaussian distribution. Moreover, having a strict upper bound in the $\Gamma_i(t)$'s eliminates potentially bad surprises when using higher-order predictor corrector schemes and, thus, allows one to use a large time step while producing accurate thermal averages and trajectories.

It is certainly also possible to make damping and, hence, thermostatting direction dependent, for example, by suppressing the damping terms parallel to the sliding direction. This is particularly important when the system has a small viscosity or when the shear rates are high, because one is likely to create artificial dynamics otherwise. Using the correct velocity profile $\langle \mathbf{v} \rangle$ prior to the simulation can also reduce the problem of perturbing the dynamics in an undesirably strong fashion. However, anticipating certain velocity profiles will always suppress other modes, e.g., assuming laminar flow in a thermostat is likely to artificially bias towards laminar flow [53] and may create additional artifacts [54–56].

Effects of Damping on Calculated Friction

Making assumptions on how heat is dissipated can also influence solid friction, although typically it is less of an issue. This can be most easily explored within the Prandtl Tomlinson model. The lessons to be learned apply to a large degree to more general circumstances. In the original formulation, see (1), damping takes place relative to the substrate. However, one may also assume that the conversion of energy into heat takes place within the top solid [57]. Thus a generalized Prandtl-Tomlinson model would be

$$m\ddot{x} + \gamma_{\rm sub}\dot{x} + \gamma_{\rm top}(\dot{x} - v_0) = -\nabla V - k(x - v_0 t) + \Gamma_{\rm sub}(t) + \Gamma_{\rm top}(t)\ , \quad (16)$$

where the indices "sub" and "top" denote the thermal coupling to substrate and top solid respectively.

To investigate the way in which the thermostat affects frictional forces, it is instructive to study slightly underdamped or slightly overdamped motion. In the following, we will set $m = 1$, $a = 1$, $V_0 = 1$, $k = 0.5\, V''_{\rm max}$. Damping, γ, and temperature, T, will be varied, but we will first consider the athermal case $T = 0$. With this choice of parameters, the maximum curvature of the potential, $V''_{\rm max}$, will be greater than k so that instabilities will occur under sliding leading to finite kinetic friction at small v_0 in the absence of thermal fluctuations. Figure 9(a) shows the friction-velocity dependence for the following choices of thermostats (i) $\gamma_{\rm s} = 1$, $\gamma_{\rm t} = 0$, (ii) $\gamma_{\rm s} = 0$, $\gamma_{\rm t} = 1/4$, (iii) $\gamma_{\rm s} = 0$, $\gamma_{\rm t} = 1$, (iv) $\gamma_{\rm s} = 0$, $\gamma_{\rm t} = 4$.

We see that the kinetic friction is rather insensitive to the precise choice of the thermostat for small values of v_0, at least as long as the temperature is sufficiently small. This is because friction is dominated by fast "pops" in that regime. This conclusion becomes invalid only if γ is sufficiently small such that a/v_0, the time it takes the driving stage to move by one lattice constant, is not long enough to transfer most of the "heat" produced during the last instability into the thermostat. At high velocities, the sliding velocity v_0 is no longer negligible when compared to the peak velocity during the instability. This renders the friction-velocity dependence susceptible to the choice of the thermostat. Damping with respect to the substrate leads to strictly monotonically increasing friction forces, while damping with respect to the top wall can result in non-monotonic friction-velocity relationships.

So far, we have considered the zero temperature case. Once finite thermal fluctuations are allowed, there is a qualitatively different friction-velocity relationship, which can be shown by choosing thermal energies as small as $k_B T = 0.1 V_0$, see Fig. 9(b). Jumps can now be thermally activated and the friction force decreases with decreasing velocity. Yet again, at small v_0 there is little effect of the thermostat on the measured friction forces. Changing $\gamma_{\rm t}$ by as much as a factor of 16 results in an almost undetectable effect at small v_0. At very small sliding velocities, v_0, the system can get very close to thermal equilibrium at every position of the top wall, which is why linear response theory is applicable in that regime. It then follows that friction and velocity are linear at sufficiently small values of v_0. This is generally valid unless the energy barriers to sliding are infinitely high, which explains the linear dependence of friction upon the velocity at small v_0 and finite T, as shown in Fig. 3.

Dissipative-Particle-Dynamics Thermostat

A disadvantage of Langevin thermostats is that they require a (local) reference system. Dissipative particle dynamics (DPD) overcomes that problem by

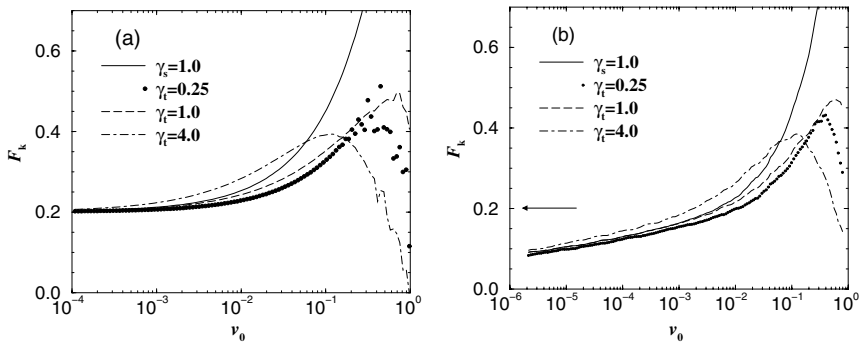


Fig. 9. Friction velocity relationship $F_{\mathrm{k}}(v_0)$ in the Prandtl Tomlinson model at (**a**) zero and (**b**) finite thermal energy, i.e., $k_{\mathrm{B}}T = 0.1V_0$. Different damping with respect to substrate γ_{s} and top solid γ_{t} for different realizations of damping. The arrow in (b) points to the zero-velocity limit in the athermal case

assuming damping and random forces in the center-of-mass system of a pair of atoms. The DPD equations of motion read

$$m\ddot{\mathbf{r}}_i = -\nabla_i V - \sum_j \gamma_{ij}(\dot{\mathbf{r}}_i - \dot{\mathbf{r}}_j) + \Gamma_{ij}(t) \;, \tag{17}$$

where $\Gamma_{ij}(t) = -\Gamma_{ji}(t)$. The usual relations for fluctuation and dissipation apply

$$\langle \Gamma_{ij,\alpha}(t) \rangle = 0 \tag{18}$$

$$\langle \Gamma_{ij,\alpha}(t)\Gamma_{kl,\beta}(t') \rangle = 2k_{\mathrm{B}}T\gamma_{ij}\left(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}\right)\delta_{\alpha\beta}\frac{\delta_{t,t'}}{\Delta t} \tag{19}$$

Note that γ_{ij} can be chosen to be distance dependent. A common method is to assume that γ_{ij} is a constant for a distance smaller than a cut-off radius $r_{\mathrm{cut,\,DPD}}$ and to set $\gamma_{ij} = 0$ otherwise. Since calculating random numbers may become a relatively significant effort in force-field based molecular dynamics, it may be sensible to make $r_{\mathrm{cut,\,DPD}}$ smaller than the cut-off radius for the interaction between the particles, or to have the thermostat act only every few time steps.

Among the advantages of DPD over Langevin dynamics are conservation of momentum and the ability to describe hydrodynamic interactions with longer wavelengths properly, [50, 51] ensuring that "macroscopic" properties are less effected with DPD than with Langevin dynamics. To see this, it is instructive to study the effect that DPD and Langevin have on a one-dimensional, linear harmonic chain with nearest neighbor coupling, which is the simplest model to study long wave length vibrations.

The Lagrange function L of an unthermostated harmonic chain is given by

$$L = \sum_{i=1}^{N} \frac{m}{2}\dot{x}_i^2 - \frac{k}{2}(x_i - x_{i-1} - a)^2 \ , \tag{20}$$

where a is the lattice constant, k the stiffness of the springs. Periodic boundary conditions are employed after a distance Na. The equations of motion at zero temperature with damping are

$$m\ddot{x}_i + \gamma\dot{x}_i = -k(2x_i - x_{i+1} - x_{i-1}) \text{ (Langevin)} \tag{21}$$

$$m\ddot{x}_i + \gamma(2\dot{x}_i - \dot{x}_{i+1} - \dot{x}_{i-1}) = -k(2x_i - x_{i+1} - x_{i-1}) \text{ (DPD)} \tag{22}$$

As usual it is possible to diagonalize these equations of motion by transforming them into reciprocal space. The equations of motion of the Fourier transforms $\tilde{x}(q,\omega)$ then read

$$-m\omega^2\tilde{x} + im\omega\gamma\tilde{x} + 4\sin^2(qa/2)k\tilde{x} = 0 \text{ (Langevin)} \tag{23}$$

$$-m\omega^2\tilde{x} + 4\sin^2(qa/2)im\omega\gamma\tilde{x} + 4\sin^2(qa/2)k\tilde{x} = 0 \text{ (DPD)} \ . \tag{24}$$

Thus while Langevin and DPD damping do not alter the eigenfrequencies of the chain, i.e., their q dependence is, $\omega_0(q) = \sqrt{4\sin^2(qa/2)/m}$, the quality factor Q, defined as the ratio of eigenfrequency and damping, does differ between the two methods,

$$Q(q) = \frac{\omega_0(q)}{\gamma} \text{ (Langevin)}$$

$$Q(q) = \frac{\omega_0(q)}{\gamma}\,\frac{1}{4\sin^2(qa/2)} \text{ (DPD)} \ . \tag{25}$$

In the long wavelength limit where $q \to 0$, Langevin dynamics will always be overdamped, while DPD dynamics will be underdamped, provided that the system is not intrinsically overdamped as is the case in the vicinity of a continuous phase transition. Although these calculations have only be done for a linear harmonic chain, the results suggest that DPD has little effect on dynamical quantities that couple to long wavelengths. One of them would be the bulk viscosity of a system, although the estimation of DPD induced artifact in viscosity requires a different treatment than the one for sound waves [51]. Nevertheless, it was found that the measured bulk viscosity in sheared fluids depended on the precise choice of γ only in a negligible manner at least as long as γ was kept reasonably small [51].

In some cases, it may yet be beneficial to work with Langevin thermostats. The reason is that (elastic) long wavelength modes equilibrate notoriously slowly. Imagine we press an elastic solid onto a fractal surface. As the DPD thermostat barely damps long-range oscillations, we must expect a lot of bumping before the center-of-mass of the top wall finally comes to rest. Conversely, Langevin dynamics can lead to faster convergence because it couples

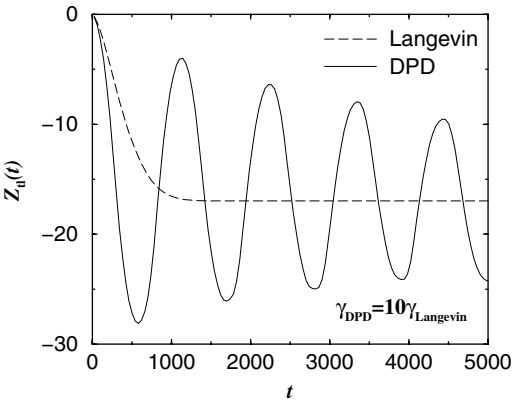


Fig. 10. Time dependence of the normal position Z_{tl} of an elastic solid, which is pressed against a self-affine substrate similar to the one shown in Fig. 7. Two different damping/thermostating schemes are employed, Langevin (*broken lines*) and DPD (*full lines*). Although the damping coefficient is 10 times greater in DPD than in Langevin, DPD-based dynamics are too strongly underdamped to relax efficiently to the right position

more strongly to long-wavelength oscillations. Figure 10 confirms this expectation. The Langevin-thermostatted system quickly reaches its mechanical position, while the DPD-thermostatted system is strongly underdamped, although the damping coefficient was 10 times larger for DPD than for Langevin.

In general, one needs to keep in mind that equilibrating quickly and producing realistic dynamics (or calculating thermal expectation values) are often mutually exclusive in simulations. It is necessary to consider carefully which aspect is more important for a given question of interest.

3.4 Determination of Bulk Viscosities

Theoretical Background: Role of Length Scales

In some cases friction between two surfaces is dominated by the bulk viscosity of the fluid embedded between the two surfaces [58]. When the surfaces are sufficiently far from one another and shear rates are low, one can usually assume that the velocity of the fluid near solid surfaces is close to the velocity of those surfaces. This scenario would be called a stick condition. As the distance, D, between the surfaces is decreased, one might have partial slip as alluded to in Fig. 6, in which the slip length, Λ, is introduced. The calculation of Λ from atomistic simulations is a subtle issue, [59] which we will not touch further upon here. When the fluid is confined even further, the concept of slip length might break down altogether and the measured friction becomes a true system function, which cannot be subdivided into smaller, independent entities. The discussion is summarized in the following equation

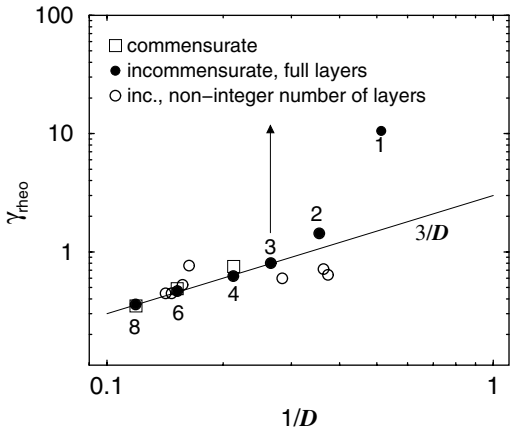


Fig. 11. Damping coefficient $\gamma_{\rm rheo} = F/Av$ obtained from simulating two atomically flat surfaces separated by a simple fluid consisting of monomers at constant temperature and normal pressure. Different coverages were investigated. The numbers in the graph denote the ratio of atoms contained in the fluid $N_{\rm fl}$ relative to the atoms contained per surface layer of one of the two confining walls $N_{\rm w}$. The walls are [111] surfaces of face-centered-cubic solids. They are rotated by 90° with respect to each other in the incommensurate cases. Full circles represent data for which $N_{\rm fl}/N_{\rm w}$ is an integer. The arrow indicates the point at which the damping coefficients for commensurate walls increases exponentially

$$F/A = \begin{cases} \eta v/D & \text{hydrodynamic regime} \\ \eta v/(D+\Lambda) & \text{moderate confinement} \\ ? & \text{strong confinement} \end{cases}, \tag{26}$$

where F/A is the force, F, per surface area, A, required to slide two solids seperated by a distance, D, at a velocity, v. η denotes the (linear-response) viscosity of the fluid between the walls. Figure 11 illustrates this point further. Shown is the linear response of a system similar to the one in the inset of Fig. 4. At large separations, the behaviour is reminiscent of hydrodynamic lubrication, i.e., the damping coefficient, $\gamma_{\rm rheo} = F/Av$, is approximately inversely proportional to D. As D is decreased, the shear response is starting to become very sensitive to the relative orientation of the two surfaces. In fact, the damping force for commensurate surfaces increases by several orders of magnitude by going from 4 layers to 3 layers of lubricant atoms. The large values for the effective damping can be understood from the discussion of lubricated commensurate surfaces in Sect. 2.4. Keep in mind that the finite size of the walls prevents the energy barrier to become inifinitely large for the commensurate walls. This is the reason why damping remains finite. Incommensurate walls do not show such size effects.

For a more detailed discussion of flow boundary conditions we refer to [2]. Here, it shall only be said that the two key ingredients from a microscopic

point of view are adhesion and the roughness of the walls. Perfectly flat walls will produce a perfect slip condition, while rough walls favor stick conditions. Increasing adhesion will also favor stick conditions at fixed roughness.

Lees-Edwards Boundary Conditions

In the hydrodynamic regime, it is unnecessary to impose shear via moving walls. It is often desirable to shear the fluid without any boundary effects. This can be achieved with the help of Lees-Edwards periodic boundary conditions, [60] which are illustrated in Fig. 12. Periodic boundary conditions are employed in all three spatial directions. However, while the center-of-mass of the central simulation cell remains fixed in space, many of its periodic images are moved parallel to the shear direction. As a consequence, even when a particle is fixed with respect to the central image, the distance to its periodically-repeated images will change with time if the vector connecting the two images contains a component parallel to the shear gradient direction.

To be specific, let $\mathbf{R}_{ij}$ denote the position in the periodically repeated cell which is the i's image to the right and the j's image on top of the central cell. (A potentially third dimension remains unaffected and will therefore not be mentioned in the following.) The position in real space of the vector $\mathbf{R}_{ij} = (X, Y)_{ij}$ would be

$$\begin{pmatrix} X \\ Y \end{pmatrix}_{ij} = \begin{pmatrix} X \\ Y \end{pmatrix}_{00} + \begin{bmatrix} 1 & \dot{\epsilon}\, t \\ 0 & 1 \end{bmatrix} \begin{pmatrix} iL_x \\ jL_y \end{pmatrix} , \tag{27}$$

$\dot{\epsilon}$ being the shear rate, and L_x and L_y being the length of the simulation cell in x and y direction, respectively. Thus, conventional periodic boundary

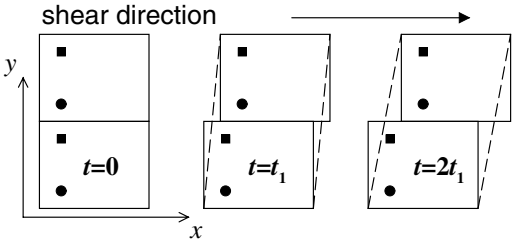


Fig. 12. Visualizations of Lees-Edwards periodic boundary conditions. At time zero, $t = 0$, regular periodic boundary conditions are employed. As time moves on, the periodic images of the central simulation cell move relative to the central cell in the shear direction as shown in the middle and the right graph. *Circle* and *square* show points in space that are fixed with respect to the (central) simulation cell. It is important to distribute the effect of shear homogeneously through the simulation cell such as indicated by the dashed lines. Otherwise, velocities will be discontinuous in shear direction whenever a particle corsses the simulation cell's boundary across the shear gradient direction. In this graph, x corresponds to the shear direction and y to the shear gradient direction

conditions can be reproduced by setting $\dot{\epsilon}$ to 0. When using Lees-Edwards periodic boundary condition, thermostatting is most naturally done with DPD thermostats, because no reference system needs to be defined.

When integrating the equations of motion, it is important to not only impose the shear at the boundaries, because this would break translational invariance. Instead, in each MD step of size Δt, we need to correct the position in shear direction. This is done, for instance, in the following fashion:

$$X_{n+1} = X_n + \Delta X_n + \dot{\epsilon}\Delta t Y_n \ , \tag{28}$$

where ΔX_n is the change in the y coordinate if no external shear were applied. This way, the effect of shear is more homogeneoulsy distributed over the simulation.

A better alternative to the implementation of Lees-Edwards boundary conditions is the formalism put forth by Parrinello and Rahman for the simulation of solids under constant stress [61]. They described the positions of particles by reduced, dimensionless coordinates r_α, where the r_α can take the value $0 \le r_\alpha < 1$ in the central image. Periodic images of a particle are generated by adding or subtracting integers from the individual components of $\mathbf{r}$.

The real coordinates of $\mathbf{R}$ are obtained by multiplying $\mathbf{r}$ with the matrix h that contains the vector spanning the simulation cell. In the present example, this would read

$$R_\alpha = \sum_\beta h_{\alpha\beta} r_\beta \tag{29}$$

$$h = \begin{bmatrix} L_x & 0 \\ 0 & L_y \end{bmatrix} \begin{bmatrix} 1 & \dot{\epsilon}\, t \\ 0 & 1 \end{bmatrix} \ . \tag{30}$$

The potential energy, V, is now a function of the reduced coordinates and the h-matrix. For the kinetic energy, one would only be interested in the motion of the particle relative to the distorted geometry so that a suitable Lagrange function, L_0, for the system would read

$$L_0 = \sum_i \frac{1}{2}\, m_i \left(\sum_\beta h_{\alpha\beta} \dot{r}_{i\beta} \right) - V(h, \{\mathbf{r}\}) \ , \tag{31}$$

in which the h matrix may be time dependent. From this Lagrangian, it is straightforward to derive the Newtonian equations for the reduced coordinates, which can then be solved according to the preferred integration schemes. One advantage of the scheme outlined in (29)–(31) is that it is relatively easy to allow for fluctuations of the size of the central cell. This is described further below.

Geometric and Topological Constraints

Note that L_x and L_y can be chosen to be time dependent. When simulating simple fluids under shear, there is no particular reason why they should be

chosen independent from one another. However, when simulating self-assembled monolayers under shear, which have received significant attention due to their technological applications, it may be necessary to allow independent fluctuations of the cell geometry along different spatial dimensions. This will be discussed in the context of diblock copolymers.

The most simple diblock copolymers are linear molecules, in which one part of the chain consists of one type of monomer, say polystyrene (PS), and the other one of another type, say polybutadiene (PB), see Fig. 13. PS and PB usually phase separate at low temperatures, however, due to their chemical connectivity, blockcopolymers cannot unmix on a macroscopic scale. They can only phase separate on a microscopic scale, whose size is determined by the length of the polymers.

When lamellar structures are formed, it is necessary to choose the dimension of the simulation cell commensurate with the intrinsic periodicity of the lamellae, in order to avoid any pressure that is unintentionally exerted due to geometric constraints. It is therefore desirable to allow the system to fluctuate parallel to "solid directions," which are introduced in Fig. 13. For these directions, it would be appropriate to employ the usual techniques related to constant stress simulations [61].

Let us consider the three-dimensional case and further work within the Parrinello-Rahman framework. A rather general three-dimensional h matrix will be considered

$$h = \begin{bmatrix} L_{xx} & 0 & L_{xz} \\ 0 & L_{yy} & L_{yz} \\ L_{xz} & L_{yz} & L_{zz} \end{bmatrix} \begin{bmatrix} 1 & \dot{\epsilon}\, t & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} . \tag{32}$$

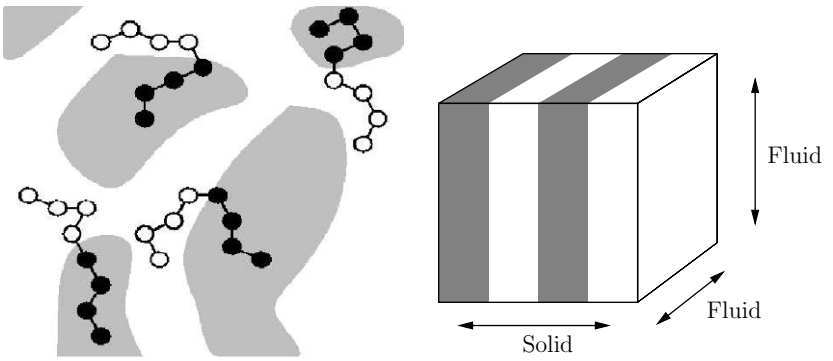


Fig. 13. Schematic representation of the microphase separation of block copolymers. The *left* graph shows atomic-scale details of the phase separation at intermediate temperatures, the *right* graph a lamellar phase formed by block copolymers at low temperatures. The block copolymers have solid-like properties normal to the lamellae, due to a well-defined periodicity. In the other two directions, the system is isotropic and has fluid-like characteristics. From [62]

It is now possible to treat the variables $L_{\alpha\beta}$ as generalized coordinates and allow them to adjust within the MD simulation. For this purpose, it is necessary to define a kinetic energy $T_{\rm cell}$ associated with the fluctuation of the cell geometry as a bilinear function of generalized velocities $\dot{L}_{\alpha\beta}$,

$$T_{\rm cell} = \sum_{\alpha\beta\gamma\delta} \frac{1}{2} M_{\alpha\beta\gamma\delta} \dot{L}_{\alpha\beta} \dot{L}_{\gamma\delta} \ , \tag{33}$$

where $M_{\alpha\beta\gamma\delta}$ must be a positive definite matrix with the unit of mass. Though the optimum choice for the M-matrix is a matter of discussion, [63] a reasonable approach is to treat the various $\dot{L}_{\alpha\beta}$ as independent, uncoupled variables and to assign the same mass $M_{\rm cell}$ to all diagonal elements, $M_{\alpha\beta\alpha\beta}$ which simplifies (33) to

$$T_{\rm cell} = \sum_{\alpha,\beta\le\alpha} \frac{1}{2} M_{\rm cell} \dot{L}_{\alpha\beta}^2 \ . \tag{34}$$

It is often sensible to choose $M_{\rm cell}$ such that the simulation cell adjusts on microscopic time scales to the external pressure.

The Lagrange function L for Lees-Edwards boundary conditions combined with Parrinello Rahman fluctuations for the cell geometry now reads

$$L = L_0 + T_{\rm cell} - p \det h \ , \tag{35}$$

where p is an isotropic pressure and $\det h$ the volume of the simulation cell. The Newtonian equations of motion for the generalized coordinates $L_{\alpha\beta}$ and $\mathbf{r}_i$ follow from the Lagrange formalism. Furthermore, it is possible to couple fluctuating cell geometries not only to constant isotropic pressure but also to non-isotropic stresses. The description of these approaches is beyond the scope of the present manuscript, but they can be found in the literature [61].

When studying sytems with mixed "fluid" and "solid" directions, it is important to keep in mind that each solid direction should be allowed to breathe and fluid directions need to be scaled isotropically or to be constrained to a constant value. Allowing two fluid directions to fluctuate independently from one another allows the simulation cell to become flat like a pancake, which we certainly would like to avoid. To give an example, consider Fig. 14, in which a lamellar block copolymer phase is sheared. The convention would be to have the shear direction parallel to x and the shear gradient direction parallel to y. There is no reason for the simulation cell to distort such that $L_{xz} = L_{yz} = 0$ would not be satisfied on average, so one may fix the values of L_{xz} and L_{yz} from the beginning. There is one solid direction plus two fluid directions. We can also constrain L_{xx} to a constant value, because the shear direction will always be fluid and there is another fluid direction that can fluctuate. This means that we should allow the simulation cell to fluctuate independently in the shear and the shear gradient direction. Yet during the reorientation process, i.e., during the intermediate stage shown in Fig. 14,

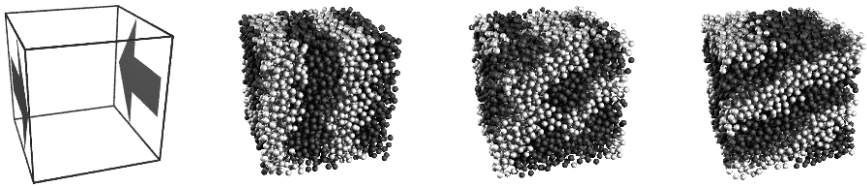


Fig. 14. A lamellar block copolymer phase is reoriented through external shear. The initial phase has the director of the lamellae parallel to the shear gradient direction. The most stable state would be to orient the director parallel to the shear and shear gradient direction. However, the reorientation process gets stuck before true equilibrium is reached. The stuck orientation is relatively stable, because the lamellae have to be broken up before they can further align with respect to the shear flow. From [64]

simulation cells do have the tendency to flatten out, because periodicity and hence solid like behaviour is lost for a brief moment in time.

It is interesting to note that the reorientation process of the lamellae does not find the true equilibrium state but gets stuck in a metastable state. The periodic boundary conditions impose a topological constraint and prevent the system from simply reorienting. It is conceivable that similar metastable states are also obtained experimentally, although the nature of the constraints differs in both cases. One means of overcoming this topological constraint is to impose a higher temperature, T_{h}, at the boundaries of the simulation cell (e.g. at $0 \leq r_y \leq 0.2$ and $0 \leq r_z \leq 0.2$) and to keep the temperature low in the remaining system [62]. This would melt the lamellar structure at the boundary and allow the remaining lamellae to reorient freely with respect to the shear flow.

4 Selected Case Studies

The last few years have seen an increasing number of tribological simulations that incorporate realistic potentials and/or realistic boundary conditions. Unfortunately, it is not possible in this chapter to give all of these studies the exposure that they may deserve. Instead, two subjects will be selected. One deals with the occurrence and breakdown of superlubricity. The other subject evolves around the chemical complexity of real lubricant mixture and ways how simulations can aid in the rational design of lubricant mixtures.

4.1 Superlubricity and the Role of Roughness at the Nanometer Scale

In Sect. 2, it is argued that two three-dimensional solids with clean, flat surfaces should be superlubric, unless conditions are extreme. Calculations of

Hirano and Shinjo, in particular, supported this picture at an early stage for three-dimensional solids and realistic model potentials [4, 5]. Theoretical considerations, moreover, predict that low-dimensional systems are less likely to be superlubric because they are more easily deformed on large length scales than highly-dimensional structures [9]. But, even for one-dimensional Frenkel Kontorova (FK) chains, the interaction between chain and substrate has to be sufficiently strong, in order to produce finite kinetic friction [34]. Reasonable parametrizations of the FK model would require the harmonic coupling between adjacent atoms in the chain, k, to be larger than the maximum curvature, $V''_{\max}$, of the embedding potential [9]. In this regime, FK models are typically superlubric [34]. Thus, all theoretical considerations point to superlubricity between unaligned solids, provided that the interfaces are sufficiently flat. While more and more systems are found to be superlubric, some experiments find clear violations of superlubricity, although theory predicts these same systems to be superlubric. One example, is the measurement of friction anisotropy at Ni(100)/Ni(100) interfaces in ultra high vacuum [65]. When misoriented by 45°, friction coefficients remained in the order of 2.5 instead of becoming unmeasurably small. The simulations by Tangney et al. [66] and by Qi et al. [67] probably clarify the seeming discrepancy between theory and simulation.

Tangney et al. [66] studied the friction between an inner and an outer carbon nanotube. Realistic potentials were used for the interactions within each nanotube and Lennard Jones potentials were employed to model the dispersive interactions between nanotubes. The intra-tube interaction potentials were varied and for some purposes even increased by a factor of 10 beyond realistic parametrizations, thus artificially favoring the onset of instabilities and friction. Two geometries were studied, one in which inner and outer tubes were commensurate and one in which they were incommensurate.

Let us rewind for a second to discuss the two idealized egg cartons mentioned in the beginning of Sect. 2.4. Aligning the two egg cartons explains static friction, but, as long as no additional "microscopic" dissipation mechanism is present, there is no reason why our egg cartons should show static friction. Whenever the top egg carton slides downwards, kinetic energy is produced that will help it to climb up the next potential energy maximum. Instabilities would be required to produce static friction. Thus, the presence of static friction does not imply kinetic friction. The simulations of the nanotubes exhibit exactly this behaviour, which is demonstrated in Fig. 15.

In interpreting Fig. 15, it is important to know that the outer tube has two open endings and that both tubes are equally long. The embedded tube only has a small fraction inside the outer tube initially. The system will try to reduce surface energy by sucking the outer tube inside the inner tube. If there is no friction between the tubes, then the force will be constant until the inner tube starts to exit at the other end at high velocity. When the nanotube is almost completely exited at the other end, the procedure will repeat and surface-energy driven oscillations will occur. This is what is found

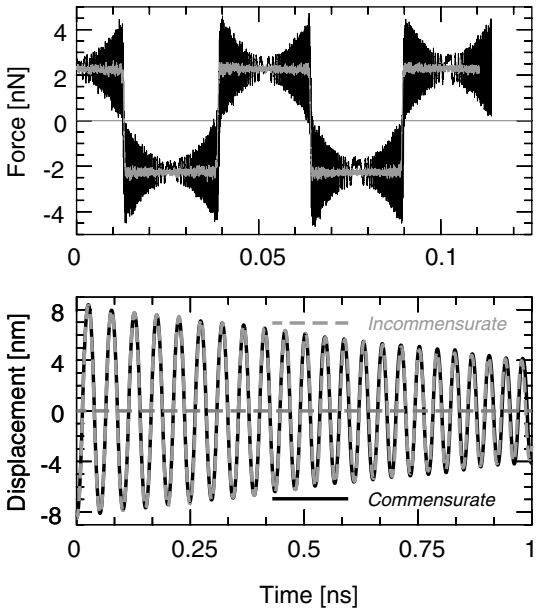


Fig. 15. *Top*: Friction force between two nanotubes as a function of time. *Bottom*: Displacement of the nanotubes as a funtion of time. Grey and black lines indicate incommensurate and commensurate geometries, respectively. From [66]

for the incommensurate tube, except for some small fluctuations in the instantaneous force and some weak viscous-type damping. The commensurate system behaves slightly different. There are large fluctuations in the instantaneous force, which have the periodicity of the lattice. The fluctuations are particularly large when the inner tube is fully immersed in the outer one. The fluctuations can be understood within the egg carton model. Moreover, from the absence of significant dissipation on long time scales, it is possible to conclude that no instabilities occur. Thus the simulated, low-dimensional rubbing system exhibits superlubricity, not only for incommensurate but even for commensurate surfaces. Finite static and zero kinetic friction forces have also been observed experimentally, albeit for a different system [68]. Although the nanotube simulations were based on reasonably realistic potentials, it needs to be emphasized that real carbon nanotubes have a lot of chemical defects, which induce the experimentally measured non-viscous type friction forces [69].

It turns out that the detailed microstructrue matters just as much for three-dimensional systems as for nanotubes. Qi et al. [67] studied atomically smooth Ni(100)/Ni(100) interfaces. Their idealized geomtries display the same superlubric behaviour as the idealized copper interfaces studied by Hirano and Shinjo [4]. However, roughening the top layer with a mere 0.8 Å rms variation, changes the behaviour completely, with friction coefficients increasing by several orders of magnitude as can be seen in Fig. 16.

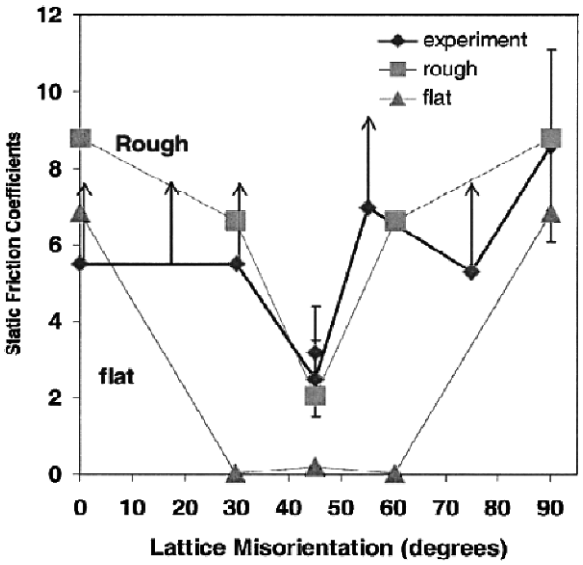


Fig. 16. Friction coefficient for differently aligned Ni(100)/Ni(100) interfaces. Rough surfaces have a 0.8 Å rms variation in roughness added to the atomically smooth surfaces. From [67]

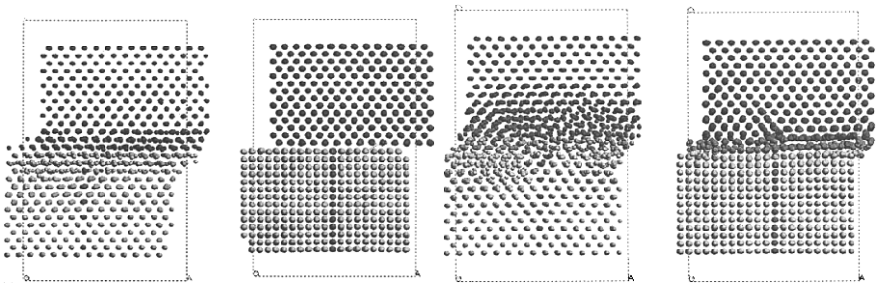


Fig. 17. Snapshots from the simulations leading to the friction coefficients shown in Fig. 16. From the left to the right: Atomically flat commensurate, atomically flat incommensurate, rough commensurate, and rough incommensurate geometries. Only the flat incommensurate surfaces remain undamaged resulting in abnormally small friction coefficients. From [67]

The microscopic origin of the increase in friction can be understood from the microstructures shown in Fig. 17. Due to sliding, all contacts deform plastically except for the atomically flat, incommensurate contact. As the atoms are no longer elastically coupled to their lattice sites, they can interlock the surfaces in a way that is roughly akin of the scenario shown in Fig. 2. Of course, when placing two commensurate, atomically smooth solids on top of each other, the identity of each of the solids disappears instantaneously, and

sliding them corresponds to shearing a single crystal, which certainly has finite resistance to shear. No intervening layer is required in this case to interlock the surfaces.

4.2 Physics and Chemistry of Lubricant Additives

Most of this chapter evolves around simple models of surfaces and lubricants. There is a lot we can learn from these generic models, but it is important to keep in mind that real lubricants are rather complex mixtures. There are many different chemicals which are added to the so-called base oil of industrial lubricants used in car engines. These additives function to reduce the formation of foam, to reduce friction, to disperse debris, to act as anti-oxidants for nascent surface material, and to reduce wear. The rational formulation of an industrial lubricant formulation not only requires a knowledge of each additive, but it is also imperative to understand the interactions between the additives.

One particular class of anti-wear additives are zinc dialkyldithiophosphates (ZDDPs), which were invented in the 1930's as anti-oxidants. Their main function, however, turned out to be the protection of cast iron surfaces from wear. Experiments revealed that the ZDDPs decompose into simple zinc phosphates (ZPs) under the conditions at which the lubricant operates and that these decomposition products form films or patches of ZPs on surfaces. In what follows, these patches will be called anti-wear pads (AWP). Although an abundance of experimental data was available regarding the ZDDPs, no coherent, molecular theory existed that could explain how the AWPs form and function [70]. Relevent experimental observations which must be accounted for by such a theory include the following: AWPs formed on the tops of asperities are harder and more elastic than those in the valleys between asperities, the spectra of the AWPs on the tops of asperities are reminiscent of ZPs with longer chain lengths, while those in the valleys are characteristic of shorter ZP chains. Additionally, the ability of the AWPs to inhibit wear is reduced when zinc is replaced by another charge-balancing cation, such as calcium. Lastly, the rapid formation of AWPs under sliding conditions and the inability of ZDDPs to protect aluminum surfaces were not understood.

Previous models for ZP-AWP formation were based on reaction schemes in which iron atoms acted catalytically in the formation of the pads. It was believed that the reaction required the high temperatures typically found in the microscopic points of contact. However, the conditions found in tribological contacts are far from ambient and it may therefore not be sufficient to only incorporate the effect of temperature. In contrast, pressures can be extremely high, and approach the theoretical yield pressure for a short period of time. With these considerations in mind, it is natural to ask, how the lubricant molecules respond to the extreme conditions encountered in tribological contacts.

Mosey et al. [71] addressed this question by exposing the relevant decomposition products of ZDDPs to pressures close to the theoretical yield strengths of steel and aluminium, respectively. Configurations from these simulations before, during, and after compression are shown in Fig. 18. It is found that

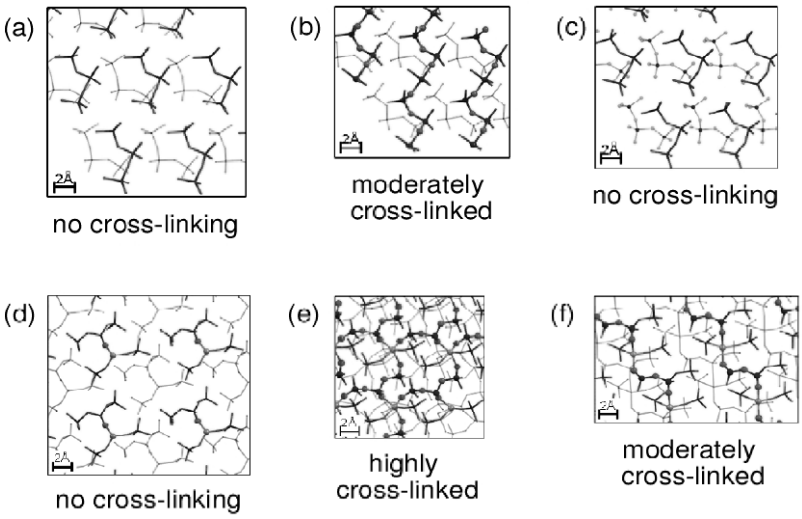


Fig. 18. Molecular configuration of triphosphates (TPs) (**a**)–(**c**) and zinc phosphates (ZPs) (**d**)–(**f**). The graphs (a) and (d) correspond to an initial low-pressure structure. The graphs (b) and (e) show configurations at high pressure, in which both TPs and ZPs form covalent cross links. After pressure is released back to ambient pressures, (c) and (f), only the ZPs remain a chemically connected network

high pressure is a sufficient condition to form chemically connected networks. However, these networks only remain intact under decompression if zinc is present as a crosslink-forming agent. The formation of the cross-links occurs through changes in the coordination at the zinc atom. Since calcium does not exhibit a variable coordination, replacing zinc with that atom decreases the degree to which cross-linking occurs, thereby reducing the ability of the film to inhibit wear. From the simulations one may conclude that the degree of chemical connectivity is particularly large, when the pressure to which the ZPs are exposed are very high. Moreover, it was found that the chemically cross-linked ZPs are much harder after the compression than before. Lastly, the formed pads are harder than aluminium surfaces suggesting that they cannot effectively redistribute pressures on aluminium surfaces and instead abdrade them. All these conclusions are in correspondance to the experimental observations.

More importantly, the simulations provide guidelines with which to screen for replacements of the environmentally problematic ZDDPs. One needs to identify molecules that can enter the microscopic points of contact as a

viscous fluid and that crosslink through the application of the typically encountered pressure and temperature conditions. (Temperature appeared to play a negligible role in the simulations.) The resulting cross-linked AWPs should be softer, but not much softer, than the substrate so that the AWPs can redistribute the pressure exerted and thereby alleviate the extreme conditions to which the asperities are exposed. This example is only one of many which show that simulations bear the potential to not only address questions of fundamental scientific interest, but also guide in the design of new materials.

Acknowledgments

Financial support from the Natural Sciences and Engineering Research Council of Canada and SHARCnet (Ontario) is gratefully acknowledged. The author thanks Carlos Campana for providing Figs. 7 and 10 and Nick Mosey for providing Fig. 18. The authors also acknowledges helpful discussions with Carlos Campana and Nick Mosey.

References

1. J. A. Harrison, S. J. Stuart, and D. W. Brenner (1999) Atomic-scale simulation of tribological and related phenomena. In *Handbook of Micro/Nanotribology* Ed. B. Bhushan (CRC Press, Boca Raton, FL), pp. 525–594
2. M. O. Robbins and M. H. Müser (2001) Computer simulations of friction, lubrication, and wear. In *Modern Tribology Handbook* Ed. B. Bhushan (CRC Press, Boca Raton, FL), pp. 717–765 (cond-mat/0001056)
3. M. H. Müser, M. Urbakh, and M. O. Robbins (2003) Statistical mechanics of static and low-velocity kinetic friction. *Adv. Chem. Phys.* **126**, pp. 187–272
4. M. Hirano and K. Shinjo (1990) Atomistic locking and friction. *Phys. Rev. B* **41**, pp. 11837–11851
5. K. Shinjo and M. Hirano (1993) Dynamics of friction - superlubric state *Surf. Sci.* **283**, pp. 473–478
6. L. Prandtl (1928) Ein Gedankenmodell zur kinetischen Theorie der festen Körper. *Z. Angew. Math. Mech.* **8**, pp. 85–106
7. G. A. Tomlinson (1929) A molecular theory of friction. *Philos. Mag. Series* **7**, pp. 905–939
8. T. Baumberger and C. Caroli (1990) Dry friction dynamics at low velocities. *Comments Cond. Mater. Phys.* **62**, p. 251
9. M. H. Müser (2004) Structural lubricity: Role of dimension and symmetry. *Europhys. Lett.* **66**, pp. 97–103
10. B. N. J. Persson (2001) Theory of rubber friction and contact mechanics. *J. Chem. Phys.* **115**, pp. 3840–3861
11. F. P. Bowden and D. Tabor (1986) *The Friction and Lubrication of Solids.* Clarendon Press, Oxford
12. S. Hyun, L. Pei, J.-F. Molinari, and M. O. Robbins (2004) Finite-element analysis of contact between elastic self-affine surfaces. *Phys. Rev. E* **70**, p. 026117

13. G. He, M. H. Müser, and M. O. Robbins (1999) Adsorbed layers and the origin of static friction. *Science* **284**, pp. 1650–1652
14. M. H. Müser, L. Wenning, and M. O. Robbins (2001) Simple microscopic theory of Amontons' laws for static friction. *Phys. Rev. Lett.* **86** pp. 1295–1298
15. J. N. Glosli and G. McClelland (1993) Molecular dynamics study of sliding friction of ordered organic monolayers. *Phys. Rev. Lett.* **70**, pp. 1960–1963
16. H. A. Harrison, C. T. White, R. J. Colton, and D. W. Brenner (1995) Investigation of the atomic-scale friction and energy-dissipation in diamond using molecular-dynamics. *Thin Solid Films* **260**, pp. 205–211
17. G. T. Gao, P. T. Mikulski, and J. A. Harrison (2002) Molecular-scale tribology of amorphous carbon coatings: Effects of film thickness, adhesion, and long-range interactions. *J. Am. Chem. Soc.* **124**, pp. 7202–7209
18. N. J. Mosey, T. K. Woo, and M. H. Müser (2005) Energy dissipation via quantum chemical hysteresis during high-pressure compression: A first-principles molecular dynamics study of phosphates. *Phys. Rev. B* **72**, p. 054124
19. D. Dowson (1979) *History of Tribology* Longman, New York
20. M. H. Müser (2002) Simple microscopic theory of Amontons's laws for static friction. *Phys. Rev. Lett.* **89**, p. 224301
21. Y. Sang, M. Dubé, and M. Grant (2001) Thermal effects on atomic friction. *Phys. Rev. Lett.* **87**, p. 174301
22. O. K. Dudko, A. E. Filippov, J. Klafter, and M. Urbakh (2002) Dynamic force spectroscopy: a Fokker-Planck approach. *Chem. Phys. Lett.* **352**, pp. 499–504
23. G. He and M. O. Robbins (2001) Simulations of the kinetic friction due to adsorbed surface layers. *Tribol. Lett.* **10**, pp. 7–14
24. J. H. Dieterich and B. D. Kilgore (1994) Direct observation of frictional contacts: New insights for state-dependent properties. *Pure Applied Geophys.* **143**, pp. 283–302
25. J. H. Dieterich and B. D. Kilgore (1996) Imaging surface contacts: power law contact distributions and contact stresses in quartz, calcite, glass and acrylic plastic. *Tectonophysics* **256**, pp. 219–239
26. G. He and M. O. Robbins (2001) Simulations of the static friction due to adsorbed molecules. *Phys. Rev. B* **64**, p. 035413
27. L. Wenning and M. H. Müser (2001) Friction laws for elastic nano-scale contacts. *Europhys. Lett.* **54**, pp. 693–699
28. A. W. Bush, R. D. Gibson, and T. R. Thomas (1975) The Elastic Contact of a Rough Surface. *Wear*, **35**, p. 87
29. B. N. J. Persson (2001) Elastoplastic contact between randomly rough surfaces. *Phys. Rev. Lett.* **87**, p. 116101
30. B. Q. Luan, S. Hyun, M. O. Robbins, and N. Bernstein (2005) Multiscale modeling of two dimensional rough surface contacts. *Mater. Res. Soc. Sympp. Proc.* **841**, pp. R7.4.1–10
31. M. Schoen, C. L. Rhykerd, D. J. Diestler, and J. H. Cushman (1989) Shear forces in molecularly thin films. *Science* **245**, pp. 1223–1225
32. M. H. Müser and M. O. Robbins (2000) Conditions for static friction between flat crystalline surfaces. *Phys. Rev. B* **61**, pp. 2335–2342
33. E. Gnecco, R. Bennewitz, T. Gyalog, and E. Meyer (2001) Friction experiments on the nanometre scale. *J. Phys.: Condens. Mat.* **13**, pp. R619–R642
34. O. M. Braun and Y. S. Kivshar (1998) Nonlinear dynamics of the Frenkel-Kontorova model. *Phys. Repp.* **306**, pp. 2–108

35. J. E. Hammerberg, B. L. Holian, J. Roder, A. R. Bishop, and S. J. Zhou (1998) Nonlinear dynamics and the problem of slip at material interfaces. *Physica D* **123**, pp. 330–340
36. W. A. Steele (1973) The physical interaction of gases with crystalline solids. *Surf. Sci.* **36**, pp. 317–352
37. M. O. Robbins and P. A. Thompson (1991) Critical velocity of stick-slip motion. *Science* **253**, p. 916
38. B. Q. Luan and M. O. Robbins (2004) Effect of inertia and elasticity on stick-slip motion. *Phys. Rev. Lett.* **93**, p. 036105
39. V. Zaloj, M. Urbakh, and J. Klafter (1998) Atomic scale friction: What can be deduced from the response to a harmonic drive? *Phys. Rev. Lett.* **81**, pp. 1227–1230
40. P. Meakin (1998) *Fractals, Scaling and Growth far from Equilibrium*. Cambridge University Press, Cambridge
41. J. Q. Broughton, F. F. Abraham, N. Bernstein, and E. Kaxiras (1999) Concurrent coupling of length scales: Methodology and application. *Phys. Rev. B* **60**, pp. 2391–2403
42. W. A. Curtin and R. E. Miller (2003) Atomistic/continuum coupling in computational materials science. *Model. Simul. Mater. Sci.* **11**, pp. R33–R68
43. Y. Saito (2004) Elastic lattice Green's function in three dimensions. *J. Phys. Soc. Japan* **73**, p. 1816
44. W. Cai, M. de Koning, V. V. Bulatov, and S. Yip (2000) Minimizing boundary reflections in coupled-domain simulations. *Phys. Rev. Lett.* **85**, pp. 3213–3216
45. S. A. Adelman and J. D. Doll (1974) (a) Generalized Langevin equation approach for atom/solid-surface scattering: Collinear atom/harmonic chain model. (b) Erratum: Generalized Langevin equation approach for atom/solid-surface scattering: Collinear atom/harmonic chain model. (c) Generalized Langevin equation approach for atom/solid-surface scattering: General formulation for classical scattering off harmonic solids. *J. Chem. Phys.* **61**, pp. 4242–4245; *ibid.*, (1975) **62**, pp. 2518–2518; *ibid.*, (1976) **64**, pp. 2375–2388
46. R. Kubo (1966) The fluctuation-dissipation theorem. *Rep. Prog. Phys.* **29**, pp. 255–284
47. T. Schneider and E. Stoll (1978) Molecular-dynamics study of a three-dimensional one-component model for distortive phase transitions. *Phys. Rev. B* **17**, pp. 1302–1322
48. P. Español and P. Warren (1995) Statistical mechanics of dissipative particle dynamics. *Europhys. Lett.* **30**, pp. 191–196
49. S. S. Sarman, D. J. Evans, and P. T. Cummings (1998) Recent developments in non-Newtonian molecular dynamics. *Phys. Rep.* **305**, pp. 1–92
50. P. Español (1995) Hydrodynamics from dissipative particle dynamics. *Phys. Rev. E* **52**, pp. 1734–1742
51. T. Soddemann, B. Dünweg, and K. Kremer (2003) Dissipative particle dynamics: A useful thermostat for equilibrium and nonequilibrium molecular dynamics simulations. *Phys. Rev. E* **68**, p. 046702
52. A. Ricci and G. Ciccotti (2003) Algorithms for Brownian dynamics. *Mol. Phys.* **101**, pp. 1927–1931
53. D. J. Evans and G. P. Morriss (1990) *Statistical Mechanics of Nonequilibrium Liquids*. Academic Press, London; http://rsc.anu.edu.au/evans/evansmorrissbook.htm

54. D. J. Evans and G. P. Morriss (1986) Shear thickening and turbulence in simple fluids. *Phys. Rev. Lett.* **56**, pp. 2172–2175
55. W. Loose and G. Ciccotti (1992) Temperature and temperature control in nonequilibrium-molecular-dynamics simulations of the shear-flow of dense liquids. *Phys. Rev. A* **45**, pp. 3859–3866
56. M. J. Stevens and M. O. Robbins (1993) Simulations of shear-induced melting and ordering. *Phys. Rev. E* **48**, pp. 3778–3792
57. P. Reimann and M. Evstigneev (2004) Nonmonotonic velocity dependence of atomic friction. *Phys. Rev. Lett.* **93**, p. 230802
58. D. Dowson and G. R. Higginson (1968) *Elastohydrodynamic Lubrication.* Pergamon, Oxford
59. L. Bocquet and J.-L. Barrat (1993) Hydrodynamic boundary-conditions and correlation-functions of confined fluids. *Phys. Rev. Lett.* **70**, pp. 2726–2729
60. A. W. Lees and S. F. Edwards (1972) The computer study of transport processes under extreme conditions. *J. Phys. C* **5**, pp. 1921–1928
61. M. Parrinello and A. Rahman (1981) (a) Polymorphic transitions in single-crystals – a new molecular-dynamics method. (b) Strain fluctuations and elastic-constants. *J. Appl. Phys.* **52**, p. 7182; *ibid.* (1982); *J. Chem. Phys.* **76**, p. 2662
62. B. Fraser, M. H. Müser, and C. Denniston (2005) Diffusion, elasticity, and shear flow in self-assembled block copolymers: A molecular dynamics study. *J. Pol. Sci. B* **43**, pp. 969–981
63. R. M. Wentzcovitch (1991) Invariant molecular-dynamics approach to structural phase-transitions. *Phys. Rev. B* **44**, pp. 2358–2361
64. B. Fraser (2005) *Ph.D. thesis.* University of Western Ontario
65. J. S. Ko and A. J. Gellman (2000) Friction anisotropy at Ni(100)/Ni(100) interfaces. *Langmuir* **16**, pp. 8343–8351
66. P. Tangney, S. G. Louie, and M. L. Cohen (2004) Dynamic sliding friction between concentric carbon nanotubes. *Phys. Rev. Lett.* **93**, 065503
67. Y. Qi, Y.-T. Cheng, T. Cagian, and W. A. Goddard III (2002) Friction anisotropy at Ni(100)/(100) interfaces: Molecular dynamics studies. *Phys. Rev. B* **66**, 085420
68. A. Socoliuc, R. Bennewitz, E. Gnecco, and E. Meyer (2004) Transition from stick-slip to continuous sliding in atomic friction: Entering a new regime of ultralow friction. *Phys. Rev. Lett.* **92**, p. 134301
69. Z. Xia and W. A. Curtin (2004) Pullout forces and friction in multiwall carbon nanotubes. *Phys. Rev. B* **69**, p. 233408
70. M. A. Nicholls, T. Do, P. R. Norton, M. Kasrai, and G. M. Bancroft (2005) Review of the lubrication of metallic surfaces by zinc dialkyl-dithlophosphates. *Tribol. Int.* **38**, pp. 15–39
71. N. J. Mosey, M. H. Müser, and T. K. Woo (2005) Molecular mechanisms for the functionality of lubricant additives. *Science* **307**, pp. 1612–1615

Simulation of Nanodroplets on Solid Surfaces: Wetting, Spreading and Bridging

A. Milchev

Institute of Physical Chemistry, Bulgarian Academy of Sciences G. Bonchev Str. Block 11, 1113 Sofia, Bulgaria
milchev@ipc.bas.bg



Andrey Milchev

A. Milchev: *Simulation of Nanodroplets on Solid Surfaces: Wetting, Spreading and Bridging*, Lect. Notes Phys. **704**, 105–126 (2006)
DOI 10.1007/3-540-35284-8_5
© Springer-Verlag Berlin Heidelberg 2006

1 Introduction 107

2 Theoretical Model 108

3 Nanodroplets on Flat Structureless Substrates 109

3.1 Estimation of the Surface Tension γ_{lv} from Capillary Wave Analysis 111

3.2 Estimation of γ_{wl} from the Local Pressure Tensor Anisotropy 111

4 Nanodroplets Adsorbed on Nanofibers 112

5 Spreading of Droplets on Horizontal Surfaces 114

6 Liquid Bridges and Induced Forces on Chemically Decorated Droplets with Varying Wettability 117

7 Film Rupture 120

8 Conclusions and Outlook 123

References 124

The structure, wetting behavior and spreading dynamics of small polymer melt droplets sitting on flat solid substrates is studied by Monte Carlo and Molecular Dynamics computer simulation, using a coarse-grained bead-spring model of flexible macromolecules. For nanodroplets adsorbed on ultrathin cylindric fibers we observe a transition from an axially-symmetric "barrel" shape to asymmetric "clam" shape upon changes in the fiber radius, the size of the droplet, or the strength of the adhesive potential. In the case of lyophobic walls, chemically decorated with lyophilic domains of circular or rectangular shape, we investigate the change in drop morphology and contact angle upon variation of the domain size. The formation and stability against rupture of *liquid bridges* between lyophilic domains just opposite of one another on parallel substrates is also examined and the bridging force is measured at varying wall separation.

1 Introduction

A wide variety of applications such as in microelectronics, catalysis, litography, as well as more recently in biology and life sciences are concerned with the wetting of surface structures. One particularly useful aspect of structured substrates is the ability to chemically pattern a substrate in order to achieve preferential adsorption (wetting) of a desired component [1]. Chemically structured surfaces that exhibit lateral patterns of varying wettability can be produced by means of photolitography [2, 3], microcontact printing [4–6], vapor deposition through grids [7], domain formation in Langmuir-Blogget monolayers [8, 9], electrophoretic colloid assembly [10], lithography with colloid monolayers [11], microphase separation in diblock copolymer films [12], etc. However, the quantitative understanding of such phenomena in terms of pertinent theory is difficult, since often the necessary material properties (e.g. interfacial tensions) are incompletely known. Also a number of questions concerning how ideal the substrate surfaces are (roughness, chemical heterogeneity), polydispersity of the droplets, impurities in the liquid, etc., may be a problem. A number of such complications can be easily eliminated in a computer simulations so that the relevant generic features of the phenomena in concern are revealed. In addition, one can vary the control parameters, such as the strength of the liquid – substrate interactions and can monitor many quantities of interest simultaneously.

In this contribution we provide a brief overview over computer experiments pertaining to the behavior of very small droplets on solid surfaces. Depending on the curvature and the chemical patterns of such substrates, one observes some fascinating phenomena even if the concrete atomistic structure of the surface is ignored. In what follows we focus on the wetting, spreading and bridging phenomena in such systems and their interpretation within the framework of statistical physics.

2 Theoretical Model

In our simulation studies both Monte Carlo (MC) and Molecular Dynamics (MD) implementations of a coarse-grained bead-spring type model have been employed whereby the detailed atomistic structure of the adsorbing substrate is ignored. Each polymer chain contains N effective monomers connected by anharmonic springs described by the finitely extendible nonlinear elastic (FENE) potential.

$$U_{FENE} = -\frac{K}{2} R^2 \ln \left[1 - \frac{(\ell - \ell_0)^2}{R^2} \right] \tag{1}$$

Here ℓ is the length of an effective bond, which can vary in between $\ell_{min} < \ell < \ell_{max}$, with $\ell_{min} = 0.4$, $l_{max} = 1$ being the unit of length, and has the equilibrium value $l_0 = 0.7$, while $R = \ell_{max} - \ell_0 = \ell_0 - \ell_{min} = 0.3$, and the spring constant K is taken as $K/k_BT = 40$. The nonbonded interactions between the effective monomers are described by the Morse potential

$$U_M = \epsilon_M \{ \exp[-2\alpha(r - r_{min})] - 2\exp[-\alpha(r - r_{min})] \} , \tag{2}$$

where r is the distance between the beads, and the parameters are chosen as $r_{min} = 0.8$, $\epsilon_M = 1$, and $\alpha = 24$. Owing to the large value of this latter constant, $U_M(r)$ decays to zero very rapidly for $r > r_{min}$, and is completely negligible for distances larger than unity. This choice of parameters is useful from a computational point of view, since it allows the use of a very efficient link-cell algorithm. Physically, these potentials (1), (2) make sense when one interprets the effective bonds as kinds of Kuhn segments, comprising a number of chemical monomers along the chain, and thus the length unit $\ell_{max} = 1$ corresponds physically rather to 1 nm than to the length of a covalent $C - C$ bond (which would only be about 1.5 Å). Therefore it also makes sense to treat the surface of the adsorbing surfaces on this coarse-grained length scale as perfectly flat and smooth: any atomistic corrugation of these surfaces would be on a scale much finer than the length scale ℓ_{max} of the Kuhn segment, and hence is not resolved in our coarse-grained model.

The MD runs are performed applying the standard velocity Verlet algorithm and the temperature is kept constant via the Langevin thermostat, which provides a very good stability of the algorithm. The coordinate of each monomer then changes according to

$$m\ddot{\boldsymbol{r}}_i = \boldsymbol{F}_i - \xi(\Delta z)\dot{\boldsymbol{r}}_i + W_i(t) , \tag{3}$$

where $m = 1$ is the mass of an effective monomer, $\boldsymbol{F}_i$ is the force deriving from all the potentials

$$\boldsymbol{F}_i = -\sum_{j(\neq i)} \partial U_{ij}/\partial \boldsymbol{r}_i , \quad U_{ij} = U_{ij}^{LJ} + U_{ij}^{\mathrm{FENE}} + U_{ij}^{\mathrm{wall}} . \tag{4}$$

In our MD studies we use a standard Lennard-Jones potential, U^{LJ}, rather than the interaction (2). $\boldsymbol{W}_i(t)$ is a Gaussian white noise term,

$$\langle \boldsymbol{W}_i(t) \cdot \boldsymbol{W}_j(t') \rangle = 6k_B T m \xi \delta_{ij} \delta(t - t') \; . \tag{5}$$

Note that we use the thermostat near the wall only,

$$\xi(\Delta z) = \xi^* \exp(\sigma - \Delta z) \; , \; \xi^* = 15 \; , \tag{6}$$

motivated by the idea that physically equilibration is achieved via heat transport to and from the wall. Typically one run is over 1.1 million integration time steps.

We also assume that the interaction between the wall atoms and the effective monomers is simply the Lennard Jones potential. For a homogeneous distribution of wall atoms in the half space $z < 0$, a simple integration then yields

$$U^{wall}(z) = 4\pi\epsilon_w \left(\frac{1}{45 z^9} - \frac{1}{6 z^3} \right) ; \tag{7}$$

putting the range parameter $\sigma = 1$ for the Lennard Jones potential as well. Similarly, in the case of a droplet on a fiber we consider a cylinder of radius R and ask for the potential at a distance D from the cylinder surface. Performing the integration over the z-coordinate along the cylinder axis one obtains a double integral,

$$U^{cyl}(R, D) = 8\pi\epsilon_w \int_0^{2\pi} d\phi \int_0^R \rho d\rho \left\{ \frac{63}{512[(D+R)^2 + \rho^2 - 2\rho(D+R)\cos\phi]^{11/2}} \right.$$
$$\left. - \; \frac{3}{16[(D+R)^2 + \rho^2 - 2\rho(D+R)\cos\phi]^{5/2}} \right\} \; . \tag{8}$$

While no such simple formula as for the planar surface applies here, (8) can be evaluated numerically without difficulty. One thus sees that the adsorbing potential acting on the effective monomers is not only a function of the distance D from the cylinder surface, but depends very strongly on the cylinder radius as well.

3 Nanodroplets on Flat Structureless Substrates

In equilibrium, a laterally homogeneous film covering a flat perfect substrate decreases its free energy by breaking up into droplets in a situation of incomplete wetting [13, 14] when the so called *spreading coefficient* S becomes negative, $S = \gamma_{wv} - \gamma_{w\ell} - \gamma_{\ell v} < 0$ and (neglecting line tension effects [14]) the contact angle θ of a (macroscopic) droplet that has formed is given by the classical Young equation

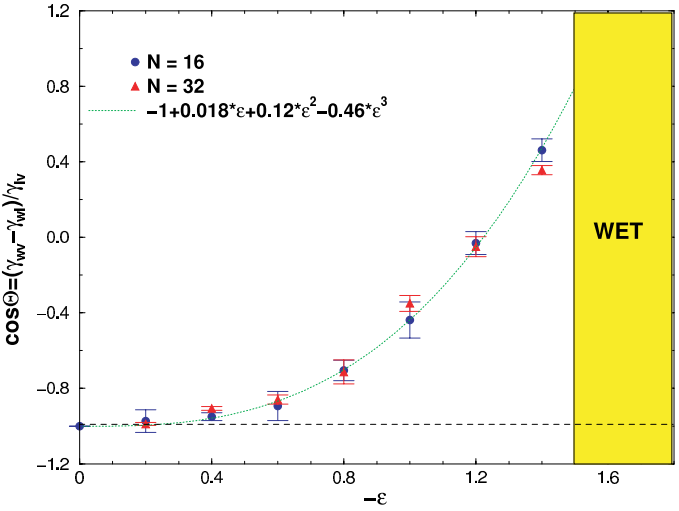


Fig. 1. Estimated values of $\cos\theta$ for $N = 16$ (*dots*) and $N = 32$ (*triangles*) plotted vs. $-\epsilon_w$ and $T = 0.49$. *Broken horizontal straight line* shows the minimum value $\cos\theta = -1$, full curve is the best fit

$$\cos\theta = (\gamma_{wv} - \gamma_{w\ell})\,/\gamma_{\ell v}\ . \tag{9}$$

In order to have nanoscopically small droplets at a planar surface perfectly metastable for a very long time, it is essential to consider a nonvolatile liquid, such as a polymer melt (a droplet of small molecule fluid could evaporate very fast, and this would also hamper a corresponding simulation, while for reasonable choices of parameters the evaporation of a polymer chain from an adsorbed droplet on the accessible time scales is never observed [15]).

In Fig. 1 we demonstrate [15] that estimates for $\cos\theta$ can be obtained from the analysis of droplet density profiles for the full range of adsorption strength $-0.2 \le \epsilon_w \le -1.4$ with reasonable accuracy. It is also seen from Fig. 1 that the variation of $\cos\theta$ with ϵ_w near $\cos\theta = 1$ is clearly linear and reaches the value $\cos\theta = 1$ at $\epsilon_w \approx 1.5$ with a finite and nonzero slope: as is well known, this is the signature of a first-order wetting transition.

If we wish to test the Young equation, (9), using the estimated value of the contact angle, then the interfacial energy $\gamma_{w\ell}$ at the wall {note that in our case $\gamma_{wv} = 0$ since the density of the vapor phase is zero to a very good approximation}, and the interfacial tension γ_{lv} at the vapor phase boundary need to be known. We outline below how they are estimated within the framework of the same model.

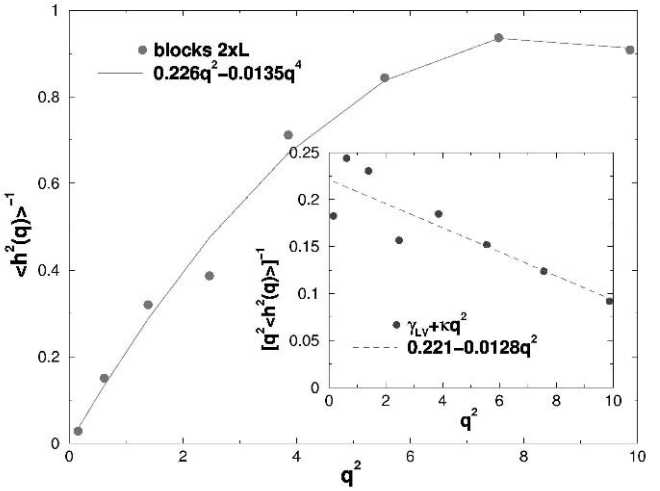


Fig. 2. Plot of $\langle h^2(q)\rangle^{-1}$ vs. q^2 where a fit through the origin, $0.226q^2 - 0.0135q^4$, is shown by full line. The inset shows a plot of $[q^2\langle h^2(q)\rangle]^{-1}$ vs. q^2 yielding $\gamma_{lv} \approx 0.22 \pm 0.02$ and $\kappa \approx -0.0128 \pm 0.003$

3.1 Estimation of the Surface Tension γ_{lv} from Capillary Wave Analysis

If the interface is treated on a coarse-grained level, where one disregards the detailed density profile from the liquid to the gas and is only interested in the local position $h(x, y)$ of the interface above the substrate, the effective interface Hamiltonian can be written as $\mathcal{H}_{cw}\{h\} = \int dxdy\{\gamma_{lv}[1 + \frac{1}{2}(\nabla h(x, y))^2] + \frac{1}{2}\kappa[\nabla^2 h(x, y)]^2 + V_{wall}(h)\}$ with γ_{lv}-cost of a flat surface, $\frac{1}{2}\gamma_{lv}(\nabla h)^2$-cost of interfacial fluctuations, κ-bending rigidity, $V_{wall}(h)$ status - substrate potential. Assuming $\bar{h} \gg 0$ in our case, cf. Fig. 3, $V_{wall}(h)$ is not varying much on the scale of $\bar{h} \pm \nabla h$, and thus we may take $V_{wall}(h) \approx V_{wall}(\bar{h})$ on the length scales of interest. One obtains γ_{lv} from $\left\langle |h(\boldsymbol{q})|^2 \right\rangle = 1/\left[\gamma_{lv}q^2 + \kappa q^4\right]$, $h(\boldsymbol{q})$ being the Fourier component of $h(x, y)$, cf. Fig. 2. Evidently, the spectrum analysis of the interface fluctuations yields reasonable values for γ_{lv} but one should be aware that for $\boldsymbol{q} \to 0$ these fluctuations take a very long time to develop [15] and one should watch carefully that equilibrium actually has been reached.

3.2 Estimation of γ_{wl} from the Local Pressure Tensor Anisotropy

The pressure p follows from the virial theorem as $p = k_B T\rho + \frac{1}{3V}\langle\sum_{i<j} \boldsymbol{f}(\boldsymbol{r}_i - \boldsymbol{r}_j) \cdot (\boldsymbol{r}_i - \boldsymbol{r}_j)\rangle_T$, where ρ denotes density, and V – the volume. In a thin film the local pressure tensor can be written as a function of the distance z from the wall as

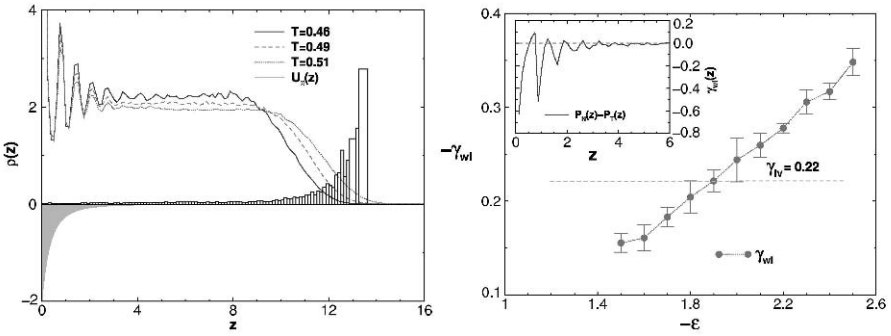


Fig. 3. (*Left*): Density profiles $\rho(z)$ for three temperatures as indicated. The wall potential is also indicated (*shaded region*). The columns show the particle mobility at $T = 0.51$ (multiplied by 10^2). (*Right*): The surface tension γ_{lw} between the melt and the wall vs. adhesion strength ϵ_w at $T = 0.51$. The interfacial tension γ_{lv} between melt and vapor, obtained from the capillary wave analysis, is shown by a *dashed line*. The inset shows the anisotropy of the local pressure tensor $P_N(z) - P_T(z)$

$$p_{\alpha\beta}(z) = \rho(z)k_BT\delta_{\alpha\beta} - \frac{1}{2A}\sum_{i\neq j}\frac{(\boldsymbol{r}_{ij})_\alpha}{|z_{ij}|}\left(\frac{\partial U(\boldsymbol{r}_{ij})}{\partial(\boldsymbol{r}_{ij})_\beta}\right)\theta\left[\frac{z-z_i}{z_{ij}}\right]\theta\left[\frac{z_j-z}{z_{ij}}\right], \tag{10}$$

where A is now the surface area of the wall in the simulated system, and θ is the Heaviside step function. The parallel and normal components of the tensor, $p_T(z) = \rho(z)k_BT - \frac{1}{4A}\langle\sum_{i\neq j}\frac{x_{ij}^2+y_{ij}^2}{r_{ij}}U'(r_{ij})\delta(z-z_i)\rangle$, and $p_N(z) = \rho(z)k_BT - \frac{1}{2A}\langle\sum_{i\neq j}\frac{z_{ij}^2}{r_{ij}}U'(r_{ij})\delta(z-z_i)\rangle - \rho(z)z\frac{d}{dz}V_{wall}(z)$ are then used to compute the surface excess free energy due to the wall: $\gamma_{\ell w} = \int_0^{z_{cutoff}} dz[P_N(z) - P_T(z)]$.

Indeed, as Fig. 3 (right) shows, the wetting transition should be located according to (9) at $\gamma_{lv} = \gamma_{wv} - \gamma_{wl} = -\gamma_{wl}$ (noting that $\gamma_{wv} = 0$ for polymers) yielding $\epsilon_w \approx -1.9$. The discrepancy with Fig. 1 where the wetting transition occurs at $\epsilon_w \approx -1.5$ is due to line tension corrections which are not taken into account by the Young equation but become significant for very small droplets. Notwithstanding, our simulational results thus suggest that the concepts developed for mesoscopic length scales (the contact angle, and its use in the Young equation, the capillary wave Hamiltonian, etc.) still work on this nano scale.

4 Nanodroplets Adsorbed on Nanofibers

With the discovery of nanotubes, it has become possible to create also small cylindrical wires from a variety of materials, and such cylindrical wires or fibers of nanoscopic size are of interest for a variety of applications [16]. The wetting properties of these nanocylinders when brought in contact with

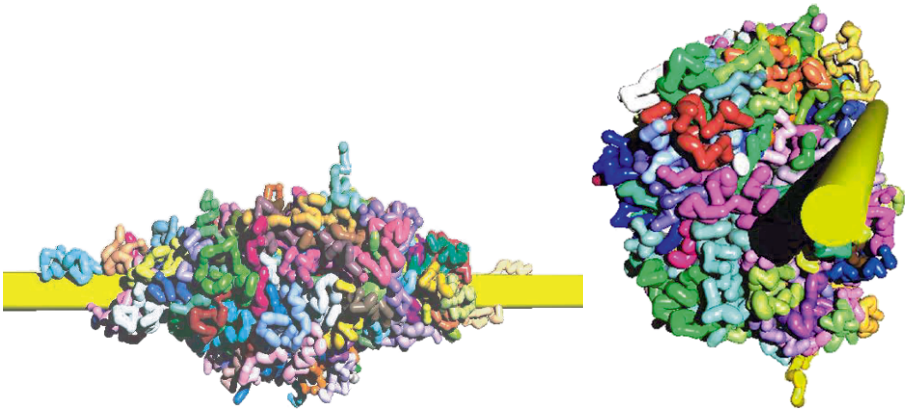


Fig. 4. (*Left*): Snapshot picture of a droplet composed of $M = 256$ flexible polymers with chain length $N = 32$ effective monomers (total number of monomers $N_{mon} = 8192$), adsorbed on a cylinder of radius $R = 0.35$. Here $T = 0.49$ and $\epsilon_w = 1.0$ in a "barrel" shape; (*Right*): The same for $R = 0.1$, illustrating the case of a "clamshell" configuration where the center of mass of the polymer droplet is far outside the cylinder

various liquids are clearly important [17]. While the wetting properties of fibers with diameters in the μm range (or larger) have been extensively explored, nanometer-size liquid droplets adsorbed on cylinders have only been considered recently [17] albeit the considered tube radii were still large on the nanometer scale. Therefore we have used the present model of polymer melt droplets to study their adsorption on tubes [18] with radius of several nm for different combinations of droplet volume, adhesion strength to the fiber surface ϵ_w, and the cylinder radius R. The interplay between surface and line tension in this case is especially delicate since, according to (8), γ_{lw} grows as R is increased.

For droplets and cylinders in the micrometer range, it is known that two topologically distinct droplet shapes occur, depending on the droplet volume, the contact angle, and the cylinder radius [19]: one shape is the "barrel" morphology (Fig. 4-lhs), i.e. the droplet surrounds the cylinder symmetrically, shaped like a barrel. The other shape is called "clamshell", i.e. the droplet sticks on one side of the cylinder only, and exhibits a shape comparable with the shell of a mussel – this case is expected for large enough R, as a natural description of the crossover toward the "sphere-cap" shape of a droplet on a flat wall – Fig. 4-rhs. While for droplets in the micrometer range there seems to be a rather well-defined transition between both morphologies, the so-called "roll-up instability" [19], both finite-size effects and the quasi-one-dimensional system of a thin fiber covered by a liquid film which one deals with here are expected to provide a rounding of the transition.

The simulation results demonstrate several important effects. If one extracts from the iso-density profiles the contact angle Θ with which the nanodroplets touch the cylinder, a general trend indicating that $\cos\Theta \to -1$ with growing volume of the drop is observed. Thus the droplet dewets the fiber because when the total number of monomers N_{mon} increases, the surface free energy of the droplet increases proportional to $N_{mon}^{2/3}$, while the binding energy to the fiber increases only as $N_{mon}^{1/3}$, and the contribution due to the line tension even stays constant. A similar dewetting effect is observed when the cylinder radius decreases, albeit one should keep in mind that not only the contact surface between drop and fiber is then diminished but also the potential (8) becomes thereby weaker. We also find a new realization of the roll-up instability, mentioned above: besides instability of the barrel shape which takes place for growing cylinder radius R, we find an additional one for extremely small radii of the nanofibers: so little energy is won by the droplet-nanowire interaction, that it becomes favorable for the droplet to "expel" the nanowire from the interior to the surface region, and thus acquire a slightly smaller total surface area.

5 Spreading of Droplets on Horizontal Surfaces

The spreading of liquid droplets on solid substrates is a longstanding problem important for applications such as lubrication, adhesion, painting, etc., and still poses a number of challenges for a theoretical understanding [13,20]. Experiments have revealed universal laws such as Tanner's law [21,22] of spreading with time t after the onset of complete wetting (the droplet lateral radius $R_{lat}(t) \propto t^n$ with $n = 1/10$ while the corresponding Young's angle vanishes $\Theta \propto t^{-m}$ with $m = 0.3$), as well as fascinating phenomena on the atomistic scale (a monolayer precursor film advancing in front of the drop, two different regimes for the growth of the radius with time, stratified profiles, etc.) For a better understanding of how these macroscopic hydrodynamic phenomena and the atomistic observations are precisely related, computer simulation studies prove valuable [23,24].

The dynamics of wetting can be determined by the variation of the contact angle with time or by measuring the lateral wetting velocity. The driving force of the process of droplet spreading on a solid surface is the macroscopic, unbalanced Young's force $F_d = \gamma_{wl} - (\gamma_{wv} + \gamma_{lv}\cos\theta) \approx S + \frac{1}{2}\gamma_{lv}\theta^2$, $S \gg \frac{1}{2}\gamma_{lv}\theta^2$, where S is the spreading parameter, which measures the energy difference between the bare substrate (γ_{wv}) and the substrate covered with a liquid film ($\gamma_{wl} + \gamma_{lv}$) (if $S < 0$ one has to find an angle θ which satisfies the condition $F_d = 0$ which is nothing but the familiar Young's equation for the contact angle of a sessile drop in equilibrium). One could expect the spreading velocity to increase with spreading parameter S, but in fact, if we consider the spreading of silicon oil on a clean glass surface (where S is large) or on a silanised glass (where S is practically zero), the spreading velocity is exactly

the same [20]! The effect of S is more subtle and passes unseen to the naked eye. It has been shown [13, 20] that for understanding of the dynamic wetting, one should consider several types of dissipation. Two of them are most important – a substantial frictional force of the precursor film F_{film}, caused by viscous fluid flow on the surface of the substrate (this force exactly balances [13] S) and a significantly smaller frictional force on the fluid wedge F_{wedge} which balances the smaller contribution of the driving force ($\frac{1}{2}\gamma_{lv}\theta^2$).

We study the time evolution of $R_{lat}(t)$ and $\Theta(t)$ by means of both an MC [23] and MD [24] simulation of polymer melt nanodroplets. The integration over time in our MD simulation is made by velocity-Verlet algorithm and the temperature is kept constant via Nosé-Hoover thermostat. The coordinates of each monomer $\boldsymbol{r}_i$ change according to $m\ddot{\boldsymbol{r}}_i = \boldsymbol{F}_i - \zeta\dot{\boldsymbol{r}}_i$, where the friction parameter ζ is determined by $\frac{d\zeta}{dt} = \left(\sum_i v_i^2 - 3Nk_BT\right)/Q$, with $\boldsymbol{v}_i$ being the velocities of all monomers. The thermal inertia parameter $Q = 300$. $\boldsymbol{F}_i = \sum_{j\neq i} dU_{ij}/d\boldsymbol{r}_{ij}$ is the sum of the all forces acting in the system. U_{ij} contains the bonded, nonbonded and wall potentials. Figure 5 shows some typical results of such a simulation run. The left columns presents contour diagrams for $\epsilon_w = 0.7$ at different times. In order to quantify the spreading dynamics of the droplet we have taken spherical averages around the center of mass of the droplet. Thus we obtain a two-dimensional density profile $\rho(Z, R(X, Y))$ which is depicted in the form of contour diagrams. For taking mean values 128 parallel runs are averaged over. One can see that there exists a well defined precursor film which is also seen from the corresponding snapshots on the (X, Y) plane (right columns of Fig. 5). It is seen from them that after some time (such as $t \approx 470$ t.u. for $\epsilon_w = 0.4$) some of the chains split off and the droplet starts to disintegrate. On Fig. 6 we present log-log plots of the lateral radius $R_{lat}(t)$ and the contact angle $\theta(t)$ versus time t for two different strengths of the adsorption potential – $\epsilon_w = 0.4$ and $\epsilon_w = 0.7$. Time $t \approx 0.1$ t.u. is necessary for the droplet to start to spread after switching the stronger adhesion. After that time Tanner's law $R_{lat}(t) \propto t^{0.1}$ accounts rather well for the data until $t_{0.4} \approx 130$ t.u. (for $\epsilon_w = 0.4$), and $t_{0.7} \approx 15.5$ t.u. (for $\epsilon_w = 0.7$), when some of the chains split off and start diffusing individually on the substrate. The power law for the Young's angle $\theta(t) \propto t^{-0.3}$ sets on at time $t \approx 1.3$ t.u. for the two values of ϵ_w. The large deviation of our θ-data from this power law we explain with the way we use to extract $\theta(t)$. (We fit a straight line to the isodensity line for density $\rho = 1$ on contour diagrams like these shown on Fig. 5 and then we estimate $\theta(t)$). The power law of the lateral radius change during the latest times ($t_{0.4} > 130$ t.u.) is $R_{lat}(t) \propto t^{0.24}$ and $R_{lat}(t) \propto t^{0.32}$ (for $t_{0.7} > 15.5$ t.u.). This is an intermediate regime of the kinetics which is changed from capillary driven, where $R_{lat}(t) \propto t^{0.1}$, to diffusive kinetics [23] , where $R_{lat}(t) \propto t^{0.5}$.

Comparing the MD data in Fig. 6 with a previous study [23], where a kinetic MC algorithm was applied with the *same* interactions used, we note that in both cases the spreading dynamics practically coincides which is highly nontrivial since the MC algorithm allows for purely diffusive motions only

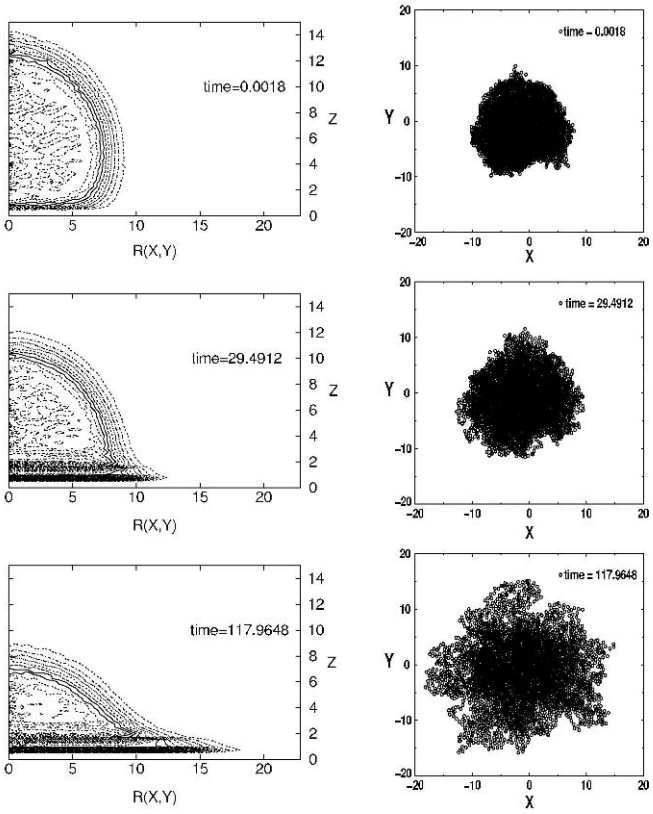


Fig. 5. Contour diagram representing the density profile $\rho(Z, R(X,Y)) = n\Delta\rho, n = 1, 2...18$, and $\Delta\rho = 0.2$ in the (Z, R) plane (*left column*) The snapshots show the configurations by projecting the coordinates (X, Y) of each monomer on the surface plane $Z = 0$ (*right column*), of a droplet containing 128 chains with 32 monomers each at $k_BT = 0.49$. The drop starts spreading upon switching the substrate potential at time $t = 0$ time units (t.u.), 1 t.u. $\approx$ 1111 MD steps, from $\epsilon_w = 0.1$ to $\epsilon_w = 0.7$

whereas in MD one also accounts for inertial effects. One may thus conclude that friction plays a decisive role in droplet spreading. Moreover, if one compares the mean squared displacements of the centers of mass of the polymers, $\langle[\boldsymbol{r}_{cm}(t) - \boldsymbol{r}_{cm}(0)]^2\rangle$, one may even "calibrate" a Monte Carlo Step in terms of the physical time units used in the MD simulation! For our coarse-grained model simulation one gets then 1 MCS $\approx$ 1.5 fs.

In summary, our MD study of a droplet spreading shows that the spreading kinetics of a nanoscopic size polymer droplet on a structureless substrate follows Tanner's law as long as the droplet does not disintegrate into individually moving polymer chains and this is independent of the spreading coefficient S (the larger S values lead to a wider precursor film).

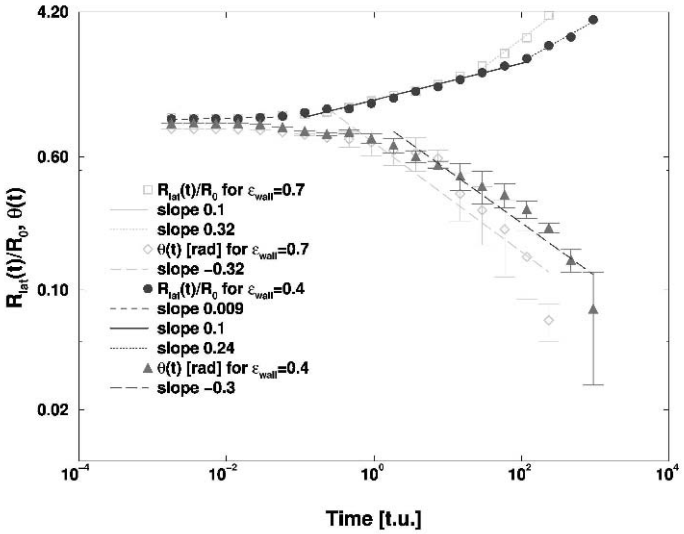


Fig. 6. Log-Log plot of the reduced lateral radius R_{lat}/R_0 (circles - for a spreading droplet at $\epsilon_w = 0.4$ and squares – for a droplet at $\epsilon_w = 0.7$) and the contact angle $\theta(t)$ (triangles – for a spreading droplet at $\epsilon_w = 0.4$ and diamonds – for a droplet at $\epsilon_w = 0.7$) vs. time t for droplets containing $N_{ch} = 128$ chains with $M = 32$ monomers each, which spread after a change of the adsorption potential strength from $\epsilon_w = 0.1$ (where the initial radius $R_0 = 4.82$) to $\epsilon_w = 0.4$ or $\epsilon_w = 0.7$ at $t = 0$

6 Liquid Bridges and Induced Forces on Chemically Decorated Droplets with Varying Wettability

For patterned surfaces in the micrometer range, on which liquid droplets or thin liquid films are adsorbed, fascinating wetting morphologies have been predicted [25–32] (including "morphological wetting transitions" [25, 29, 30]) and observed [7]. Liquid bridges, for instance, which may span opposite lyophilic domains in a narrow slit are also of interest in the context of forces between droplets (or bubbles, respectively) [33, 34], long range forces between colloidal particles [35, 36], etc. The theoretical treatments mostly apply phenomenological quasi-macroscopic concepts (in terms of interfacial tensions, contact angles, [38, 39] etc.); but for droplets on the micrometer scale the lack of knowledge on the line tension [13, 39–49] is a serious drawback already, and the understanding of the density distribution of the droplet near the contact line is a difficult problem [47, 48]. The present drive towards nanoscale technology creates also more interest in droplets on the nanometer scale, and in fact some of the techniques mentioned above are well suited to create surface patterns on the nanoscale [11, 12].

Although sometimes macroscopic concepts do allow reasonable predictions down to the nanoscale [15, 23, 24, 50], there is no guarantee that the phenomenological theories are quantitatively accurate for bridge formation between polymer nanodroplets. We study this problem by Molecular Dynamics simulation so as to complement the existing knowledge on the problem by insight on a molecular level.

In our computer [51] experiments we create a circular or rectangular lyophilic domain on the adsorbing substrate whereby the constant ϵ_w, describing the wall potential $V_{wall}(z)$ in (7), has a value $\epsilon_w = 0.2$ for the radius $0 < R < R_D$ and a smaller value (an otherwise lyophobic surface) $\epsilon_w = 0.05$ outside the circle (stripe).

The variation of the droplet iso-density profiles and the contact angle θ with the width R_D of the lyophilic domain is shown in Fig. 7. It is evident from Fig. 7 that the shape of the adsorbed droplet changes when the radius R_D decreases. As expected, the contact angle θ varies continuously with R_D, due

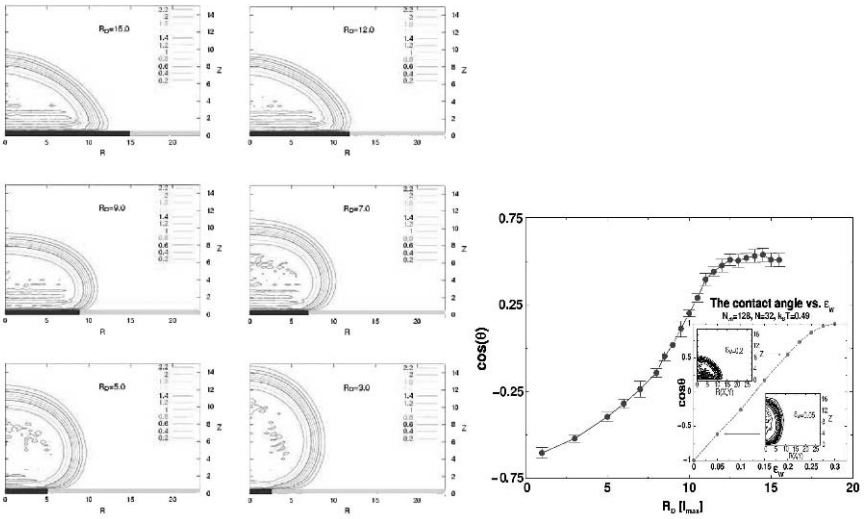


Fig. 7. (*Left*): Contour diagrams representing the density profiles $\rho(Z, R(X, Y)) = n\Delta\rho$, $n = 1, 2, \ldots 11$, $\Delta\rho = 0.2$, in the (Z, R) plane, for a droplet containing 128 chains with 32 monomers each, at a temperature $k_BT = 0.49$. The droplet is in contact with an ideally flat lyophobic substrate (represented by the *thick grey line* in the bottom) decorated with one lyophilic circle of radius R_D (the *thick black line* in the bottom). Profiles are shown for $R_D = 15, 12, 9, 7, 5$ and 3, respectively. These profiles are obtained by averaging over 10 runs of $2.1 \cdot 10^6$ integration steps each. One MD time step is $\delta t = 0.0009$ MD time units. (*Right*): Cosine of the contact angle θ plotted versus the radius R_D of the lyophilic domain. The inset shows the dependence of the contact angle of the droplet at an infinitely extended homogeneous substrate. The density profiles correspond to $\epsilon_{\rm w} = 0.05$ (lyophobic) and $\epsilon_{\rm w} = 0.2$ (lyophilic) substrate

to the interaction of the contact line and the boundary between the lyophilic and lyophobic regions at the substrate, and is no longer identical to the contact angle that applies for an infinitely extended flat lyophilic substrate.

If chemically decorated lyophilic domains are placed opposite to one another in a slit pore of width L (the distance between the parallel walls), the droplets sitting on these domains may coalesce at sufficiently small L so that a liquid "bridge" between the walls is formed. If the droplets forming on such domain were large enough, one could then write, ignoring the line tension, the Helmholtz free energy as,

$$H^{bridge} = \gamma_{lv} A_{lv} - (\gamma_{vw} - \gamma_{lw}) A_{lw} + \Delta p V \tag{11}$$

where A_{lv} and A_{lw} denote the contact areas liquid-vapor and liquid-wall, $\Delta p = p_v - p_l$ is the pressure difference across the liquid-vapor interface, and the liquid is confined in volume V. One can minimize (11) with respect to the shape of the liquid – vapor surface and is then lead to the Laplace equation, $\Delta p = -2\gamma_{lv}/\mathcal{R}$, with $\mathcal{R}$ – the mean curvature of the $l - v$ – interface, implying that for fixed pressure difference one will get an optimal bridge shape with constant curvature. Depending on the sign of Δp, the bridge can bulge both outward ($\Delta p < 0$) as well as inward ($\Delta p > 0$). Further minimization with respect to the configuration of the three-phase $l - v - w$ contact line yields then the generalized Young equation, $\gamma_{vw}(\boldsymbol{r}) = \gamma_{lw}(\boldsymbol{r}) + \gamma_{lv} \cos[\theta(\boldsymbol{r})]$ where the contact angle $\theta(\boldsymbol{r})$ varies with position $\boldsymbol{r}$ in the surface plane. In the simplest case when the centers of the two opposite completely wettable ($\theta_{phil} = 0$) domains (stripes of length $L_\parallel$) of the same width $2R_D$ are not mutually shifted and the substrate outside the domains is nonwettable ($\theta_{phob} = \pi$), one finds [27] $H^{bridge}/(\gamma_{lv} L_\parallel) = 4\mathcal{R} \arcsin(\frac{L}{2\mathcal{R}}) - 4R_D$ in the case of constant volume of the bridge. Then one can get a simple expression for the force $\mathcal{F}$ exerted by the bridge which spans the two parallel walls of the slit, $\mathcal{F} = -\partial H^{bridge}/\partial L = -2\sin(\theta) - 4R_D \cos(\theta)/L$ (θ is the contact angle in the bottom left-hand corner of the bridge – see Fig. 8). Thus there exists a characteristic angle $\theta^* = -\arctan^{-1}(2R_D/L) > \pi/2$ for which $\mathcal{F}$ changes sign. The variation of this force with distance L - Fig. 9 - is of great interest since its measurement in real experiments has proved to be difficult [34,35]. Indeed, from Fig. 9 it is evident that when the distance L is large enough, $L \geq 52$ in our case, the bridge is no longer stable and breaks into two channels each covering the entire lyophilic stripe. At shorter separations the bridge bulges inward and the normal force $\mathcal{F}$ is attractive, displaying a local minimum at $L = 11$. When the plates get closer $\mathcal{F}$ changes sign at $L \approx 8$ and becomes repulsive (the bridge bulges outward) at even closer separation between the walls.

It should be clear that the force seen in Fig. 9 is entirely due to the interplay of the various interfacial interactions that control the shape of the liquid bridge. One should point out that both the course of the bridging force against the inter-plate distance L as well as the particular bridge configurations corresponding to various parts of the force-distance curve closely match some

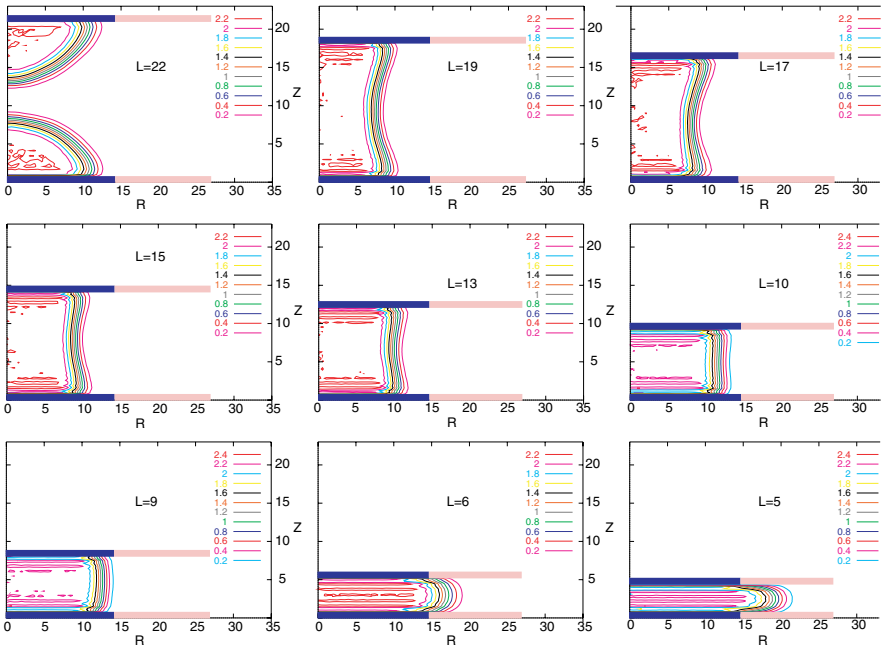


Fig. 8. Contour diagrams representing the density profiles $\rho(Z, R(X, Y)) = n\Delta\rho$, $n = 1, 2, \ldots, 11(\Delta\rho = 0.2)$ in the (Z, R) plane for two droplets adsorbed on opposite walls ($L = 22$) or a liquid bridge between the decorated walls (marked by *dark bars*) at $(L = 19, 17, 15, 13, 10, 9, 6, 5)$ respectively

recent results [32] for macroscopic structureless droplets obtained by different surface minimization techniques.

7 Film Rupture

The kinetics of break-up of a thin polymer film which forms a free-standing liquid bridge between two lyophilic stripes opposite to one another on an otherwise lyophobic substrates can be visualized by a series of snapshot pictures [37], as shown in Fig. 10. Since the two gas-liquid interfaces of the thin film are slightly concave, the thickness of the film is smallest at mid-distance $L/2$ i.e. $z = 0$. Due to thermal capillary wave-type fluctuations of the interfaces, at some spot in this region the barrier against nucleation of a hole in the film is so small that a hole is formed almost immediately after the system was prepared in the considered state. Figure 10 shows that this hole starts out with a cross section close to circular but rather soon develops an elliptical shape: obviously hole growth is easier in the y-direction rather than close to the walls with the lyophilic stripes where the liquid meniscus makes the film thicker. Actually at times $t \geq 1500$ almost all monomers gather in elongated

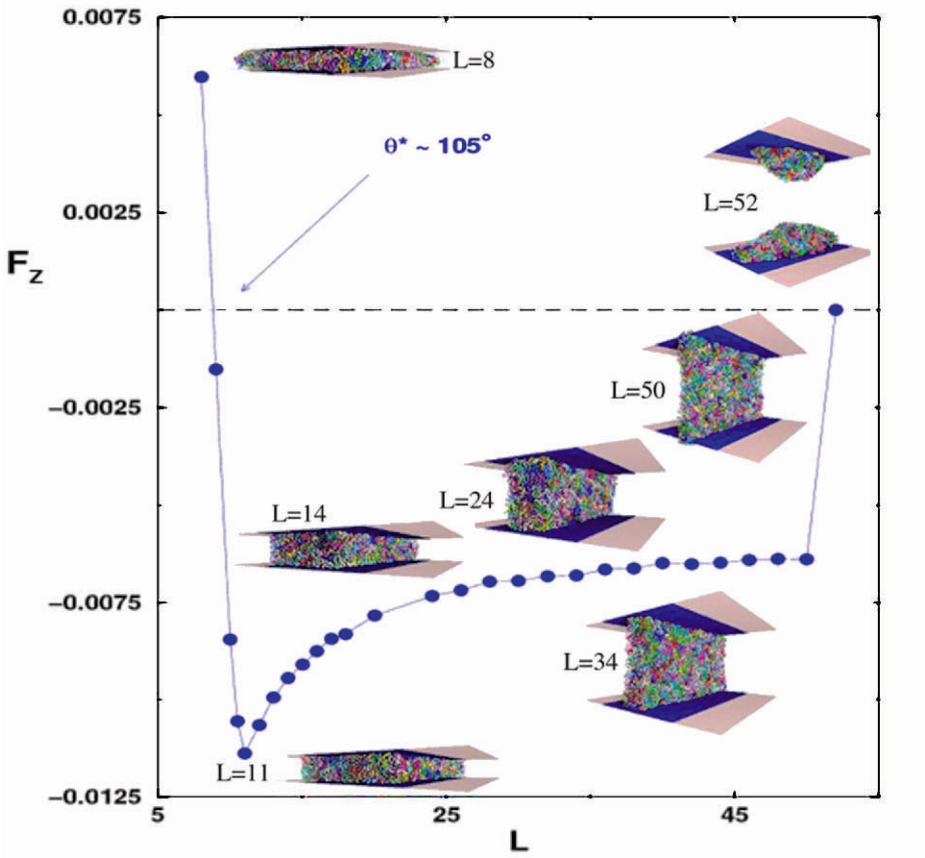


Fig. 9. The normal force acting between two parallel walls bridged by a liquid plotted vs distance L. The substrates are lyophobic and decorated by rectangular lyophilic stripes. In the insets are shown snapshots of a film bridging two substrates at different values of L . The *dark* areas represent rectangular lyophilic parts of the substrates. The contact angle at zero force is $\theta^* \approx 105^\circ$

domains attached to the two lyophilic stripes, and all what is left from the liquid film spanning the whole distance $L = 52$ between the two stripes is a rounded liquid bridge column, which ruptures in the example shown in Fig. 10 soon after the last snapshot shown.

Figure 11 presents the time evolution of the hole diameter of the rupturing film on a log-log plot. The linear growth of the hole diameter with time, $D_y \propto t$, that is observed until $t \approx 300$ MD units, differs markedly from the exponential growth of hole sizes with time in freely suspended thin viscous [52] films and from the growth laws discussed for spinodal dewetting of supported thin films [53]. Of course, the free-standing polymer films whose rupture is considered in the literature [52] have macroscopic rather than nanoscopic

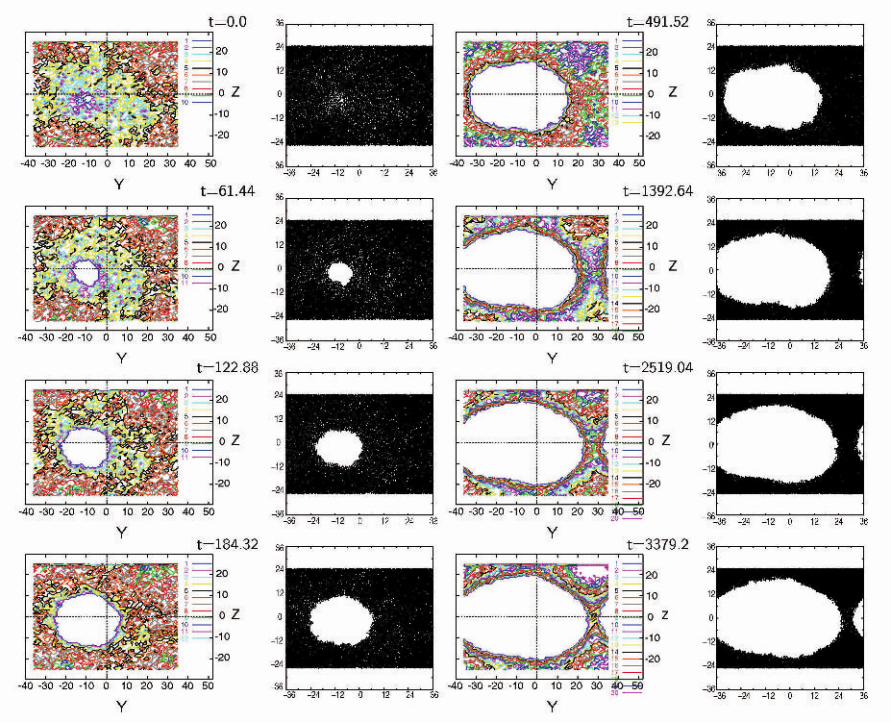


Fig. 10. Snapshot pictures of the time evolution (time is measured in the standard MD time units) of a thin film between two lyophilic stripes at distance $L = 52$ in the z-direction, while along the film periodic boundary conditions are used. Left series of pictures shows contours of the local accumulative particle density. Right series of pictures were generated simply by projecting the y, z coordinates of each monomer in the film onto one yz-plane. The early time evolution, namely times $t = 0.0, 61.44, 122.88$ and 184.32, and the late time evolution ($t = 491.52, 1392.64, 2519.04$ and 3379.2, are shown respectively)

lateral dimensions. During the stage where the hole in such films grows exponentially with time the thickness of the remaining film stays constant everywhere. However, in our case the film thickness W is distinctly nonuniform, due to the concave shape of the cross section through the film. In fact, we observe that the film thickness is about $W \approx 10$ near the walls but only about $W \approx 5$ near $z = 0$. Studying the flow of material in a viscous liquid film with such inhomogeneous linear dimensions from a hole to the walls is a complicated problem of hydrodynamics that we are not going to address here. Similarly, in the late stages of hole growth (region II in Fig. 11) mass has to be transported from the center of the liquid bridge (at $z = 0$) to the regions near the walls, and there the excess mass contained in the conical "foots" of the liquid bridge has to spread sideways into the positive or negative y-direction along

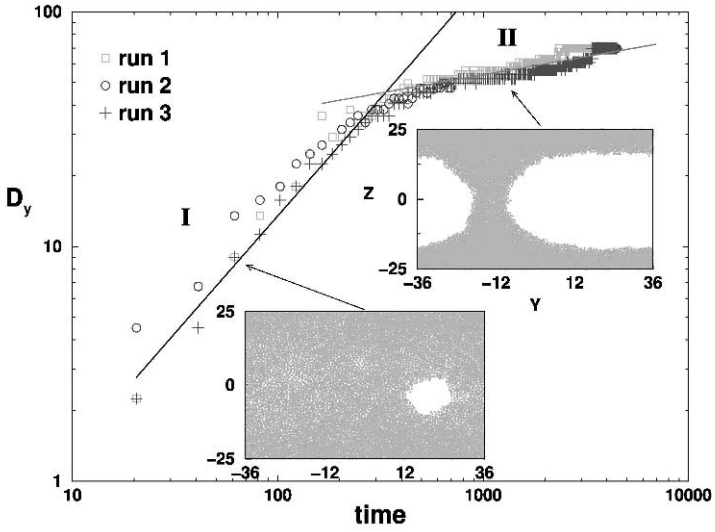


Fig. 11. Hole diameter D_y plotted vs. time, for the same parameters as shown in Fig. 10. Three equivalent runs are included. Straight lines on the log log plot illustrate power laws $D_y \propto t$ in regime I and $D_y \propto t^{0.15}$ in regime II, respectively

the elongated liquid channel. This spreading of liquid material is somewhat reminiscent of the spreading of liquid droplets on wet substrates [23, 24].

8 Conclusions and Outlook

We have presented here a feasibility study of simulating the equilibrium properties of polymer droplets adsorbed on a wall or a fiber under conditions of incomplete wetting and demonstrated that classical concepts still work on nanoscopic scale.

The spreading kinetics of a polymer droplet on a structureless substrate follows Tanner's law as long as the droplet does not disintegrate into individually moving polymer chains and thus is independent of S, while the larger S values lead to a wider precursor film.

Motivated by the work of Lipowsky et al. [27–31], we have studied in the present work nanoscopic fluid droplets both on isolated lyophilic stripes on otherwise lyophobic substrates, and ultrathin polymer films held between two such stripes on surfaces a small distance L apart. Unlike the macroscopic theory [27–31] which neglects statistical fluctuations due to its mean-field type character, but which also disregards line tension contributions as well as any effects on the atomistic scale (e.g. due to the packing of the molecules at a flat substrate, etc.), our MD simulations of a coarse-grained model of short flexible polymer chains in principle include all such effects and hence

can provide check to what extent the macroscopic theory still can account for such phenomena at the nanoscale.

Of course, the present modeling can be viewed as a first step only, and many interesting aspects have been disregarded. Thus, our coarse-grained polymer model disregards effects of chain stiffness, and even on a coarse-grained level is suitable for neutral polymers only. Also the substrate was strongly idealized (perfectly flat and ideal, no atomistic corrugation or mesoscopic roughness, etc.). We also emphasize that the choice made to consider very short chains ($N = 20$) excludes any effects specific for the dynamics of long polymers (such as reptation) which may effect the macroscopic growth laws as well. We do hope that the present work will stimulate both further theoretical and experimental studies of these interesting problems.

Acknowledgments

The author is indebted to J. Yaneva and K. Binder for the close cooperation in deriving most of the results quoted in the present chapter. Our studies have been partially supported by the NANOPHEN Project, Contract No. INCO-CT-2005-016696, of the EU, and by the Deutsche Forschungsgemeinschaft (DFG) under grant number 436 BUL 113/130.

References

1. N. L. Abbott, J. P. Folkers, and G. M. Whitesides (1992) Manipulation of the Wettability of Surfaces on the 0.1- to 1-Micrometer Scale Through Micromachining and Molecular Self-Assembly. *Science* **257**, p. 1380
2. R. Wang et al. (1997) Light-induced amphiphilic surfaces. *Nature* **388**, p. 431
3. G. Möller, M. Hake, and H. Motschmann (1998) Controlling Microdroplet Formation by Light. *Langmuir* **14**, p. 4955
4. G. P. Lopez, H. A. Biebuyck, C. D. Frisbie, and G. M. Whitesides (1993) Imaging of features on surfaces by condensation figures. *Science* **260**, p. 647
5. J. Drelich, J. D. Miller, A. Kumar, and G. M. Whitesides (1994) Wetting Characteristics of Liquid Drops at Heterogeneous Surfaces. *Colloids Surf. A* **93**, p. 1
6. F. Morhard, J. Schumacher, A. Lenenbach, T. Wilhelm, R. Dahint, M. Grunze, and D. S. Everhart (1997) Optical diffraction – a new concept for rapid on-line detection of chemical and biochemical analytes. *Electrochem. Soc. Proc.* **97**, p. 1058
7. H. Gau, S. Herminghaus, P. Lenz, and R. Lipowsky (1999) Liquid Morphologies on Structured Surfaces: From Microchannels to Microchips. *Science* **283**, p. 46
8. R. Wang, A. N. Parikh, J. D. Beers, A. P. Shreve, and B. Swanson (1999) Nonequilibrium Pattern Formation in Langmuir-Phase Assisted Assembly of Alkylsiloxane Monolayers. *J. Phys. Chem. B* **103**, p. 10149
9. M. Gleiche, L. F. Chi, and H. Fuchs (2000) Nanoscopic channel lattices with controlled anisotropic wetting. *Nature* **403**, p. 173

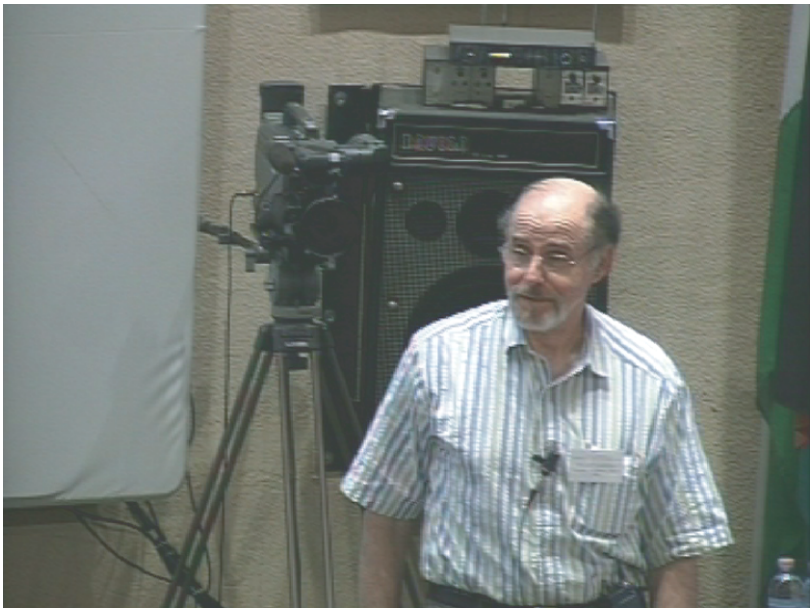
10. R. C. Hayward, D. A. Saville, and I. A. Aksay (2000) Electrophoretic assembly of colloidal crystals with optically tunable micropatterns. *Nature* **404**, p. 56
11. F. Burmeister, C. Schäfte, T. Matthews, M. Böhmisch, J. Boneberg, and P. Leiderer (1997) Colloid Monolayers as Versatile Lithographic Masks. *Langmuir* **13**, p. 2983
12. J. Heier, E. J. Kramer, S. Walheim, and G. Krausch (1997) Thin Diblock Copolymer Films on Chemically Heterogeneous Surfaces. *Macromolecules* **30**, p. 6610
13. P. G. de Gennes (1985) Wetting: statics and dynamics. *Rev. Mod. Phys.* **57**, p. 827
14. S. Dietrich (1988) In *Phase Transitions and Critical Phenomena*, vol. 12, edited by C. Domb and J. L. Lebowitz, Academic Press, London, p. 1
15. A. Milchev and K. Binder (2001) Polymer melt droplets adsorbed on a solid wall: A Monte Carlo simulation. *J. Chem. Phys.* **114**, p. 8610
16. E. Dujardin, T. W. Ebbesen, A. Krishnan, and M. M. J. Treacy (1998) Wetting of Single-Shell Nanotubes. *Adv. Mater.* **10**, p. 1472
17. C. Bauer and E. Dietrich (2000) Shapes, contact angles, and line tensions of droplets on cylinders. *Phys. Rev. E* **62**, p. 2428
18. A. Milchev and K. Binder (2002) Polymer nanodroplets adsorbed on nanocylinders: A Monte Carlo study. *J. Chem. Phys.* **117**, p. 6852
19. B. J. Carrol (1991) Spreading of liquid droplets on cylindrical surfaces: Accurate determination of contact angle. *J. Appl. Phys.* **70**, p. 493
20. F. Brochard-Wyart (1999) In *Soft Matter Phys.*, Eds. M. Daoud and C. E. Williams, Springer, Berlin, p. 1
21. L. H. Tanner (1979) The spreading of silicone oil drops on horizontal surfaces. *J. Phys. D: Appl. Phys.* **12**, p. 1473
22. B. R. Duffy and S. K. Wilson (1997) A third-order differential equation arising in thin-film flows and relevant to Tanner's Law. *Appl. Math. Lett.* **10**, p. 63
23. A. Milchev and K. Binder (2002) Droplet spreading: A Monte Carlo test of Tanner's law. *J. Chem. Phys.* **116**, p. 7691
24. J. Yaneva, A. Milchev, and K. Binder (2003) Dynamics of a spreading nanodroplet: Molecular Dynamics simulation. *Macromol. Theory Simul.* **12**, p. 573
25. P. Lenz and R. Lipowsky (1998) Morphological Transitions of Wetting Layers on Structured Surfaces. *Phys. Rev. Lett.* **80**, p. 1920
26. R. Lipowsky, P. Lenz, and P. S. Swain (2000) Wetting and Dewetting of Structured or Imprinted Surfaces. *Colloids Surf. A* **61**, p. 3
27. P. S. Swain and R. Lipowsky (2000) Wetting between structured surfaces: liquid bridges and induced forces. *Europhys. Lett.* **43**, p. 203
28. A. Valencia, M. Brinkmann, and R. Lipowsky (2001) Liquid Bridges in Chemically Structured Slit Pores. *Langmuir* **17**, p. 3390
29. R. Lipowsky (2001) Structured Surfaces and Morphological Wetting Transitions. *Interface Science* **9**, p. 105
30. R. Lipowsky (2001) Morphological wetting transitions at chemically structured surfaces. *Curr. Opin. Colloid Interface Sci.* **6**, p. 40
31. M. Brinkmann and R. Lipowsky (2002) Wetting morphologies on substrates with striped surface domains. *J. Appl. Phys.* **92**, p. 4296
32. E. J. De Souza (2004) Capillary forces between structured substrates. Thesis, Max-Planck Institut für Metallforschung, Stuttgart

33. J. L. Parker, P. M. Claesson, and P. Allard (1994) Bubbles, cavities, and the long-ranged attraction between hydrophobic surfaces. *J. Phys. Chem.* **98**, p. 8468
34. P. Attard and S. J. Miklavcic (2001) Effective Spring Constant of Bubbles and Droplets. *Langmuir* **17**, p. 8217; P. Attard (1996) Bridging Bubbles between Hydrophobic Surfaces. *Langmuir* **12**, p. 1693
35. O. I. Vinogradova, G. E. Yakubov, and H.-J. Butt (2001) Forces between polystyrene surfaces in water-electrolyte solutions: Long-range attraction of two types? *J. Chem. Phys.* **114**, p. 8124
36. D. Andrienko, P. Patricio, and O. I. Vinogradova (2004) Capillary bridging and long-range attractive forces in a mean-field approach. *J. Chem. Phys.* **121**, p. 4414
37. J. Yaneva, A. Milchev, and K. Binder (2005) Polymer droplets on substrates with striped surface domains: Molecular Dynamics simulation of equilibrium structure and liquid bridge rupture. *J. Phys. Condens. Matter* **17**, p. S4199
38. T. Young (1805) *Phil. Trans. R. Soc. London* **5**, p. 65
39. J. S. Rowlinson and B. Widom (1982) *Molecular Theory of Capillarity*, Clarendon, Oxford
40. M. V. Berry (1985) Simple fluids near rigid solids – statistical mechanics of density and contact angle. *J. Phys. A* **7**, 231
41. B. A. Pethica (1977) The contact angle equlibrium. *J. Colloid Interface Sci.* **62**, p. 567
42. G. Navascués and P. Tarazona (1981) Line tension effects in heterogeneous nucleation theory. *J. Chem. Phys.* **75**, p. 2441
43. B. V. Toshev, D. Platikanov, and A. Scheludko (1988) Line tension in three-phase equilibrium systems. *Langmuir* **4**, p. 489
44. I. Szleifer and B. Widom (1992) Surface-Tension, Line Tension, and Wetting. *Mol. Phys.* **75**, p. 925
45. C. Varea and A. Robledo (1992) Statistical mechanics of the line tension. *Physica A* **183**, p. 12
46. J. Drelich (1996) The Significance and Magnitude of the Line Tension in Three-Phase (Solid/Liquid/Fluid) Systems. *Colloids Surf. A* **116**, p. 43
47. T. Getta and S. Dietrich (1998) Line tension between fluid phases and a substrate. *Phys. Rev. E* **57**, p. 655
48. C. Bauer and S. Dietrich (1999) Quantitative study of laterally inhomogeneous wetting films. *Eur. Phys. J. B* **10**, p. 767
49. F. Bresme and N. Quirke (1998) Computer Simulation Study of the Wetting Behavior and Line Tensions of Nanometer Size Particulates at a Liquid-Vapor Interface. *Phys. Rev. Lett.* **80**, p. 3791
50. F. Varnik and K. Binder (2002) Shear viscosity of a supercooled polymer melt via nonequilibrium molecular dynamics simulations. *J. Chem. Phys.* **117**, p. 6336
51. J. Yaneva, A. Milchev, and K. Binder (2004) Polymer nano-droplets forming liquid bridges in chemically structured slit pores: a computer simulation. *J. Chem. Phys.* **121**, p. 12632
52. G. Debrégeas, P. Martin, and F. Brochard-Wyart (1995) Viscous Bursting of Suspended Films. *Phys. Rev. Lett.* **75**, p. 3886
53. A. Milchev and K. Binder (1997) Dewetting of thin polymer films adsorbed on solid substrates: A Monte Carlo simulation of the early stages. *J. Chem. Phys.* **106**, p. 1978

Monte Carlo Simulations of Compressible Ising Models: Do We Understand Them?

D.P. Landau[1], B. Dünweg[2], M. Laradji[3], F. Tavazza[1], J. Adler[4], L. Cannavaccioulo[5], and X. Zhu[1]

[1] Center for Simulational Physics, The University of Georgia, Athens, GA 30622, U.S.A.
dlandau @ hal.physast.uga.edu
[2] MPIP, Mainz, Germany
[3] Dept. of Physics, University of Memphis, Memphis, TN, U.S.A.
[4] Dept. of Physics, Technion-Israel Institute of Technology, Haifa, Israel
[5] IFF, Forschungszentrum Juelich, Juelich, Germany



David P. Landau

D.P. Landau et al.: *Monte Carlo Simulations of Compressible Ising Models: Do We Understand Them?*, Lect. Notes Phys. **703**, 127–138 (2006)
DOI 10.1007/3-540-35284-8_6
© Springer-Verlag Berlin Heidelberg 2006

1 Introduction 129
2 Model and Method 130
3 Results 132
4 Conclusions 137
References 137

Extensive Monte Carlo simulations have begun to shed light on our understanding of phase transitions and universality classes for compressible Ising models. A comprehensive analysis of a Landau-Ginsburg-Wilson hamiltonian for systems with elastic degrees of freedom resulted in the prediction that there should be four distinct cases that would have different behavior, depending upon symmetries and thermodynamic constraints. We shall provide an account of the results of careful Monte Carlo simulations for a simple compressible Ising model that can be suitably modified so as to replicate all four cases.

1 Introduction

Although the Ising model in three dimensions has not been solved exactly, there are now highly accurate numerical results for simple cubic Ising lattices. Physical systems, however, are not confined to a rigid lattice and there is thus substantial interest in understanding how the removal of the rigid lattice constraint might modify the nature of the magnetic phase transition. Theoretical studies of the compressible Ising model began decades ago, but a complete understanding of the effects of allowing elastic interactions to change near neighbor distances is still lacking. (A review of the theoretical background can be found elsewhere [1–3]). Of course, the equivalence between the Ising model and the binary alloy model adds to the general interest in the problem. One simple example is the Si/Ge alloy for which the covalent interactions give rise to strongly directional nearest neighbor bonds that dominate the behavior. (Ising spins $\sigma_i = \pm 1$ are then equivalent to Si or Ge atoms, respectively; and hence, the concentration of Ge atoms plays the role of the magnetization in the corresponding magnetic system. In the same vein, the chemical potential is equivalent to the magnetic field in the magnetic interpretation.) The appropriate model contains both two body and three body interactions so that both the bond lengths and bond angles are somewhat constrained. The particular values of the constants were determined by fitting to the properties of Si/Ge alloys, but the goal here is actually to study the behavior of a generic compressible Ising model. Of course, it remains to be seen if this type of model fully encompasses the range of physical behavior characteristic of "compressible Ising models" or if the problem is more subtle.

A comprehensive analysis of a Landau-Ginsburg-Wilson hamiltonian for systems with elastic degrees of freedom determined [2] that there were four distinct cases that would exhibit quite different behavior. These depend upon symmetries (e.g. the coupling between the elastic and magnetic degrees of freedom) as well as the thermodynamic constraints and are listed in Table 1.

Table 1. Predicted nature of the critical behavior for compressible Ising models

Type	Ferromagnet	Antiferromagnet
Constant Pressure	Mean Field-like	1st order
Constant Volume	two transitions, Mean Field-like	Fisher renormalized

2 Model and Method

These studies began with the intent of exploiting the Ising model-binary alloy equivalence and thus looking at Si/Ge alloys as a physical example. Therefore, spins were placed on a distortable diamond net with fully periodic boundary conditions. The simulation cell consisted of $L \times L \times L$ unit cells with 8 sites per unit cell. The total number of sites is thus $8L^3$. Several different hamiltonians were used, but in all cases they contained distance dependent two-body couplings and angle-dependent terms that could be represented by three-body expressions. For example, initially we used the Keating [4] interatomic potential, whose "stiffness" parameters E and A were determined by the macroscopic elastic constants of the crystal and which has been used to describe the structural properties of mixed systems [1].

$$\mathcal{H} = \mathcal{H}_{ext.} + \mathcal{H}_{chem.} + \mathcal{H}_{el.,bonds} + \mathcal{H}_{el.,angles} \tag{1}$$

with

$$\mathcal{H}_{ext.} = -\mu_A \sum_i \delta_{S_i,+1} - \mu_B \sum_i \delta_{S_i,-1} \tag{2}$$

(Kronecker's δ),where μ_A and μ_B are chemical potentials

$$\mathcal{H}_{chem.} = \sum_{i-j} \varepsilon(S_i, S_j) \ , \tag{3}$$

where the $\varepsilon(S_i, S_j)$ are chemical binding energies

$$\mathcal{H}_{el.,bonds} = \sum_{i-j} E(S_i, S_j) \left({\boldsymbol{r}_{ij}}^2 - R_0^2(S_i, S_j) \right)^2 \tag{4}$$

where $E(S_i, S_j)$ are bond "stiffnesses", $R_0(S_i, S_j)$ ideal lattice neighbor distances, and

$$\mathcal{H}_{el.,angles} = \sum_{i-j-k} A(S_i, S_j, S_k)(\boldsymbol{r}_{ij} \cdot \boldsymbol{r}_{kj} + R_0(S_i, S_j)R_0(S_k, S_j)/3)^2. \tag{5}$$

where $A(S_i, S_j, S_k)$ are "angular stiffnesses". In these equations, i–j denotes a bond between nearest neighbors i and j, while i–j–k denotes the angle of the nearest–neighbor bonds i–j and j–k with vertex at site j. The vector $\boldsymbol{r}_{ij} = \boldsymbol{r}_i - \boldsymbol{r}_j$, where $\boldsymbol{r}_i$ is the position of site i. Additional degrees of freedom are the linear sizes of the simulation cell Λ_x, Λ_y and Λ_z which fluctuate in

order to fix the pressure, e.g. $P = 0$. These simulations were repeated and continued [5] with the Stillinger-Weber potential [6] instead. This potential also contains two-body and three-body interactions, but in a slightly different form, and has the decided advantage that a simple Stillinger-Weber alloy expands upon heating whereas a Keating model shows unphysical shrinkage as the temperature is elevated.

The sampling algorithm is implemented as follows: For a single particle of type S_i at position $\boldsymbol{r}_i$, we randomly choose a new type S_i' and a slightly displaced new position $\boldsymbol{r}_i'$, while keeping the other particles and the simulation box dimensions fixed. This random trial move is accepted or rejected according to the usual Metropolis criterion. After all particles have been considered, we randomly choose new box dimensions Λ_x', Λ_y' and Λ_z'. The Metropolis acceptance criterion involves

$$\Delta\mathcal{H}_{eff} = \Delta\mathcal{H} - Nk_BT \ln \frac{\Lambda_x'\Lambda_y'\Lambda_z'}{\Lambda_x\Lambda_y\Lambda_z} , \tag{6}$$

where $\Delta\mathcal{H}$ is the energy change associated with this global distortion of the system and the latter term describes the change in translational entropy when the volume changes. (Since the total number of particles was constant, although the number of Si and Ge were not individually fixed, the simulations were performed in the semi-grand canonical ensemble.) Constant volume simulations were performed by simply turning off volume changing moves so that the volume was fixed instead of the pressure. The fixed volume was chosen to be about 3/4 of the way between the volumes for pure Si and pure Ge at $T = 0$. For some simulations parallel tempering [7,8] was implemented. Runs of length 10^5 – 10^7 MCS or more were used.

Since the Ising model and the corresponding binary alloy model interpretations are equivalent, either the magnetic field H or the chemical potential difference $\Delta\mu$ can be used for presenting results. Similarly, either the magnetization M or the Ge concentration c_{Ge} can be used to measure the order parameter. However, it should be noted that the elastic degrees of freedom result in a non-equivalence of the constant-M and constant-H ensembles. This has to do with the fact that the usual grand-canonical particle bath has the properties of a fluid, while it would have to have the properties of a solid (i.e., in particular, apply coherency stresses) in order to ensure equivalence. This difference was first noted by Vandeworp and Newman [9].

Monte Carlo data were obtained by sweeping the chemical potential at low temperature or by sweeping the temperature at fixed chemical potential for higher temperatures. Because of pronounced hysteresis, at low temperatures thermodynamic integration was used to find the intersection of the free energies and hence the location of the transition. At higher temperatures histogram reweighting [10] was used to locate the transition. Of course, in a finite system transitions are rounded and shifted [11], so we generated Monte Carlo data for different system sizes with the intent of extrapolating to the thermodynamic limit.

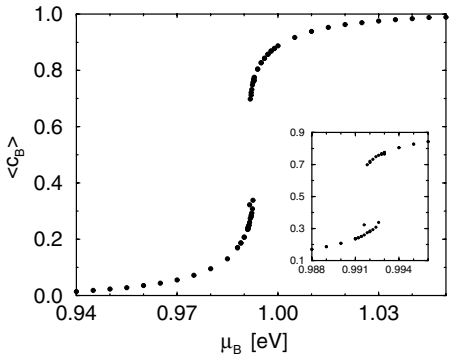


Fig. 1. Concentration dependence on chemical potential as determined from Monte Carlo simulations at $T = 0.026$ eV for the ferromagnetic compressible Ising model at constant pressure

3 Results

For the compressible ferromagnet at constant pressure the transition at low temperatures was strongly first order and free energy integrations were needed to locate the transition. At higher temperatures a jump in the magnetization could be found at the transition, coupled with slight hysteresis. An example is shown in Fig. 1 where we have chosen to portray the data using the Ising-binary alloy analogy. The inset shows hysteresis that is present at intermediate temperatures but is not immediately obvious unless a fine scale is used for observation. Near the critical point two dimensional histogram reweighting with finite size scaling was used to locate the critical point and determine critical behavior quite accurately. (Since the spin-up/spin-down symmetry of the rigid Ising model is no longer valid this required a search in the two-dimensional chemical potential-temperature space.) Excellent scaling was obtained using the mean-field exponents, and clear, systematic deviations were found for other choices of exponents. In addition, the 4th order cumulants were found to cross at a value that was distinctly different from the rigid Ising value ($(U_4)^* \sim 0.47$) but which matched the mean field value quite well. See Fig. 2. The inescapable conclusion was that the critical behavior was mean field-like. This result actually **preceded** the analysis of the Landau-Ginsburg-Wilson Hamiltonian analysis mentioned earlier.

As mentioned earlier, quite similar results were obtained for the compressible Ising ferromagnet when the Stillinger-Weber potential is substituted for the Keating potential. Although the scales of interest for the temperature and chemical potential change, the asymptotic behavior of the normalized thermodynamic quantities as the critical point is approached does not. Finite scaling of the reduced fourth order cumulant, shown in Fig. 3, demonstrates how well mean-field exponents work in this case as well.

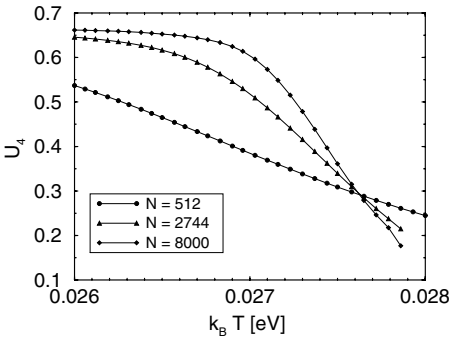


Fig. 2. Crossing of 4th order cumulants of the concentration for the compressible Ising ferromagnet at constant pressure

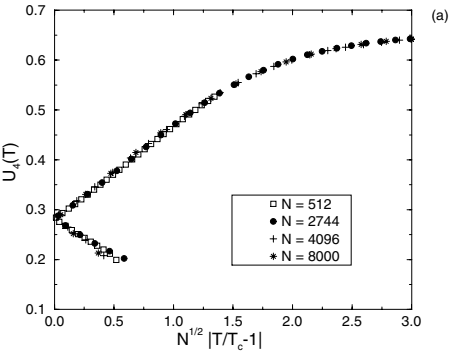


Fig. 3. Finite size scaling of the 4th order cumulants of the concentration for the compressible Ising ferromagnet at constant pressure

For the constant volume case, the analysis of the Landau-Ginsburg-Wilson Hamiltonian [2] suggests that at low temperatures two transitions should occur: one from a state that was primarily Si to one that is about equal in Si and Ge concentration, followed by a second transition to a state that was almost pure Ge. In contrast, however, the Monte Carlo data obtained by sweeping chemical potential at low temperatures revealed a relatively smooth increase of the Ge-concentration with increasing chemical potential and no apparent transition (see Fig. 4). However, a more careful, high resolution study of the regions in the small dashed boxes in Fig. 4 showed that the low temperature data indeed indicated very slight, smeared out hysteresis loops as can be seen in Fig. 5 for the range of chemical potential between $\sim 0.45 - 0.49$.

Similar hysteresis could be seen for both "high" and "low" values of chemical potential. It was reproducible and was rather insensitive to the length of the runs. This was matched by a hysteresis in the internal energy as well. The hysteresis slowly diminished as the temperature increased. At substantially higher temperature the hysteresis in the concentration disappeared for the

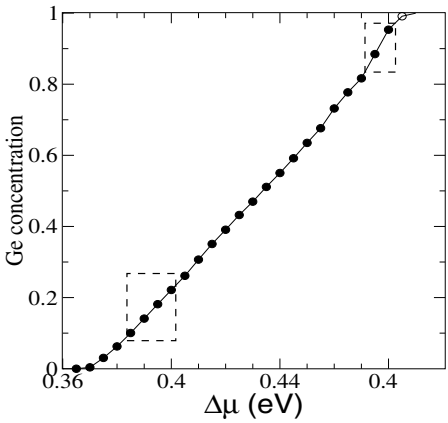


Fig. 4. Concentration dependence on chemical potential as determined from Monte Carlo simulations at $T = 0.006$ eV

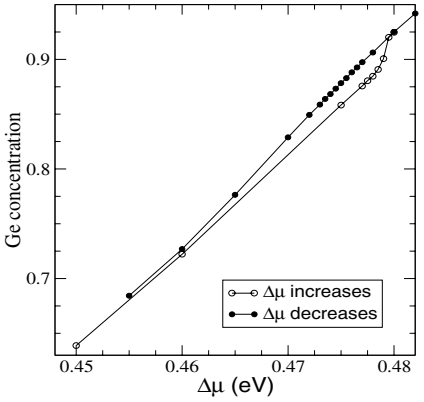


Fig. 5. Monte Carlo data showing hysteresis as the chemical potential was swept in opposite directions ($T = 0.0026$ eV)

sizes that could be studied, but an examination of the distribution of nearest neighbors of each species showed that a double peaked distribution remained below some chemical potential dependent temperature. It thus appeared as though there was a phase transition separating the low temperature, medium concentration and the disordered states over a wide range of chemical potential (see Fig. 6) but the nature of the transition was ephemeral.

These results could only be understood by augmenting the numerical data with pictures of the system obtained by visualization methods. Using the visualization program AViz developed at the Technion [12,13], we obtained many snapshots of the system for different simulation conditions. What was found was quite unexpected [14]. Snapshots of the system (see e.g. Fig. 7) showed that the minority phase developed slabs of approximately fixed thickness, and

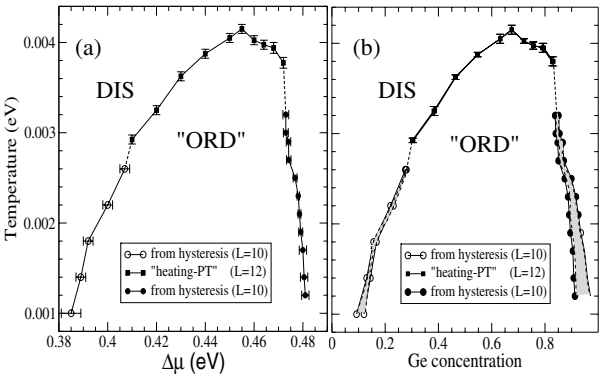


Fig. 6. Phase diagrams: (*left*) chemical potential-temperature space; (*right*) concentration-temperature space (*shaded* areas show regions where hysteresis is visible)

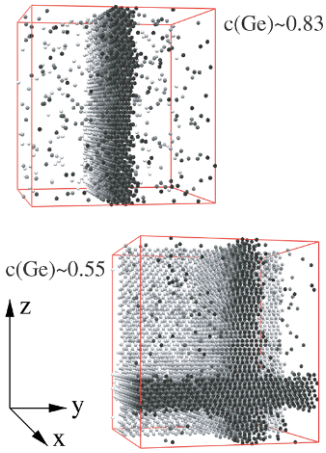


Fig. 7. Formation of slabs of the minority species. (*Top*) $\Delta\mu = 0.472$ eV; (*bottom*) $\Delta\mu = 0.410$ eV. $T = 0.0029$ eV. for both cases. Only Si atoms are shown, and they are greyscale coded vs depth

instead of growing with enhanced concentration the minority species would simply form another slab. This behavior was smeared out for small systems but became quite clear for larger systems. Particularly long time scales were associated with the formation or dissolution of these slabs as the temperature or chemical potential were swept, and this property gave rise to the rather strange "smeared" hysteresis loops. Forming the planes was particularly slow and this produced asymmetry in the hysteresis. At higher temperatures the formation of the slabs seemed to indicate the transition to an "ordered" state. Moreover, the presence of the slabs introduces a special length scale into the problem that had not been included in the LGW Hamiltonian study. The

resultant phase diagram, shown in Fig. 6, shows that instead of the two predicted phase lines a single, closed phase boundary is present. The range of coexistence along the 1st order phase boundary, as determined from the hysteresis, is quite narrow, and for the lattice sizes that were accessible it seemed to disappear completely at high temperature. The phase boundary appears to be first order everywhere except at the single point where the derivative $dT/dH = 0$. Although it is now clear what went wrong in the analysis of this case in 2 (incorrect assumption of a finite interfacial tension), the physical mechanism which brings about the observed slab formation still awaits elucidation [15].

For studies of compressible antiferromagnets, essentially the same model (distortable diamond net with a Stillinger-Weber potential) was used but the sign of the interaction between unlike species was modified to produce an antiferromagnetic ground state. (Of course, in this case the resultant alloy is no longer related in any way to Si/Ge and must be regarded merely as a model of theoretical relevance.) In both cases the resultant phase boundaries separating the ordered and disordered phases were 2nd order everywhere. Careful analysis of the critical behavior revealed rigid Ising exponents. For example, Fig. 8 shows a finite size scaling plot, made with rigid Ising model critical exponents, for the compressible Ising antiferromagnet at constant volume [17].

Similar behavior was found for the compressible Ising antiferromagnet at constant pressure [3]. Moreover, as shown in Fig. 9, cumulant crossings for both cases occurred at the value found previously for a rigid Ising model. Since **no indication** of crossover to any other kind of behavior was seen, substantially larger systems would probably be required to see some other possible kind of asymptotic behavior, if it exists! In such a case the simula-

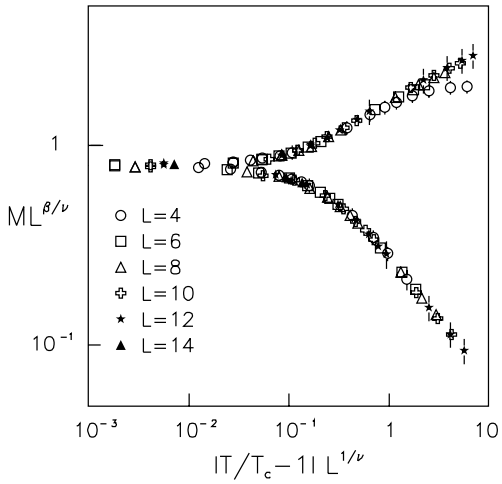


Fig. 8. Finite size scaling of the order parameter for the compressible Ising antiferromagnet at constant volume

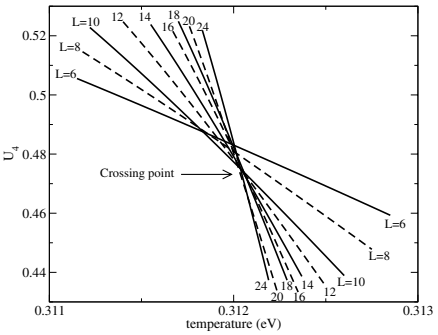


Fig. 9. Cumulant crossing for the compressible antiferromagnet at constant pressure

tional needs for further studies of the present model would be well beyond the capability of current computer resources.

4 Conclusions

A Monte Carlo study of Ising models on a distortable diamond lattice yielded results that were in some cases unexpected and which disagree with theoretical predictions. For the ferromagnet at constant pressure mean-field critical behavior was found, whereas for constant volume a closed first order line was found instead of the two predicted first order lines, terminating in critical points. In the "ordered" phase, the less favorable species forms "slabs" of approximately fixed thickness; as the concentration of the less favorable species is enhanced, the number of slabs increases. The compressible Ising antiferromagnet showed a closed, Ising like 2nd order phase boundary separating the ordered and disordered states. Of course, we have studied only a single "family" of compressible Ising models, and it may well be that the general classifications of such models is even more complex than we envisage.

Acknowledgments

We are indebted to K. Binder for illuminating discussions. This research was supported by NSF grant No DMR-0341874 and by the U.S.–Israel Binational Science Foundation.

References

1. B. Dünweg and D. P. Landau (1993) *Phys. Rev. B* **48**, p. 14182

2. B. Dünweg (2000) *Computer Simulationen zu Phasenübergängen und Kritischen Phänomenen*, Habilitationsschrift, U. Mainz
3. X. Zhu, F. Tavazza, D. P. Landau, and B. Dünweg (2005) *Phys. Rev. B* **72**, p. 104102
4. P. N. Keating (1966) *Phys. Rev.* **145**, p. 627
5. M. Laradji, D. P. Landau, and B. Dünweg (1995) *Phys. Rev. B* **51**, p. 4898
6. F. H. Stillinger and T. A. Weber (1985) *Phys. Rev. B* **31**, p. 5262
7. K. Hukushima and K. Nemoto (1996) *Phys. Soc. of Japan* **65**, p. 1604
8. R. H. Swendsen and J. S. Wang (1986) *Phys. Rev. Lett.* **57**, p. 2607
9. E. M. Vandeworp and K. E. Newman (1995) *Phys. Rev. B* **52**, p. 4086; E. M. Vandeworp and K. E. Newman (1997) *Phys. Rev. B* **55**, p. 14222
10. A. M. Ferrenberg and R. H. Swendsen (1989) *Phys. Rev. Lett.* **63**, p. 1195
11. V. Privman (Ed.) (1990) *Finite Size Scaling and Numerical Simulation of Statistical Systems.* World Scientific, Singapore
12. J. Adler, A. Hashibon, and G. Wagner (2002) *Recent Developments in Computer Simulation Studies in Condensed Matter Physics, XIV*, ed. D. P. Landau, S. P. Lewis, and H.-B. Schüttler, Springer, Heidelberg, p. 160
13. J. Adler (2003) *Comput. in Sci. and Engineering* **5**, p. 61
14. F. Tavazza, D. P. Landau, and J. Adler (2004) *Phys. Rev. B* **70**, p. 184103
15. This prediction was based upon the assumption (2) that the interfacial tension between the Si-rich and Ge-rich phases is finite. This assumption has turned out to be wrong. This gives rise to a complete suppression of capillary waves [16], and also to a phase behavior which strongly differs from that assumed in 2
16. B. J. Schulz, B. Dünweg, K. Binder, and M. Müller (2005) *Phys. Rev. Lett.* **95**, p. 096101
17. L. Cannavacciuolo and D. P. Landau (2005) *Phys. Rev. B* **71**, p. 134104

Computer Simulation of Colloidal Suspensions

H. Löwen

Heinrich-Heine-Universität Düsseldorf, Universitätsstraße 1, 40225 Düsseldorf, Germany
hlowen @thphy.uni-duesseldorf.de



Hartmut Löwen

H. Löwen: *Computer Simulation of Colloidal Suspensions*, Lect. Notes Phys. **704**, 139–161 (2006)
DOI 10.1007/3-540-35284-8_7
© Springer-Verlag Berlin Heidelberg 2006

1 Introduction 141
2 Effective Interactions 143
3 Approximative Density Functionals 148
4 Charged Colloidal Dispersions 149
5 Star Polymers 153
6 Colloids and Polymers: Depletion Interactions 155
7 Conclusions 157
References 158

By definition, soft matter systems react sensitively upon external mechanical perturbations. This material class includes mesoscopic complex fluids such as colloidal suspensions. It is a major challenge to understand the fascinating properties of colloids from first principles, i.e., by deriving its properties from the microscopic interactions. Here, concepts borrowed from statistical physics are described, which are capable to overbridge the gap from microscopic over mesoscopic to macroscopic length scales. This is illustrated explicitly for charged colloidal suspensions and for star polymer solutions. A particular emphasis is placed on density functional theory.

1 Introduction

Soft matter [1] is synonymous with "complex fluids" and "colloids" although emphasis is put on different aspects in using these substitute names. The textbook definition of *colloids* is that there is at least one mesoscopic length scale in the range between 1nm and 1μm on which the system exhibit discontinuities. The main difference between a *molecular* and a *colloidal* system becomes immediately clear in comparing two liquids known from everyday life: water and milk. Water looks like a clear and structureless fluid on a length scale down towards nanometers and one needs molecular resolution to detect the water and alcohol molecules. Milk, on the other hand, consists of fat globules exhibiting already a structure on a length scale of 10 microns. Furthermore, milk contains submicron-sized caseine micelles which are the crucial building blocks in producing cheese. Hence the length scales of the characteristic structure are quite different in the two cases: water is a molecular and milk a colloidal liquid.

Discontinuities on a mesoscopic scale can happen between different phases. Accordingly [2], there are eight different kinds of colloidal dispersions depending on whether the disperse phase and the dispersion medium are in the solid, liquid or gas phase. The name "colloids" is also frequently used in a more specialized sense for colloidal suspensions which are solid particles embedded in a molecular liquid. Typical examples are printing ink, paints, blood, urine, spittle, adhesives (e.g. glue where the name "colloid" stems from), viruses, and muddy water.

The physics of colloids is the domain between molecular physics occuring on a length scale smaller than one nanometer and the traditional solid state physics of small crystallites which are larger than a micron. As a characteristic feature of colloids, their bulk to interface ratio is much larger than that of a crystallite. This can readily be seen by cutting a macroscopic solid into subunits and counting the resulting area. Hence colloidal and interfacial properties are very much inter-related.

If the term *complex fluids* is used, emphasis is put on the complexity of the description which involves very different (microscopic and mesoscopic) length scales. Finally, the expression *soft matter* puts emphasis on the mechanical

properties of colloidal materials. They react sensitively on mechanical perturbations as compression or shear. An example is given below.

Let us first focus on colloidal suspensions, i.e., solid mesoscopic particles embedded in a molecular solvent. A theorist would immediately approximate the solid colloidal particles by isotropic spheres. This is of course the leading order in a systematic approximation of the particle shape but it is not that crazy as it looks at a first glance. By sophisticated preparation methods one is nowadays indeed able to realize excellent model suspensions of monodisperse submicron-sized latex or polystyrol spheres [3]. For high colloidal concentrations, these spheres self-organize is crystalline arrays, i.e., they undergo a freezing transition. Such a colloidal crystal has a lattice constant a in the mesoscopic regime which leads to different elastic properties as compared to molecular solids. This is illustrated strikingly by considering the shear modulus G of a colloidal crystal. Roughly speaking, G scales with a typical energy scale (say the thermal energy k_BT) divided by a typical volume of the elementary crystal cell, $G \approx k_BT/a^3$. Hence a colloidal crystal has a shear modulus which is 9–12 orders of magnitude smaller than that of an ordinary crystal! This implies that colloidal crystals are vulnerable to shear and explains why the term "soft matter" is appropriate for colloidal samples.

Polymers are other prominent examples of soft matter. They are macromolecules composed of many monomeric units. A typical example is a linear hydrocarbon chain. But there are more complicated topologies conceivable, such as branched polymers (called dendrimers) or star polymers which consist of f linear chains attached to a common microscopic center [4]. The monomers can be charged resulting in a highly charged macromolecule which is called polyelectrolyte.

In this chapter I shall discuss systematic coarse-graining procedures which lead to effective interactions between the largest, mesoscopic particles in multicomponent, multiscale fluid mixtures. These effective interactions follow from a rigorous "integrating out" of microscopic degress of freedom. This concept allows for a simple understanding of trends in the phase behaviour, structure and dynamics of colloids and polymers. Moreover, the effective interaction can be used in standard simulations of samples involving only the large particles which now play the role of molecules in atomistic simulations. After a formal Statistical Mechanics justification of the coarse-graining procedure in Sect. 2, we shall briefly propose approximative density functional in Sect. 3 which are necessary to implement calculations within the coarse graining picture. The coarse graining concept will then be successively applied to interacting electric double-layers (Sect. 4), to solutions of star polymers (Sect. 5). A major part of this chapter (in particular that with an emphasis of computer simulations) is already published elsewhere in a recent review of the author Hansen [5]. Other useful review articles concerning the matter of effective interactions are those from Likos [6] and Belloni [7], aspects of charged suspensions are reviewed by Hansen and Löwen [8] and a recent review on computer simulations of colloids is provided by Dijkstra [9].

2 Effective Interactions

From a more theoretical point of view, the great challenge in soft matter is to understand and predict the macroscopic properties starting from the microscopic interactions. Already for pure molecular systems such an "ab initio" calculation poses a very hard problem. For a soft matter system this is even more complicated due to the presence of intermediate mesoscopic length scales.

Typical quantities of interest are the osmotic pressure of a colloidal dispersion, the elastic moduli or the phase behavior of colloidal suspensions as a function of internal thermodynamic or external parameters. A general framework to overbridge the different length scales is highly desirable because: i) a fundamental understanding of the thermodynamics of soft matter including biological macromolecules can be reached and ii) new material properties could be predicted.

It is clear that methods of classical equilibrium statistical mechanics should be applicable as most of the constituents are described as classical particles. Bridging the different length scales is most conveniently done in different steps from microscopic to mesoscopic and then from mesoscopic to macroscopic length scales. The first step can be made by using the important concept of the effective interaction. The second step is performed using ideas from classical many-body theory. Let us first outline these concepts briefly in general and then illustrate them for examples, in particular.

An efficient statistical description of multi-component systems involving particles of widely different sizes requires a controlled-coarse-graining which may be achieved by integrating ("tracing") out the degrees of freedom of the majority components of "small" particles, which may be solvent molecules, microscopic ions ("micro-ions") or monomers of macro-molecules. For the sake of simplicity, consider an asymmetric binary "mixture" of N_1 "large" spherical particles, with centres of positions $\{\boldsymbol{R}_i\}$ $(1 \leq i \leq N_1)$, and $N_2 \gg N_1$ "small" particles at positions $\{\boldsymbol{r}_j\}$ $(1 \leq j \leq N_2)$. Restriction will be made to thermodynamic equilibrium states. If classical statistics apply, integration over momenta is trivial, and the focus will be on configurational averages. The total potential energy of the mixture may be conveniently split ino three terms:

$$U(\{\boldsymbol{R}_i\},\{\boldsymbol{r}_j\}) = U_{11}(\{\boldsymbol{R}_i\}) + U_{22}(\{\boldsymbol{r}_j\}) + U_{12}(\{\boldsymbol{R}_i\},\{\boldsymbol{r}_j\}) \tag{1}$$

At a fixed inverse temperature $\beta = 1/k_B T$, the configurational part of the Helmholtz free energy F of the two-component system may be formally expressed as:

$$\begin{aligned} \exp(-\beta F) &= Tr_1\, Tr_2\, \exp(-\beta\, U) \\ &= Tr_1\, \exp(-\beta\, U_{11})\, Tr_2\, \exp(-\beta(U_{12}+U_{22})) \\ &= Tr_1\, \exp(-\beta\, U_{11})\, \exp(-\beta\, F_2(\{\boldsymbol{R}_i\})) \\ &= Tr_1\, \exp(-\beta\, V_{11}(\{\boldsymbol{R}_i\})) \end{aligned} \tag{2}$$

where the short-hand trace notation implies integration over the configuration space of species 1 or 2, i.e.

$$Tr_\alpha = \frac{1}{N_\alpha!}\int d^{3N_\alpha} r \tag{3}$$

$V_{11}(\{\boldsymbol{R}_i\})$, the effective interaction energy of the large particles, is the sum of their direct (or bare) interaction energy U_{11}, and of the configurational free energy of the fluid of small particles in the "external" field of the large particles F_2; the latter depends parametrically on the configuration $\{\boldsymbol{R}_i\}$ of the large particles

$$V_{11}(\{\boldsymbol{R}_i\}) = U_{11}(\{\boldsymbol{R}_i\}) + F_2(\{\boldsymbol{R}_i\}) \tag{4}$$

and can be written as:

$$F_2(\{\boldsymbol{R}_i\}) = -k_B T \ln[Tr_2 \exp(-\beta(U_{12} + U_{22})] \tag{5}$$

Up to now, no approximation has been made.

Three key aspects of the effective interaction V_{11} must be underlined. Firstly, *any* physical quantity $\mathcal{A}(\{\boldsymbol{R}_i\})$ depending only on the coordinates of the big particles can be formally averaged via the effective interaction

$$Tr_1\, Tr_2 \mathcal{A}(\{\boldsymbol{R}_i\}) \exp(-\beta\, U) = Tr_1 \mathcal{A}(\{\boldsymbol{R}_i\}) \exp(-\beta\, V_{11}(\{\boldsymbol{R}_i\})) \tag{6}$$

Hence, once the effective interaction is known, any averages (e.g. pair correlations) of the big particles can be extracted directly.

Secondly, due to the presence of a free energy, F_2, V_{11} is obviously *state-dependent*, and has an entropic contribution of the small particles ($F_2 = U_2 - T\, S_2$).

Finally, although the direct interaction U_{11} may be pair-wise additive, this is no longer true of V_{11}. The free energy $F_2(\{\boldsymbol{R}_i\})$ generally has many-body-contributions, so that V_{11} will be of the more general form (with the change of notation $N_1 \to N$ and $V_{11} \to V_N$):

$$V_N(\{\boldsymbol{R}_i\}) = V_N^{(0)} + \sum_{i\le j}\sum v_2(\boldsymbol{R}_i, \boldsymbol{R}_j) + \sum_{i\le j\le k}\sum\sum v_3(\boldsymbol{R}_i, \boldsymbol{R}_j, \boldsymbol{R}_k) + \ldots . \tag{7}$$

$V_N^{(0)}$ is a state dependent but configuration-independent "volume" term, which has no bearing on the local structure of the large particles, but through its contribution to the thermodynamic properties, it can, in some cases, strongly influence their phase behaviour [10].

Let us now give some important examples of how to apply the concept of coarse-graining to soft matter systems. Obviously one has first of all to specify which statistical degrees of freedom should be considered and which of those should be integrated out ("small particles") and which of them should be left in the effective interaction ("big particles"). Depending on this choice

Table 1. Type of microscopic degrees of freedom which are integrated out for different kinds of macroparticles.

soft matter system	microscopic degrees which are coarse-grained	resulting physical effect
charged colloids	counterions, salt-ions	screening of Coulomb repulsions
polymers (linear chains, star polymers, dendrimers)	monomers	polymers viewed as soft spheres
colloid-polymer mixtures	polymer coils	depletion attraction between colloids
binary mixtures of big and small colloids	small colloids	depletion attraction and accumulation repulsion
nanoparticles in solvent	solvent particles	discrete solvent effects
polyelectrolyte stars	counterions, salt-ions	entropic interaction between centers

one can cover quite different physical phenomena which are summarized in Table 1. These include counterion screening of charged suspensions, depletion interactions in mixtures and polymer modeling by soft spheres. Most of those effects will be described in detail in the next sections.

Expression (4) for the effective interaction, or potential of mean force, was derived in the canonical ensemble, where the total numbers of small and large particles are fixed (closed system). In many practical situations the binary system is in osmotic equilibrium with a pure phase of the small particles (e.g the solvent), and the appropriate ensemble for such an open system is the semi-grand canonical ensemble where N_1 and the chemical potential μ_2 of the small particles (rather than N_2) are fixed. The corresponding thermodynamic potential is the semi-grand potential $\Omega_2 = \Omega_2(T, N_1, \mu_2; \boldsymbol{R}_i)$, and the effective interaction energy of the large particles will then be:

$$V_{11}(\boldsymbol{R}_i) = U_{11}(\boldsymbol{R}_i) + \Omega_2(\{\boldsymbol{R}_i\}) \tag{8}$$

which will again be state-dependent, a function of temperature, volume V and μ_2 (rather than $\rho_2 = N_2/V$).

In summary, the initial two-component system, involving a large number of microscopic degrees of freedom, has been reduced to an effective one-component system involving only the degrees of freedom of the mesoscopic particles. The price to pay is that the effective interaction energy is state-dependent and generally involves many-body terms. Approximations must now be invoked to calculate the highly non-trivial F_2 or Ω_2 term, i.e. the part of the interaction energy between the large particles induced by the small particles. Three different strategies have so far been used in practical implementations:

(a) For any given configuration $\{\boldsymbol{R}_i\}$ of the large particles, the small particles are subjected to the "external" potential $U_{12}(\{\boldsymbol{R}_i\},\{\boldsymbol{r}_j\})$, and hence form an inhomogeneous fluid, characterized by a local density $\rho(\boldsymbol{r};\{\boldsymbol{R}_i\})$. The thermodynamic potentials F_2 or Ω_2 are functionals of $\rho(\boldsymbol{r})$, and full use can be made of the classical density functional theory (DFT) of non-uniform fluids for the small particles [11,12]. DFT guarantees the existence of a excess free energy density functional $F_{exc}[\rho(\boldsymbol{r}]$ such that the free energies F_2 or Ω_2 can be written exactly as functionals:

$$F_2[\rho(\boldsymbol{r}] = F_{id}[\rho(\boldsymbol{r}] + F_{exc}[\rho(\boldsymbol{r}] + \sum_{j=1}^{N_1} \int \rho(\boldsymbol{r}) u_{12}(\boldsymbol{r}_i - \boldsymbol{R}_j)\, d\boldsymbol{r} \tag{9}$$

and

$$\Omega_2[\rho(\boldsymbol{r}] = F_2[\rho(\boldsymbol{r}] - \mu_2 \int \rho(\boldsymbol{r}) d\boldsymbol{r} \tag{10}$$

Here, $F_{id}[\rho(\boldsymbol{r}]$ is the functional of an ideal gas which is known exactly

$$F_{id}[\rho(\boldsymbol{r}] = k_B T \int \rho(\boldsymbol{r}) \left[\ln(\Lambda_2^3\, \rho(\boldsymbol{r})) - 1\right] d\boldsymbol{r} \tag{11}$$

with Λ_2 being the thermal wave-length of the small particles. The density functionals given is (9) and (10) give the physical free energies F_2 or Ω_2 if the equilibrium density $\rho(\boldsymbol{r};\{\boldsymbol{R}_i\})$ of the small particles is inserted into the functional which follows from the variational minimization principle

$$\left.\frac{\delta \Omega_2[\rho^*(\boldsymbol{r})]}{\delta \rho^*(\boldsymbol{r})}\right|_{\rho^*=\rho} = 0 \tag{12}$$

where $\rho^*(\boldsymbol{r})$ is a properly parametrized trial density. The only difficulty is that in general the exact functional $F_{exc}[\rho(\boldsymbol{r}]$ is not known. Tractable approximations are known for hard spheres and soft potential fluids which are summarized in the next chapter. The optimization (12) may be implemented by steepest descent or conjugate gradient techniques, and the resulting effective potential energy between large particles can then be used directly in standard MC or MD simulations [13]. In the latter case, the forces $\boldsymbol{F}_i$ acting on the large particles may be directly calculated from a classical version of the Hellmann-Feynman theorem:

$$\begin{aligned} \boldsymbol{F}_i &= -\boldsymbol{\nabla}_i V_{11}(\{\boldsymbol{R}_j\}) \\ &= -\boldsymbol{\nabla}_i U_{11}(\{\boldsymbol{R}_j\} - \langle \boldsymbol{\nabla}_i U_{12}(\{\boldsymbol{R}_j\},\{\boldsymbol{r}_l\}\rangle_{\{\boldsymbol{R}_j\}} \end{aligned} \tag{13}$$

where the angular bracket denotes an equilibrium average over the degrees of freedom of the small particles, for a forced configuration $\{\boldsymbol{R}_i\}$ of the large ones. If the interaction energy U_{12} between the two species is pairwise additive ($U_{12} = \sum_{i=1}^{N_1}\sum_{j=1}^{N_2} u_{12}(|\boldsymbol{r}_i - \boldsymbol{R}_j|)$), the force $\boldsymbol{F}_i$ is directly expressible in terms of the local equilibrium density $\rho(\boldsymbol{r})$:

$$\boldsymbol{F}_i = -\boldsymbol{\nabla}_i U_{11}(\{\boldsymbol{R}_j\}) - \int \rho(\boldsymbol{r})\, \boldsymbol{\nabla}_i\, u_{12}(\boldsymbol{r}_i - \boldsymbol{R}_j)\, d\boldsymbol{r} \tag{14}$$

The optimization can also be achieved "on the fly", along lines directly inspired by the Car-Parrinello method for ion-electron systems [14]. Successive minimization and large particle updating steps are replaced by a single dynamical evolution, which involves the physical motion of the large particles and fictitious dynamics of the local density of small particles, parametrized by a plane wave expansion [15].

(b) The previous DFT optimization method calculates directly the total effective energy of interaction between the large particles, or the resulting forces acting on each of these particles, without dividing V_N up into pair triplet and higher order interactions, as written in (7). Another strategy is to attempt to compute these various contributions separately. At very low concentration of large particles, the effective pairwise interaction v_2 is expected to be dominant. In order to map out v_2 as a function of the distance r between two large particles, one may use standard MC or MD algorithms to simulate a bath of small particles in the field of two fixed large particles. Equation (13) may then be used to calculate the mean forces acting on the two mesoparticles (which are opposite if the latter are identical) for each distance $r = |\boldsymbol{R}_1 - \boldsymbol{R}_2|$. The effective pair potential $v_2(r)$ finally follows from an integration of the forces. This procedure must be repeated for each distance r, but there are no time-scale or ergodicity problems, since the two large particles are fixed. The same goal can be achieved by appealing once more to DFT for the inhomogeneous fluid of small particles, subjected to the force field of two fixed large particles. The optimization may be carried out in r-space, using an adequate Eucledian or non-Eucledian [16] grid on which the local density of small particles is defined. For two identical large particles, the local density has obvious cylindrical symmetry, but under favourable conditions, a considerable simplification occurs by fixing one of the large particles and considering an infinitely dilute solution of large particles in a bath of small particles around the fixed large particle. The density profile of the large particles in the zero concentration limit is directly related to the effective pair potential between two large particles in a bath of small particles [17], i.e.,

$$v_2(r) = -k_B\, T \lim_{\rho_1 \to 0} \ln\left(\frac{\rho_1(r)}{\rho_1(r \to \infty)}\right) \tag{15}$$

The advantage is that the two density profiles $\rho_1(r)$ and $\rho_2(r)$ are now spherically symmetric, but the method requires the prior knowledge of an accurate density functional for an asymmetric binary mixture. This strategy may be generalized to the calculation of three-body and higher order effective interactions, by considering the density profiles of large and small particles around two or more fixed large particles [18].

(c) Although the effective interaction energy (4) or (8) is not, in general, pairwise additive at finite concentrations of the large particles, it would be

very convenient, for computational purposes, to reduce it, at least approximately, to a pairwise additive form. Contrarily to the two-body potential $v_2(r)$ discussed in the previous paragraph, which is only valid in the low density limit of large particles, the effective pair potential corresponding to finite concentrations is expected to be density-dependent, and will, in some average sense, incorporate the contributions of higher order terms in (7). Such effective density-dependent pair potentials can, in some cases, be derived from approximate functionals or from inversion procedures, examples of which will be described in Sect. 5.

3 Approximative Density Functionals

In this section we summarize different modern approximations for the excess free energy density functional $F_{exc}[\rho(\boldsymbol{r}]$ for different small-small interaction pair potentials $u_{22}(r)$, namely hard spheres and soft particles. The case of Coulomb interactions will be re-discussed in the next chapter.

For *hard spheres* of diameter σ the best current functional approximation is that of Rosenfeld's fundamental measure theory [19]. It can be constructed also for hard sphere mixtures but here we restrict ourselves to a one-component hard sphere fluid. In this approximation one takes

$$\mathcal{F}_{exc}[\rho] = k_B T \int d\boldsymbol{r} \Phi[\{n_\alpha(\boldsymbol{r})\}] \tag{16}$$

where one introduced a set of weighted densities

$$n_\alpha(\boldsymbol{r}) = \int_\Omega d\boldsymbol{r}' \rho(\boldsymbol{r}') w_\alpha(\boldsymbol{r} - \boldsymbol{r}') \tag{17}$$

Here, the index $\alpha = 0, 1, 2, 3, V1, V2$ labels six different weighted densities and six different associated weight functions. Explicitly these six weight functions are given by

$$w_0(\boldsymbol{r}) = \frac{w_2(\boldsymbol{r})}{\pi\sigma^2} \tag{18}$$

$$w_1(\boldsymbol{r}) = \frac{w_2(\boldsymbol{r})}{2\pi\sigma} \tag{19}$$

$$w_2(\boldsymbol{r}) = \delta\left(\frac{\sigma}{2} - r\right) \tag{20}$$

$$w_3(\boldsymbol{r}) = \Theta\left(\frac{\sigma}{2} - r\right) \tag{21}$$

$$w_{V1}(\boldsymbol{r}) = \frac{w_{V2}(\boldsymbol{r})}{2\pi\sigma} \tag{22}$$

and

$$w_{V2}(\boldsymbol{r}) = \frac{\boldsymbol{r}}{r}\delta\left(\frac{\sigma}{2} - r\right) \tag{23}$$

Note that the index V denotes a vector weight function. We can express this fact by writing $w_{V1} \equiv \boldsymbol{w}_{V1}$, $n_{V1} \equiv \boldsymbol{n}_{V1}$,... Finally the function Φ is given by

$$\Phi = \Phi_1 + \Phi_2 + \Phi_3 \tag{24}$$

with

$$\Phi_1 = -n_0 \ln(1 - n_3) \tag{25}$$

$$\Phi_2 = \frac{n_1 n_2 - \boldsymbol{n}_{V1} \cdot \boldsymbol{n}_{V2}}{1 - n_3} \tag{26}$$

and

$$\Phi_3 = \frac{n_3^2 (1 - (\boldsymbol{n}_{V2}/n_2)^2)^3}{24\pi(1 - n_3)^2} \tag{27}$$

The six weight functions are connected to the geometrical (fundamental) Minkowski measures [20]. There are several arguments in favor of the Rosenfeld approximation: as an example we mention that the freezing transition can be calculated by plugging in a constant density field for the fluid phase and a lattice sum of Gaussian peaks in the solid phase. If the width of the Gaussians and the prefactor are taken as variational parameters one gets a first-order freezing transition with coexisting packing fractions of $\eta_f = 0.491$ and $\eta_s = 0.540$ which are very close to "exact" simulation data $\eta_f = 0.494$, $\eta_s = 0.545$.

In the complementary case of very soft interactions, on the other hand, it has recently been shown that a mean-field approximation for the density functional is a very good approximation [21–23]. If the pair potential $u_{22}(r)$ is finite at the origin, then it can be shown that a mean-field functional is exact in the limit of very large densities. It works, however, amazingly well also for finite densities. In the mean-field approximation one takes:

$$\mathcal{F}_{exc}[\rho] = \frac{1}{2} \int d\boldsymbol{r} \int d\boldsymbol{r'} \rho(\boldsymbol{r}) \rho(\boldsymbol{r'}) u_{22}(|\boldsymbol{r} - \boldsymbol{r'}) \tag{28}$$

All other intermediate cases are more difficult. Some success is to map harsh interaction onto effective hard spheres employing some ideas from the construction of Rosenfeld's functional [24]. Treating attractive tails has mainly been limiting to mean-field-like approaches as well.

4 Charged Colloidal Dispersions

Electric double-layers around mesoscopic colloidal particles of various shapes (spheres, rods, platelets, ...) or around polyelectrolytes make the generally dominant contribution to the effective interaction between highly-charged particles, which will be referred to as polyions [7, 8]. Most simulations are based on a primitive model, whereby the discrete nature of the aqueous solvent is neglected, and a macroscopic value of the dielectric permittivity ϵ is assumed.

At very low polyion concentration, strategy b) of the previous section may be adopted to compute an effective pair interaction between two polyions, which is screened by microscopic counterions of opposite sign, as well as coions in the presence of added salt. The resulting effective pair potential turns out to be invariably repulsive of the screened Coulomb form predicted a long time ago by Derjaguin, Landau Verwey and Overbeek (DLVO) [25] as long as the microions are monovalent. However if divalent counterions are present, they are more strongly correlated, and this may lead to a short-range attraction between equally-charged polyions, due to an overscreening effect [26]. Although most of the work on effective pair interactions has focussed so far on spherical polyions, some recent MC simulations have investigated the case of lamellar colloids [27]. The triplet interaction between spherical polyions has similarly been calculated by MD simulations of co and counterions in the field of three fixed polyions [28], and turns out to be attractive under most circumstances. In the opposite limit of high concentrations, each polyion is confined to a cage of neighbouring polyions, so that many-body interactions are expected to be important, and pairwise additivity of the effective interaction is expected to break down. It is then reasonable to consider a Wigner-Seiz cell model, where a cell of geometry adapted to the shape of the polyions (e.g. a spherical cell for spherical polyions) contains one polyion at its centre, surrounded by co and counterions, such that overall charge neutrality is ensured, and with appropriate boundary conditions for the electric field on the surface of the cell. A physically reasonable boundary condition is to impose that the normal-component of the electric field vanishes on the surface. The initial problem involving many polyions is thus approximately reduced to the much simpler problem of a single polyion surrounded by its electric double-layer. Although all information on correlations between polyions is lost, the cell model allows a calculation of the thermodynamic properties of concentrated suspension, from MC or MD simulations of the inhomogeneous fluid of microions contained in the cell, as well as an estimate of the effective polyions charge, taking into account the phenomenon of counterion "condensation" [29, 30]. Such simulations provide stringent tests for approximate DFT calculations, including Poisson-Boltzmann (PB) theory.

At moderate polyion concentrations, the two previous strategies break down. Strategy a) of the previous section, based on the step by step or "on the fly" optimization of an appropriate free energy functional of the microion density profiles, is the most appropriate [15]. The free energy functional $F_2[\rho_+(\boldsymbol{r}), \rho_-(\boldsymbol{r}), \{\boldsymbol{R}_i\}]$ of the co- and counterion densities is conveniently split into ideal, Coulomb, external and correlation parts:

$$\begin{aligned} F_2[\rho_+, \rho_-] = F_{id}[\rho_+] + F_{id}[\rho_-] + F_{Coul}[\rho_c] \\ + F_{ext}[\rho_+] + F_{ext}[\rho_-] + F_{corr}[\rho_+, \rho_-] \end{aligned} \tag{29}$$

where:

$$F_{id}[\rho_\alpha] = k_B T \int \rho_\alpha(\boldsymbol{r}) \left[\ln(\Lambda_\alpha^3 \, \rho_\alpha(\boldsymbol{r})) - 1\right] d\boldsymbol{r} \tag{30}$$

$$F_{Coul}[\rho_\alpha] = \frac{e^2}{2} \int d\boldsymbol{r} \int d\boldsymbol{r}' \frac{\rho_c(\boldsymbol{r}) \, \rho_c(\boldsymbol{r}')}{|\boldsymbol{r} - \boldsymbol{r}'|} \tag{31}$$

$$\begin{aligned} F_{ext}[\rho_\alpha] &= \int \varphi_{ext}(\boldsymbol{r}) \, \rho_\alpha(\boldsymbol{r}) \, d\boldsymbol{r} \\ &= \sum_{i=1}^{N_1} \int u_{1\alpha}(\boldsymbol{r} - \boldsymbol{R}_i) \, \rho_\alpha(\boldsymbol{r}) \, d\boldsymbol{r} \end{aligned} \tag{32}$$

In (31), $\rho_c(\boldsymbol{r}) = z_+ \rho_+(\boldsymbol{r}) + z_- \rho_-(\boldsymbol{r})$ is the charge density of the microions (of valence z_α). The polyion-microion potentials $u_{1\alpha}$ in (32) contain a hard core repulsion and a long-range Coulomb attraction (counterions) or repulsion (coions). Rapid variations of the densities profiles $\rho_\alpha(\boldsymbol{r})$ near the surfaces of the polyions, which would pose numerical problems in r-space (grid) or k-space (large $\boldsymbol{k}$ Fourier components) may be avoided by the use of appropriate classical polyion-microion pseudopotentials [15]. The correlation term F_{corr} may be expressed within the local density approximation (LDA) [15]. If it is neglected, the functional (29) reduces to the mean-field Poisson-Boltzmann (PB) form. Optimization based on the functional (29)–(32) has been achieved with the "on the fly" MD strategy for spherical polyions with counterions only (no salt) [15], and the presence if salt (i.e. with co and counterions) [31]. The effective forces between colloids are reasonably well represented by a pair-wise additive screened-Coulomb form provided the (effective) polyion charge and the screening length are treated as adjustable parameters. Other applications include rigid rod-like polyions [32], and flexible polyelectrolytes [13], the latter being investigated by MC simulations coupled with steepest descent optimization, to allow a more efficient exploration of polyelectrolyte configuration space. If F_{corr} is neglected in the functional (29), and the ideal terms are replaced by their quadratic expansion in powers of $\Delta\rho_\alpha(\boldsymbol{r}) = \rho_\alpha(\boldsymbol{r}) - \rho_\alpha$ (where ρ_α is the bulk concentration of microions), the total functional is quadratic in the $\rho_\alpha(\boldsymbol{r})$, and the Euler-Lagrange equations resulting from the extremum conditions (12) can be solved analytically [15]. The resulting total effective energy of the polyions is then strictly pair-wise additive, and the effective pair potentials are of the linearly screened DLVO form. The entire procedure is justified only for relatively weak microion inhomogeneities (i.e. $|\Delta\rho_\alpha(\boldsymbol{r})|/\rho_\alpha < 1$), i.e. for low absolute polyion valence $|Z_p|$. If the polyion charge is distributed over a number ν of interaction sites, each carrying a charge $Z_p \, e/\nu$, linear screening may be an adequate approximation for each interaction site. The resulting "Yukawa site" model, where all sites on neighbouring particles interact via a screened Coulomb (or Yukawa) pair potential, has been used to simulate charged rods [32] or charged discs representing clay particles [33].

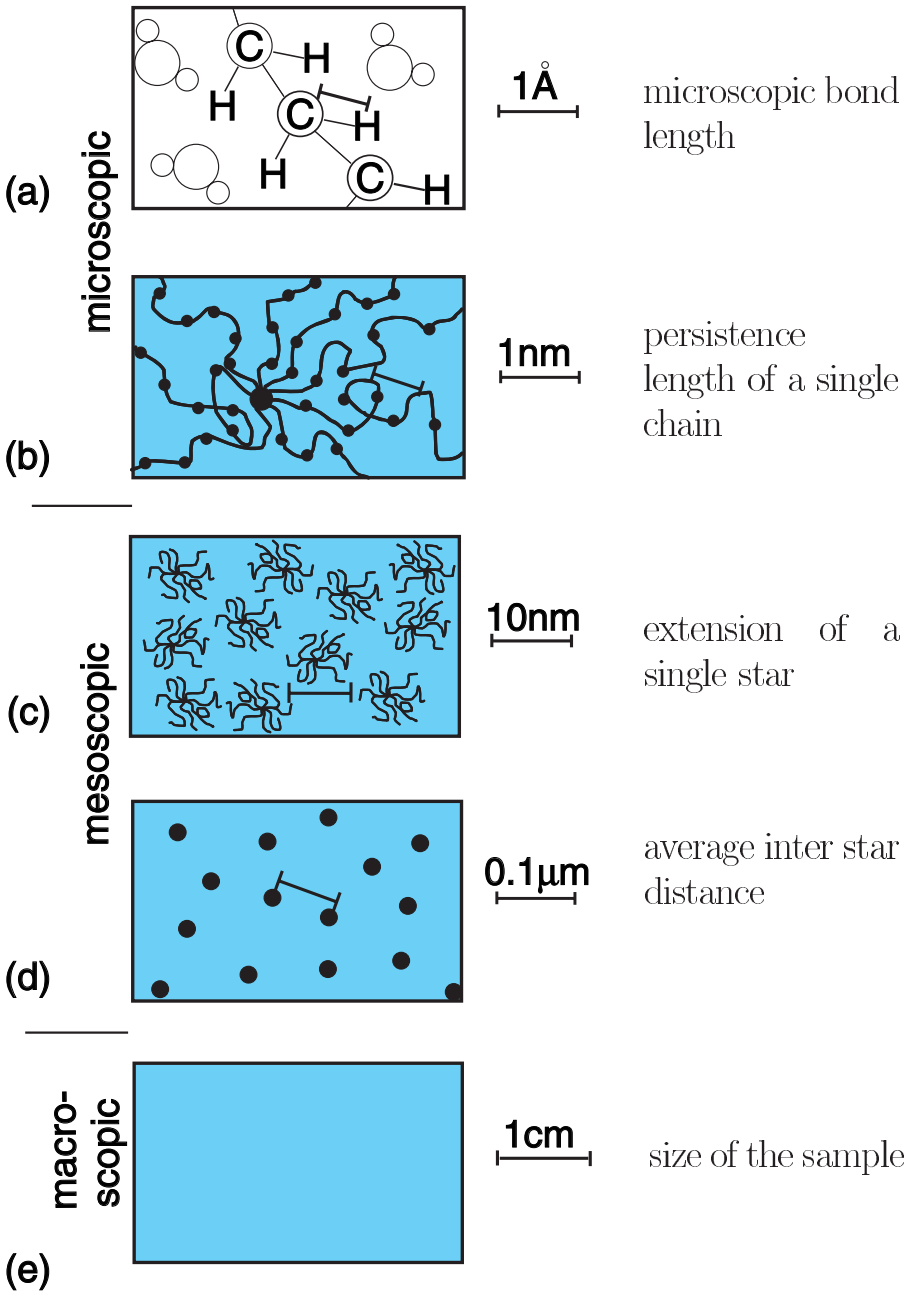


Fig. 1. Star polymer solution on different length scales. (**a**): microscopic picture, water and hydrocarbon chains are shown, the chemical bonds have a range of typically 1Å. (**b**): On a larger scale, the persistence length of the chains is relevant. (**c**): the spatial extension σ of a single polymer star. (**d**): all the coils are point particles on this scale governed by the mean intercoil distance (**e**): size of the macroscopic sample

5 Star Polymers

Star polymers [4] consist of f linear polymer chains which are chemically anchored to a common centre (f is called functionality or arm number). Obviously, linear polymers are a special case of star polymers when $f = 1, 2$ depending whether the end or middle segment is taken as "centre". Dendrimers, on the other hand, can be viewed as iterated star polymers: periodically, any linear chain branches off into n additional chains (n is called degree of branching) which is repeated g times (g is called generation number). For $f > 3$, in contrast to linear chains, star polymers and dendrimers possess a natural centre which serves as an appropriate statistical degree of freedom.

Let us first focus on *star polymers in a good solvent*. A full monomer-resolved computer simulation is completely out of reach of present-day computers: If N is the number of stars and M the number of monomers per chain, a total number of NfM particles has to be simulated, f times more than for a solution of linear chains and fM times more than for simple fluids. The strategy b) of Sect. 2, however, can be efficiently used to make progress. First consider only two stars at fixed separation r and average the force acting on their centres during an ordinary MC or MD simulation of the monomers. Such a simulation involves $2fM$ particles only. A typical simulation snapshot is shown in Fig. 2. This is repeated for different r. By integrating the distance-resolved data for the force, the effective interaction potential $v(r)$ is obtained. This interaction is repulsive, since the presence of another star reduces the number of configurations available to the chains. For small arm numbers $f \leq 10$, the simulation results confirm an effective pair potential of the log-Gauss form:

$$v(r) = \frac{5}{18} k_{\mathrm{B}} T f^{3/2} \begin{cases} -\ln(\frac{r}{\sigma}) + \frac{1}{2\tau^2\sigma^2} & \text{for} \quad r \leq \sigma \ ; \\ \frac{1}{2\tau^2\sigma^2} \exp\left(-\tau^2 \frac{r^2-\sigma^2}{\sigma^2}\right) & \text{for} \quad r > \sigma \ , \end{cases} \tag{33}$$

where σ is the corona diameter of a single star measuring the spatial extent of the monomeric density. For large distances r, the interaction is Gaussian as for linear chains. It then crosses over, at the corona diameter of the star, to a logarithmic behaviour for overlapping coronae as predicted by scaling theory [34] which implies a very mild divergence as $r \to 0^+$. The matching at $r = \sigma$ is done such that the force $-dv/dr$ is continuous. In (33), $\tau(f)$ is known from a fit to computer simulation results; for $f = 2$ we obtain $\tau = 1.03$ in line with a Gaussian potential used for linear chains.

For larger arm numbers, $f > 10$, on the other hand, a geometric blob picture of f cones around the star centre, each containing one linear chain is justified [35]. The effective force for nearly touching coronae decays exponentially with r, the associated decay length is the outermost blob-diameter $2\sigma/\sqrt{f}$. This motivates a log-Yukawa form of $v(r)$ [36]:

$$v(r) = \frac{5}{18} k_B T f^{3/2} \begin{cases} -\ln(\frac{r}{\sigma}) + \frac{1}{1+\sqrt{f}/2} & \text{for} \quad r \leq \sigma \\ \frac{\sigma}{1+\sqrt{f}/2} \frac{\exp(-\sqrt{f}(r-\sigma)/2\sigma)}{r} & \text{for} \quad r > \sigma \end{cases} \tag{34}$$

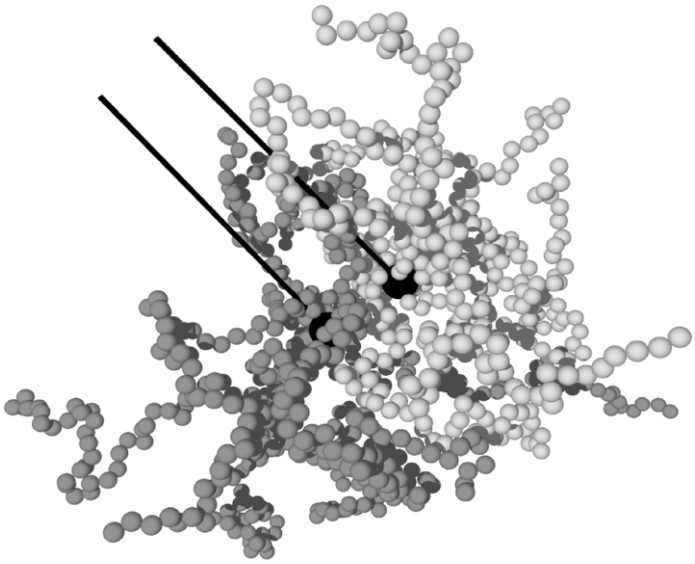


Fig. 2. Typical configuration for two stars with $f = 10$ and $M = 50$ monomers per chain as obtained from a snapshot during a Molecular Dynamics simulation. The distance between the two black lines is the centre-to-centre separation r. By courtesy of A. Jusufi

again matched at the corona diameter $r = \sigma$ such that the force is continuous. This potential was verified in monomer-resolved simulations [37] for a large range of arm numbers.

Using scaling theory and monomer-resolved simulations of a triangular configuration of three stars [38], triplet interactions were shown to be negligibly small outside the corona and at most 11 percent of the pairwise forces for penetrating triplets inside the corona; consequently the effective pair-wise description for the many-body system is adequate provided the number density ρ_s of the stars is not much higher than the overlap density $1/\sigma^3$. Large scale simulations involving many stars were performed using the pair potential of (34) [39, 40]. Due the crossover of $v(r)$ at $r = \sigma$ from a harsh Yukawa to a soft logarithmic behaviour, uncommon structural and thermodynamical properties were obtained. First, the main peak of the liquid structure factor changes non-monotonically with increasing density [40]. Secondly, the bulk phase diagram exhibits [39] a reentrant melting behaviour for $34 < f < 44$ and stable anisotropic crystal lattices. The latter finding has been supported by recent experiments on various block-copolymer micelles [41–43].

Next let us briefly discuss *star polymers in a poor solvent*. The only work in this direction is close to the Θ-point where the chains are weakly interacting. Consequently the resulting effective repulsion is weaker than in good solvent. More quantitatively, an effective potential between two plates is available

within a self-consistent field approach for polymers grafted on flat plates where the grafting density is high and the self-avoidance is weak [44]. This was extended to spherical particles by employing the Derjaguin approximation [45,46] providing an analytical expression for the effective pair potential $v(r)$. In the limit of small core sizes, this expression has been successfully tested against scattering data for $f = 64$ arm stars in a solvent close to Θ conditions [47]. What is still unexplored is a systematic approach for arbitrary solvent quality which continously switches between good solvent quality to the Θ point and beyond.

Much more stretched configurations are achieved for *polyelectrolyte stars* ("porcupines") due to the strong Coulomb repulsion of the charged monomers along the chains. If one brings two polyelectrolyte stars together they hardly interdigitate but retract. A variational analysis [49] for the effective force, which includes Coulomb interactions and entropies of the counterions, reveals that the entropy of the counterions which are inside the coronae of the two polyelectrolyte stars dominates the interaction, confirming an old idea of Pincus [48]. The analytical theory was quantitatively verified by computer simulations with explicit monomers and counterions [49]. Inside the corona, the resulting effective force could be fitted by an inverse-power law $\propto r^{-\gamma}$ where the exponent γ slightly depends on the actual charging conditions but is always around $0.7 - 0.8$. By integration, an effective potential is obtained which stays finite at the origin and behaves inside the corona as $v(r) = v(0) - Cr^{1-\gamma}$ with a positive constant C. However, the actual value $v(0)$ for completely overlapping stars is much larger than k_BT so that significant overlap is rare. Due to the softness of the interaction, similar structural anomalies as obtained for star polymers are expected including a non-monotonic variation of the first peak in the structure factor for increasing density and reentrant melting.

6 Colloids and Polymers: Depletion Interactions

If a sterically-stabilized colloidal particle is brought into a non-adsorbing polymer solution, the latter are depleted in a zone around the colloidal surfaces due to the colloid-polymer repulsion. The width of this zone is of the order of the radius of gyration $d_p/2$ of the polymers. If one now brings two colloidal particles close to each other, the two depletion zones overlap, which brings about a free energy gain of the polymers relative to a situation of non-overlapping zones, resulting in an effective attraction between the colloids, the so-called depletion attraction. Alternatively one can view the attraction arising from an unbalanced osmotic pressure exerted on the colloidal particles by the surrounding polymers.

The simplest model for colloid-polymer mixtures including the depletion effect is the so-called Asakura-Oosawa (AO) [50] or Asakura-Oosawa-Vrij (AOV) [51] model which assumes hard core interactions between the colloids of diameter d_c, further hard-core interactions between the polymers and the

colloids with a range $(d_c + d_p)/2$, but no interaction at all between polymers. The ideality of the polymers is a crucial approximation which is fulfilled only for dilute polymer solutions, but it allows to investigate many of the statistical properties of the AO model analytically. For instance, the effective interaction $v(r)$ between a colloidal pair can be calculated to be the product of the polymer osmotic pressure $P_p = k_B T \rho_p$ and the overlap volume of the two depletion zones consisting of two spherical half-caps. Explicitly it reads

$$\frac{v(r)}{k_B T} = \begin{cases} \infty & \text{for} \quad r \leq d_c \\ \rho_p \frac{\pi}{6}(d_c + d_p)^3 \left[1 - \frac{3r}{2(d_c+d_p)} + \frac{1}{2}\frac{r^3}{(d_c+d_p)^3}\right] & \text{for} \quad d_c < r \leq d_c + d_p \\ 0 & \text{for} \quad r \geq d_c + d_p \end{cases} \quad (35)$$

Furthermore, by a simple geometric consideration, it can be shown that effective triplet and higher-order many-body forces vanish provided the size ratio between colloids and polymers $q = d_p/d_c$ is smaller than 0.154. In this case, the AO model is formally equivalent to an effective one-component system with a short ranged attraction, which immediately opens the way for large-scale simulations.

The phase diagram of the AO model was explored by computer simulations on two different levels: first, one-component calculations using the effective pair potential (35) have been performed [52], which are exact for $q < 0.154$. Secondly, more recently, Dijkstra has simulated the full effective Hamiltonian including effective many-body forces to arbitrary order for $q = 1$ [53]. The emerging phase diagram involves three phases: gas (i.e. colloidal poor), liquid (i.e. colloidal rich) and an fcc colloidal crystal. A liquid phase is stable if the ratio q is larger than $q_c \approx 0.5$.

On the other hand, theoretical progress was made by constructing a free volume theory for the fluid bulk free energies [54] which provides a reliable estimate for the gas-liquid transition. A free-energy density functional for the AO colloid-polymer mixtures, valid for arbitrary inhomogeneous situations, was constructed [55] in the spirit of Rosenfeld's fundamental measure approach [19], which reproduces the effective interaction (35) for a colloid pair and the free volume theory of [54]. This density functional was applied to wetting phenomena of planar walls. A novel type of wetting involving growth of only few colloidal liquid layers on top of the wall as liquid-gas coexistence is approached was predicted by density functional theory [56] and confirmed by computer simulations [53]. This wetting scenario only shows up for ratios larger than q_c, so that one can speculate that it is produced by the intrinsic many-body nature of the effective forces.

Obviously, the AO model has the short-coming of idealized interactions. More realistic models involve a non-zero polymer-polymer interaction and a softer polymer-wall interaction [57]. On the other hand, full two-component

simulations of colloids and polymers were performed [58, 59] where the polymers are defined on a lattice. Clearly these include any effective many-body interactions. A second computationally less demanding technique is to calculate effective pair interactions between a colloid and a polymer first by a monomer-resolved reference simulation. This strategy was followed in the more general context of mixtures of colloids and star polymers for small size ratios q. Supported by theoretical scaling arguments the following pair interaction between a hard-sphere colloid and a star polymer was obtained [60, 61]:

$$v_{cp}(r) = k_B T \Lambda f^{3/2} \left(\frac{d_c}{2r+d_c}\right) \times \begin{cases} -\ln\left(\frac{2r-d_c}{\sigma}\right) + \left(\frac{(2r-d_c)^2}{\sigma^2} - 1\right)((1+4\kappa)/(1+2\kappa)) + \zeta & \text{for } r \le (d_c+\sigma)/2; \\ \zeta\,\mathrm{erfc}(\kappa(2r-d_c)/\sigma)/\mathrm{erfc}(\kappa) & \text{else,} \end{cases} \tag{36}$$

Here, Λ and κ are known parameters depending on the functionality f of the star, $\zeta = \sqrt{\pi}\mathrm{erfc}(\kappa)\exp(\kappa^2)/(\kappa(1+2\kappa^2)$, σ denotes the corona diameter of the star and $\mathrm{erfc}(x)$ is the complementary error function. For $r \to d_c/2$ the potential diverges logarithmically as for the star-star interaction (33). Linear polymer chains are obtained as the special case $f = 2$ where $\Lambda = 0.46$ and $\kappa = 0.58$. The two-component system with effective pair interactions was investigated in detail by further simulation and liquid integral equation theory. For different arm numbers f, the fluid-fluid demixing transition was calculated [60] in good agreement with experimental data. Furthermore, the freezing transitions was discussed. Above a critical arm number of $f_c \approx 10$, fluid-fluid demixing was preempted by freezing [62]. More recently, the fluid-fluid interfacial tension was calculated on the basis of the realistic effective interactions between colloids and linear polymers (being the special case $f = 1$ in (35) [63].

In case of polymer size comparable or larger than the colloidal diameter d_c, effective many-body forces play a significant role. Complementary methods such as monomer-resolved liquid intergral equations methods combined with the PRISM approach [64] or field-theoretic calculations [65] have provided valuable insight into the structure of colloid-polymer mixtures. The limit of large q contains completely different physics, since the colloidal spheres represent then small perturbations for the long polymer chains [59].

7 Conclusions

In conclusion, we have demonstrated that the concept of effective interactions allows large-scale simulations and provides additional insight into the physical mechanisms governing colloidal dispersions and polymer solutions.

The open problems are in the application of effective interaction to dynamical questions both in equilibrum and nonequilibrium [66]. This is much

harder to establish since there is no clean statistical mechanics guideline in this case.

Acknowledgements

I thank J. P. Hansen, E. Allahyarov, J. Dzubiella, A. Jusufi, C. N. Likos, A. A. Louis and A. Lahcen for help and for valuable comments.

References

1. For a recent review, see: T. A. Witten (1999) Insights from soft condensed matter. *Rev. Mod. Phys.* **71**, pp. S367–S373
2. R. J. Hunter (1989) *Foundations of Colloid Science* Volume I, Oxford Science Publications, Clarendon Press, Oxford
3. See e.g.: S. Neser, C. Bechinger, P. Leiderer, T. Palberg (1997) Finite-Size Effects on the Closest Packing of Hard Spheres. *Phys. Rev. Letters* **79**, pp. 2348–2351
4. For a review see: G. S. Grest, L. J. Fetters, J. S. Huang, D. Richter (1996) Star Polymers: Experiment, theory and simulation. *Advances in Chemical Physics* Volume **XCIV**, p. 67
5. J.-P. Hansen, H. Löwen (2002) *Effective interactions for large-scale simulations of complex fluids,* in: Bridging Time Scales: Molecular Simulations for the Next Decade P. Nielaba, M. Mareschal, G. Ciccotti (Eds.), Springer, Berlin, pp. 167–198, ISBN 3-540-44317-7
6. C. N. Likos (2001) Effective interactions in soft condensed matter physics. *Physics Reports* **348**, pp. 267–439
7. L. Belloni (2000) Colloidal interactions. *J. Phys.: Condens. Matter* **12**, pp. R549–R587
8. J. P. Hansen, H. Löwen (2000) Effective interactions between electric double-layers. *Ann. Rev. Phys. Chem.* **51**, pp. 209–242
9. M. Dijkstra (2001) Computer simulations of charge and steric stabilised colloidal suspensions. *Current Opinion in Colloid and Interface Science* **6**, pp. 372–382
10. See e.g. R. van Roij, M. Dijkstra, J. P. Hansen (1999) Phase diagram of charge-stabilized colloidal suspensions: van der Waals instability without attractive forces. *Phys. Rev. E* **59**, pp. 2010–2025
11. H. Löwen (1994) Melting, freezing and colloidal suspensions. *Phys. Reports* **237**, pp. 249–324
12. For an extensive review of classical DFT see R. Evans in Fundamentals of Inhomogeneous Fluids. edited by D. Henderson (Marcel Decker, New York, 1992)
13. J. P. Hansen and E. Smargiassi (1996) in Monte Carlo and Molecular Dynamics of Condensed Matter Systems, edited by K. Binder and G. Ciccotti Societa Italiana di Fisica, Bologna
14. See e.g. G. Galli and M. Parrinello (1991) in Computer Simulations in Materials Science. P. 282, edited by M. Meyer and V. Pontikis Kluwer, Dordrecht
15. H. Löwen, J. P. Hansen, P. A. Madden (1993) Nonlinear counterion screening in colloidal suspensions. *J. Chem. Phys.* **98**, pp. 3275–3289

16. F. Gygi, G. Galli (1995) Real-space adaptive-coordinate electronic-structure calculations. *Phys. Rev. B* **52**, pp. R2229–R2232
17. R. Roth, R. Evans, S. Dietrich (2000) Depletion potential in hard-sphere mixtures: Theory and applications. *Phys. Rev. E* **62**, pp. 5360–5377
18. D. Goulding, S. Melchionna (2001) Accurate calculation of three-body depletion interactions. *Phys. Rev. E* **64**, p. 011403 (1-9)
19. Y. Rosenfeld, M. Schmidt, H. Löwen, P. Tarazona (1997) Fundamental-measure free energy density functional for hard spheres: Dimensional crossover and freezing. *Phys. Rev. E* **55**, pp. 4245-4263
20. H. Löwen (2000) in: "Spatial Statistics and Statistical Physics", edited by K. Mecke and D. Stoyan, Springer Lecture Notes in Physics **554**, pp. 295–331, Berlin
21. A. Lang, C. N. Likos, M. Watzlawek, H. Löwen (2000) Fluid and solid phases of the Gaussian core model. *J. Phys.: Condensed Matter* **12**, pp. 5087–5108
22. C. N. Likos, A. Lang, M. Watzlawek, H. Löwen (2001) Criterion for determining clustering versus reentrant melting behavior for bounded interaction potentials. *Phys. Rev. E* **63**, p. 031206 (1-9)
23. A. A. Louis (2000) Effective potentials for polymers and colloids: beyond the van der Waals picture of fluids? *Philos. Trans. R. Soc. Lond. A* **359**, pp. 939–960
24. M. Schmidt (1999) Density-functional theory for soft interactions by dimensional crossover. *Phys. Rev. E* **60**, pp. R6291–R6294
25. B. V. Derjaguin, L. D. Landau (1948) Acta Physicochim. USSR **14**, 633 (1941); E. J. W. Verwey and J. T. G. Overbeek, *Theory of the Stability of Lyophobic Colloids*. Elsevier, Amsterdam
26. A. Delville, R. J. M. Pellenq (2000) Electrostatic attraction and/or repulsion between charged colloids : a (NVT) Monte-Carlo study, Molecular Simulation **24**, pp. 1–24; R. Messina, C. Holm, K. Kremer (2000) Strong Attraction between Charged Spheres due to Metastable Ionized States. *Phys. Rev. Lett.* **85**, pp. 872–875; T. Terao, T. Nakayama (2001) Charge inversion of colloidal particles in an aqueous solution: Screening by multivalent ions *Phys. Rev. E* **63**, 041401 (1–6); B. Hribar, V. Vlachy (2001) A Monte Carlo Study of Micellar Solutions with a Mixture of Mono- and Trivalent Counterions. *Langmuir* **17**, pp. 2043–2046
27. A. Delville (1999) (N,V,T) Monte Carlo Simulations of the Electrostatic Interaction between Charged Colloids: Finite Size Effects. *J. Phys. Chem. B* **103**, pp. 8296–8300; A. Delville, P. Levitz (2001) Direct Derivation of the Free Energy of Two Charged Lamellar Colloids from (N,V,T) Monte Carlo Simulations. *J. Phys. Chem. B* **105**, pp. 663–667
28. H. Löwen, E. Allahyarov (1998) The role of effective triplet interactions in charged colloidal suspensions. *J. Phys.: Condensed Matter* **10**, pp. 4147–4160
29. R. D. Groot (1991) Ion condensation on solid particles: Theory and simulations. *J. Chem. Phys.* **95**, pp. 9191–9203
30. M. J. Stevens, M. L. Falk, M. O. Robbins (1996) Interactions between charged spherical macroions. *J. Chem. Phys.* **104**, pp. 5209–5219
31. H. Löwen, G. Kramposthuber (1993) Optimal effective pair potential for charged colloids. *Europhys. Lett.* **23**, pp. 673–678
32. H. Löwen (1994) Interaction between charged rod-like colloidal particles. *Phys. Rev. Lett.* **72**, pp. 424–427; (1994) Charged rod-like colloidal suspensions: an ab initio approach. *J. Chem. Phys.* **100**, pp. 6738–6749

33. S. Kutter, J. P. Hansen, M. Sprik, E. Boek (2001) Structure and phase behavior of a model clay dispersion: A molecular-dynamics investigation. *J. Chem. Phys.* **112**, pp. 311–322
34. T. A. Witten, P. A. Pincus (1986) Colloid Stabilization by Long Grafted Polymers. *Macromolecules* **19**, pp. 2509–2513
35. M. Dauod, J. P. Cotton (1982) Star Shaped Polymers: A Model for the Conformation and Its Concentration Dependence. *J. Phys. (Paris)* **43**, pp. 531–538
36. C. N. Likos, H. Löwen, M. Watzlawek, B. Abbas, O. Jucknischke, J. Allgaier, D. Richter (1998) Star Polymers Viewed as Ultrasoft Colloidal Particles. *Phys. Rev. Letters* **80**, pp. 4450–4453
37. A. Jusufi, M. Watzlawek, H. Löwen (1999) Effective Interaction between Star Polymers. *Macromolecules* **32**, pp. 4470–4473
38. C. von Ferber, A. Jusufi, C. N. Likos, H. Löwen, M. Watzlawek (2000) Triplet interactions in star polymer solutions. *Europhys. Journal E* **2**, pp. 311–318
39. M. Watzlawek, C. N. Likos, H. Löwen (1999) Phase Diagram of Star Polymer Solutions. *Phys. Rev. Letters.* **82**, pp. 5289–5292
40. M. Watzlawek, H. Löwen, C. N. Likos (1998) The anomalous structure factor of dense star polymer solutions. *J. Phys.: Condensed Matter* **10**, pp. 8189–8205
41. G. A. McConnell, A. P. Gast (1997) Melting of Ordered Arrays and Shape Transitions in Highly Concentrated Diblock Copolymer Solutions. *Macromolecules* **30**, pp. 435–444
42. T. P. Lodge, J. Bang, M. J. Park, K. Char (2004) Origin of the Thermoreversible fcc-bcc Transition in Block Copolymer Solutions. *Phys. Rev. Lett.* **92**, p. 145501 (1-4)
43. M. Laurati, J. Stellbrink, R. Lund, L. Willner, D. Richter, E. Zaccarelli (2005) Starlike Micelles with Starlike Interactions: A Quantitative Evaluation of Structure Factors and Phase Diagram. *Phys. Rev. Lett.* **94**, p. 195504 (1-4)
44. S. T. Milner, T. A. Witten, M. E. Cates (1988) Theory of the Grafted Polymer Brush. *Macromolecules* **21**, pp. 2610–2619
45. J. Mewis, W. J. Frith, T. A. Strivens, W. B. Russel (1989) The rheology of suspensions containing polymerically stabilized particles. *A. I. Ch. E. J.* **35**, pp. 415–422
46. U. Genz, B. D'Aguanno, J. Mewis, R. Klein (1994) Structure of Sterically Stabilized Colloids. *Langmuir* **10**, pp. 2206–2212
47. C. N. Likos, H. Löwen, A. Poppe, L. Willner, J. Roovers, B. Cubitt, D. Richter (1998) Ordering phenomena of star polymer solutions approaching the Θ state. *Phys. Rev. E* **58**, pp. 6299–6307
48. P. Pincus (1991) Colloid Stabilization with Grafted Polyelectrolytes. *Macromolecules* **24**, pp. 2912–2919
49. A. Jusufi, C. N. Likos, H. Löwen (2002) Conformations and Interactions of Star-Branched Polyelectrolytes. *Phys. Rev. Letters* **88**, p. 018301 (1-4)
50. S. Asakura, F. Oosawa (1954) On Interaction between Two Bodies Immersed in a Solution of Macromolecules. *J. Chem. Phys.* **22**, pp. 1255–1256
51. A. Vrij (1976) Polymers at interfaces and the interactions in colloidal dispersions. *Pure Appl. Chem.* **48**, pp. 471–483
52. M. Dijkstra, J. M. Brader, R. Evans (1999) Phase behaviour and structure of model colloid-polymer mixtures. *J. Phys.: Condensed Matter* **11**, pp. 10079–10106

53. M. Dijkstra, R. van Roij (2002) Entropic Wetting and Many-Body Induced Layering in a Model Colloid-Polymer Mixture. *Phys. Rev. Letters* **89**, p. 208303 (1–4)
54. H. N. W. Lekkerkerker, W. C. K. Poon, P. N. Pusey, A. Stroobants, P. B. Warren (1992) Phase behaviour of colloid+polymer mixtures. *Europhys. Lett.* **20**, pp. 559–564
55. M. Schmidt, H. Löwen, J. M. Brader, R. Evans (2000) Density Functional for a Model Colloid-Polymer Mixture. *Phys. Rev. Letters* **85**, pp. 1934–1937
56. M. Schmidt, H. Löwen, J. M. Brader, R. Evans (2002) Density functional theory for a model colloid-polymer mixture: bulk fluid phases. *J. Phys.: Condensed Matter* **14**, pp. 9353–9382
57. A. A. Louis, R. Finken, J. P. Hansen (2000) Crystallization and phase separation in nonadditive binary hard-sphere mixtures. *Rev. Phys. E* **61**, pp. R1028–R1031
58. E. J. Meijer, D. Frenkel (1995) Computer simulation of colloid-polymer mixtures. *Physica A* **213**, pp. 130–137
59. A. Johner, J. F. Joanny, S. Diez Orrite, J. Bonet Avalos (2001) Gelation and phase separation in colloid-polymer mixtures. *Europhys. Letters* **56**, pp. 549–555
60. J. Dzubiella, A. Jusufi, C. N. Likos, C. von Ferber, H. Löwen, J. Stellbrink, J. Allgaier, D. Richter, A. B. Schofield, P. A. Smith, W. C. K. Poon, P. N. Pusey (2001) Phase separation in star polymer-colloid mixtures. *Phys. Rev. E* **64**, p. 01040 (1–4)
61. A. Jusufi, J. Dzubiella, C. N. Likos, C. von Ferber, H. Löwen (2001) Effective interactions between star polymers and colloidal particles. *J. Phys.: Condensed Matter* **13**, pp. 6177–6194
62. J. Dzubiella, C. N. Likos, H. Löwen (2002) Star-polymers as depleting agents of colloidal hard spheres. *Europhys. Letters.* **58**, pp. 133–139
63. R. L. C. Vink, A. Jusufi, J. Dzubiella, C. N. Likos (2005) Bulk and interfacial properties in colloid-polymer mixtures. *Phys. Rev. E* **72**, p. 030401 (1–4)
64. M. Fuchs, K. S. Schweizer (2001) Macromolecular theory of solvation and structure in mixtures of colloids and polymers. *Phys. Rev. E* **64**, p. 021514 (1–19)
65. A. Hanke, E. Eisenriegler, S. Dietrich (1999) Polymer depletion effects near mesoscopic particles. *Phys. Rev. E* **59**, pp. 6853–6878
66. H. Löwen (2001) Colloidal soft matter under external control. *J. Phys.: Condensed Matter* **13**, pp. R415–R432

Phase Transitions of Model Colloids in External Fields

P. Nielaba[1], S. Sengupta[2], and W. Strepp[1]

[1] Physics Department, University of Konstanz, 78457 Konstanz, Germany
peter.nielaba@uni-konstanz.de
[2] S.N. Bose National Center for Basic Sciences, Calcutta 700098, India



Peter Nielaba

P. Nielaba et al.: *Phase Transitions of Model Colloids in External Fields*, Lect. Notes Phys. **704**, 163–189 (2006)
DOI 10.1007/3-540-35284-8_8
© Springer-Verlag Berlin Heidelberg 2006

1 Introduction 165

2 Elastic Constants from Microscopic Strain Fluctuations 167

2.1 The Method 167
2.2 Two-Dimensional Systems 169

3 Melting of Hard Disks in Two Dimensions 171

4 Phase Transitions of Model Colloids in External Periodic Light Fields 175

4.1 Model and Method 175
4.2 Results and Discussion 180

References 186

The nature of the melting transition for a system of hard disks with translational degrees of freedom in two spatial dimensions has been analyzed by a combination of computer simulation methods and a finite size scaling technique. The behavior of the system is consistent with the predictions of the Kosterlitz-Thouless-Halperin-Nelson-Young (KTHNY) theory.

Hard and soft disks in external periodic (light-) fields show rich phase diagrams including freezing and melting transitions when the density of the system is varied. Monte Carlo simulations for detailed finite size scaling analysis of various thermodynamic quantities like the order parameter, its cumulants etc., have been used in order to map the phase diagram of the system for various values of the density and the amplitude of the external potential.

For hard disks we find clear indication of a reentrant liquid phase over a significant region of the parameter space. The simulations therefore show that the system of hard disks behaves in a fashion similar to charge stabilized colloids which are known to undergo an initial freezing, followed by a remelting transition as the amplitude of the imposed, modulating field produced by crossed laser beams is steadily increased. Detailed analysis of the simulation data shows several features consistent with a recent dislocation unbinding theory of laser induced melting. The differences and similarities of systems with soft potentials (DLVO, $1/r^{12}$, $1/r^{6}$) and the relation to experimental data is analyzed.

Interpreting hard disks as the simplest model of an atomic fluid, quantum effects on the phase diagram are investigated by path integral Monte Carlo simulations.

1 Introduction

Colloidal dispersions can roughly be classified as solutions of mesoscopic solid particles with a stable (non-fluctuating), often spherical, shape, embedded in a molecular fluid solvent. Examples are aqueous suspensions of polystyrene, latex spheres or rods. These particles can be prepared and characterized in a controlled way, and the interactions are tunable. Due to their simplicity, these systems can be considered as the simplest complex fluids and as prototypes of soft matter systems. They are nearly ideal model systems in statistical physics and are thus sometimes "bridges" to the "world" on nanometer length scales.

In addition to this, it is a pleasure to speak in Erice on the properties of such model systems, since several hundreds or even thousands of years ago important discoveries of the properties of simple geometric systems have been made in Sicily by scientists like Archimedes, and in other regions of the world, see Fig. 1. In particular spheres or -in two dimensions- disks, have obviously been of great interest already at that time, and they still are important particles in model systems of present day statistical physics.

During the last decades, crystallization and melting of colloidal suspensions, both in two and three dimensions (2D and 3D), has been a continuous

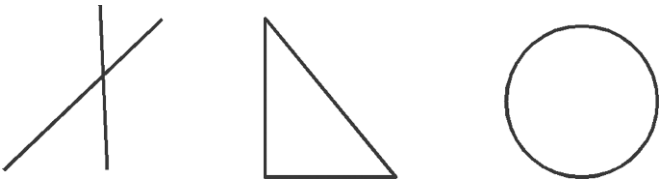


Fig. 1. Simple geometric systems already studied several hundreds of years ago in Sicily and other regions of the world

matter of interest. From the experimental point of view the research mostly focused on the analysis of structure and dynamics of the colloidal systems on different length and time scales through static or dynamic light scattering techniques. On the other hand, in theory the nature of the melting transition in 2D has been controversially discussed at least since the work of Kosterlitz and Thouless (KTHNY-theory). Obviously elastic constants play a crucial role in the solid to liquid phase transition: In 2D the KTHNY theory even claims that the melting process is entirely controlled by the elastic constants. However, both experimental and simulation studies of elastic constants are quite rare. Therefore, the development of tools for the determination of elastic constants in (colloidal) model systems is important.

Here we give an overview on a new method for the computation of elastic constants, applications to various systems, and on the behavior of colloids in external light fields.

The simulational approach makes use of a new coarse-graining procedure which has been successfully tested for a hard disk system. In this technique, elastic strains are calculated from the instantaneous configurations of the particles and averaged over sub-blocks of various linear dimensions $L_b \leq L$ of a system of total linear dimension L. From these data the correlation function of strain fluctuations in the thermodynamic limit can be extracted and the elastic constants then inferred from well known fluctuation formulae.

Hard and soft disks in external periodic (light-) fields show rich phase diagrams including freezing and melting transitions when the density of the system is varied. By Monte Carlo simulation methods we have investigated the interesting phase diagrams of such systems for various values of the density and the amplitude of the external potential.

Systems such as hard disks are not only models for suspensions of colloidal particles (with diameters on the μm scale) which behave according to classical statistical mechanics under all circumstances of practical interest, but also are the simplest possible model for fluids formed by atoms or small molecules (with diameters on the Angstrom scale). Using the latter interpretation, it makes sense to explore the extent to which quantum statistical mechanics rather than classical statistical mechanics is needed to understand their behavior. Quantum effects on the phase diagram are analyzed by path integral Monte Carlo simulations (PIMC).

2 Elastic Constants from Microscopic Strain Fluctuations

2.1 The Method

One is often interested in long length scale and long time scale phenomena in solids (eg. late stage kinetics of solid state phase transformations; motion of domain walls interfaces; fracture; friction etc.). Such phenomena are usually described by continuum theories. Microscopic simulations [1, 2] of finite systems, on the other hand, like molecular dynamics, lattice Boltzmann or Monte Carlo, deal with microscopic variables like the positions and velocities of constituent particles and together with detailed knowledge of interatomic potentials, hope to build up a description of the macro system from a knowledge of these micro variables. How does one recover continuum physics from simulating the dynamics of N particles? This requires a "coarse-graining" procedure in space (for equilibrium) or both space and time for non-equilibrium continuum theories. Over what coarse graining length and time scale does one recover results consistent with continuum theories? We attempted to answer these questions [3] for the simplest nontrivial case, namely, a crystalline solid, (without any point, line or surface defects [4]) in equilibrium, at a non-zero temperature far away from phase transitions.

Consider a general system described by a scalar order parameter $\phi(\mathbf{r})$ in d dimensions and the free energy functional

$$F[\phi(\mathbf{r})] = k_B T \int d^d r \left(\frac{1}{2} r \phi^2 + \frac{1}{2} c \nabla \phi(\mathbf{r})^2 \right). \tag{1}$$

This implies a correlation function $G(q)$ in the high temperature phase $T > T_c$ ($\langle \phi \rangle = 0$) of the Ornstein-Zernike form

$$\beta G(q) = \langle \phi_q \phi_{-q} \rangle = \chi^\infty \frac{1}{1 + (q\xi)^2} \tag{2}$$

(k_B is the Boltzmann constant, T the temperature, and $\beta = 1/k_B T$).

In the computer simulation we determine ϕ averaged within a sub-block of size $L_b \leq L$,

$$\overline{\phi} = L_b^{-d} \int^{L_b} d^d r \ \phi(\mathbf{r}) \tag{3}$$

For a sketch of the geometry see Fig. 2. Then the fluctuations of $\overline{\phi}$ obtained within this block are given by

$$\left\langle \overline{\phi}^2 \right\rangle_{L_b} L_b^d = L_b^{-d} \int^{L_b} d^d r' \, d^d r \ \langle \phi(r) \phi(r') \rangle \tag{4}$$

$$= \int^{L_b} d^d r \, \beta G(r) \equiv \chi^{L_b} \ , \tag{5}$$

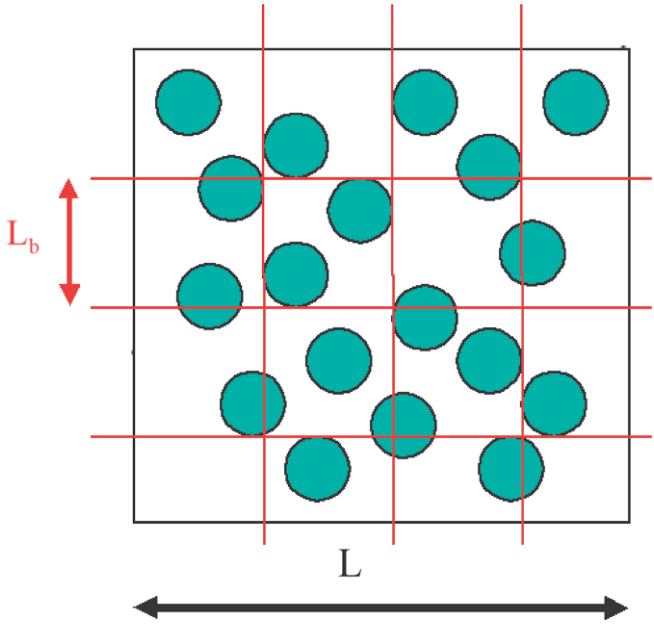


Fig. 2. Schematic picture of the subdivision of a system with linear dimension L into sub-blocks of liner dimension L_B

where $G(r)$ is the inverse Fourier transform of the correlation function defined in (2).

This concept has been proven to be very useful for Ising-type systems. Here we discuss the generalization of the concept to systems where the important degree of freedom is the strain field, which is of tensor character rather than scalar.

Fluctuations of the instantaneous local Lagrangian strain $\epsilon_{ij}(\mathbf{r}, \mathbf{t})$, determined with respect to a static "reference" lattice, are used to obtain accurate estimates of the elastic constants of model solids from atomistic computer simulations. The computed strains are systematically coarse – grained by averaging them within subsystems (of size L_b) of a system (of total size L) in the canonical ensemble, $\bar{\epsilon}_{ij} = L_b^{-d} \int^{L_b} d^d r \epsilon_{ij}(\mathbf{r})$. Using the finite size scaling ideas outlined below we predict the behavior of the fluctuations $\langle \bar{\epsilon}_{ij} \bar{\epsilon}_{kl} \rangle$ as a function of L_b/L and extract elastic constants of the system *in the thermodynamic limit* at nonzero temperature. Our method is simple to implement, efficient and general enough to be able to handle a wide class of model systems including those with singular potentials without any essential modification.

Imagine a system in the constant NVT (canonical) ensemble at a fixed density $\rho = N/V$ evolving in time t. For any "snapshot" of this system taken from this ensemble, the local instantaneous displacement field $\mathbf{u}_\mathbf{R}(t)$ defined over the set of lattice vectors $\{\mathbf{R}\}$ of a reference lattice (at the same density ρ)

is: $\mathbf{u_R}(t) = \mathbf{R}(t) - \mathbf{R}$, where $\mathbf{R}(t)$ is the instantaneous position of the particle tagged by the reference lattice point $\mathbf{R}$. Let us concentrate only on perfect crystalline lattices; if topological defects such as dislocations are present the analysis below needs to be modified. The instantaneous Lagrangian strain tensor ϵ_{ij} defined at $\mathbf{R}$ is then given by [4],

$$\epsilon_{ij} = \frac{1}{2}\left(\frac{\partial u_i}{\partial R_j} + \frac{\partial u_j}{\partial R_i} + \frac{\partial u_i}{\partial R_k}\frac{\partial u_k}{\partial R_j}\right) \tag{6}$$

The strains considered here are always small and so we, hereafter, neglect the non-linear terms in the definition given above for simplicity. The derivatives are required at the reference lattice points $\mathbf{R}$ and can be calculated by any suitable finite difference scheme once $\mathbf{u_R}(t)$ is known. We are now in a position to define coarse grained variables $\bar{\epsilon}_{ij}$ which are simply averages of the strain over a sub-block of size L_b. The fluctuations of these variables then define the size dependent compliance matrices $S^{L_b}_{ijkl} = L_b^d < \bar{\epsilon}_{ij}\bar{\epsilon}_{kl} >$.

2.2 Two-Dimensional Systems

Before proceeding further, we introduce [3] a compact Voigt notation (which replaces a pair of indices ij by one α) appropriate for two dimensional strains - the case considered in this subsection. Using $1 \equiv x$ and $2 \equiv y$ we have,

$$\begin{array}{cccc} ij = & 11 & 22 & 12 \\ \alpha = & 1 & 2 & 3 \end{array} \tag{7}$$

The nonzero components of the compliance matrix are

$$\begin{aligned} S^{L_b}_{11} &= L_b^2\langle\bar{\epsilon}_{xx}\bar{\epsilon}_{xx}\rangle = S^{L_b}_{22} \\ S^{L_b}_{12} &= L_b^2\langle\bar{\epsilon}_{xx}\bar{\epsilon}_{yy}\rangle = S^{L_b}_{21} \\ S^{L_b}_{33} &= 4L_b^2\langle\bar{\epsilon}_{xy}\bar{\epsilon}_{xy}\rangle \end{aligned} \tag{8}$$

It is useful to define the following linear combinations

$$\begin{aligned} S^{L_b}_{++} &= L_b^2\langle\bar{\epsilon}_{+}\bar{\epsilon}_{+}\rangle = 2(S^{L_b}_{11} + S^{L_b}_{12}) \\ S^{L_b}_{--} &= L_b^2\langle\bar{\epsilon}_{-}\bar{\epsilon}_{-}\rangle = 2(S^{L_b}_{11} - S^{L_b}_{12}) \end{aligned} \tag{9}$$

where $\bar{\epsilon}_{+} = \bar{\epsilon}_{xx} + \bar{\epsilon}_{yy}$ and $\bar{\epsilon}_{-} = \bar{\epsilon}_{xx} - \bar{\epsilon}_{yy}$. Once the block averaged strains $\bar{\epsilon}_{ij}$ are obtained, it is straight-forward to calculate these fluctuations (for each value of L_b).

The block averaged compliance matrices approach the limiting values for large L_b [3],

$$S^{L_b}_{\gamma\gamma} = S^{\infty}_{\gamma\gamma}\left[\Psi_2(xL/\xi) - \left(\Psi_2(L/\xi) - C\left(\frac{a}{L}\right)^2\right)x^2\right] + \mathcal{O}(x^4)\,, \tag{10}$$

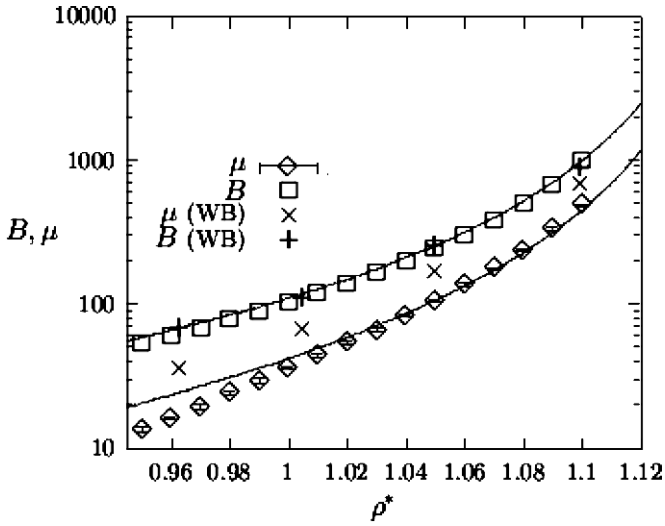


Fig. 3. The bulk (B) and shear (μ) moduli in units of k_BT/σ^2 for the hard disk solid. Our results for B (μ) are given by squares (diamonds). The values for the corresponding quantities from [6] are given by + and ×. The line through the bulk modulus values is the analytical expression obtained from the free volume prediction for the pressure. The line through our shear modulus values is obtained from the free volume bulk modulus using the Cauchy relation $\mu = B/2 - p$. From Sengupta et al. [3, 7]

where γ takes the values $+,-$ or 3 and $x = L_b/L$ and we have suppressed subscripts on the correlation length ξ and the constant C for clarity. $\Psi_2(\alpha)$ is defined as,

$$\Psi_2(\alpha) = \frac{2}{\pi}\alpha^2 \int_0^1 \int_0^1 dxdy\ K_0(\alpha\sqrt{x^2+y^2})\ , \tag{11}$$

where K_0 is a Bessel function. The above equation (10) can now be used to obtain the *system size independent* quantities $S^{\infty}_{\alpha\beta}$, ξ and C.

Once the finite size scaled compliances are obtained the elastic constants viz. the Bulk modulus $B = \rho\partial p/\partial\rho$ and the shear modulus μ are obtained simply using the formulae [5]

$$\beta B = \frac{1}{2S^{\infty}_{++}} \qquad \beta\mu = \frac{1}{2S^{\infty}_{--}} - \beta p \qquad \beta\mu = \frac{1}{2S^{\infty}_{33}} - \beta p \tag{12}$$

where we assume that the system is under an uniform hydrostatic pressure p.

As an example we present our results for elastic constants of the hard disk system in Fig. 3. The two expressions for the shear modulus in (12) give almost identical results and this gives us confidence about the internal consistency of our method. We have also compared our results to those of Wojciechowski et al. [6]. We find that while their values of the pressure and bulk modulus

are in good agreement with ours (and with free volume theory) they grossly overestimate the shear modulus. This is probably due to the extremely small size of their systems and/or insufficient averaging. Our results for the sub-block analysis shows that finite size effects are non-trivial for elastic strain fluctuations and they cannot be evaluated by varying the total size of the system from $N = 24$ to $N = 90$, an interval which is less than half of a decade. One immediate consequence of our results is that the Cauchy relation [6] $\mu = B/2 - p^*$ is seen to be valid up to $\pm 15\%$ over the entire density range we studied though there is a systematic deviation which changes sign going from negative for small densities to positive as the density is increased. This is in agreement with the usual situation in a variety of real systems [8] with central potentials and highly symmetric lattices. We have also compared our estimates for the elastic constants with the density functional theory (DFT) of Ryzhov and Tareyeva [9]. We find that both the bulk and the shear moduli are grossly overestimated by them – sometimes by as much as 100%.

This method has been applied to models of colloidal systems containing quenched point impurities and to colloidal mixtures. In order to analyze the effect of point impurities on the elastic properties of a triangular system of hard disks, we applied our method to the case of quenched impurities with various concentrations [10]. We note a substantial hardening of the material.

We also have considered colloidal mixtures with different diameters in two [10] and three [11] dimensions and the composition dependence of their phase behavior and the elastic properties [12, 13]. Interesting high pressure structures are found for colloidal mixtures in two and three dimensions.

A priori it is not obvious if such systems are softer or harder compared to the corresponding monodisperse systems, and a systematic study is required in order to design materials with well defined elastic properties at a later stage. Besides this, already in two spatial dimensions interesting structures have been found which significantly deviate from the traditional triangular lattice for certain diameter ratios.

Anisotropic situations caused by thin films of thickness D have been considered as well [14], and interesting modifications of the solid structures close to walls are found.

By application of such a method to configurations obtained experimentally by video microscopy methods it was possible to analyze precisely experimental results on the elasticity of colloidal systems [15].

3 Melting of Hard Disks in Two Dimensions

One of the first continuous systems to be studied by computer simulations [16] is the system of hard disks of diameter σ interacting with the two body potential,

$$\phi(r) = \begin{cases} \infty & r \leq \sigma \\ 0 & r > \sigma \end{cases} \tag{13}$$

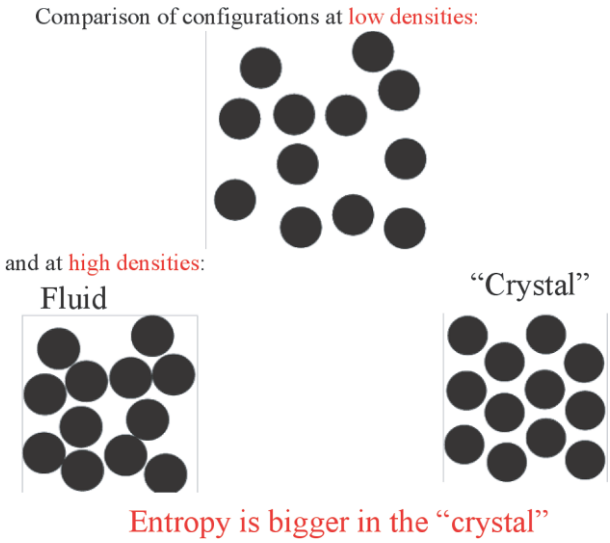


Fig. 4. Schematic picture of configurations in hard disk systems at low and high densities

where, σ (taken to be 1 in the rest of the paper) the hard disk diameter, sets the length scale for the system and the energy scale is set by $k_BT = 1$. Despite its simplicity, this system was shown to undergo a phase transition from solid to liquid as the density ρ was decreased. The reason for this phase transition with increasing density roughly is the higher entropy of the solid structure at high densities, where the particles can fluctuate around their average lattice positions, compared to the entropy of a liquid, in which the particles have positional disorder, but are essentially locked and cannot explore much phase space anymore. For a schematic picture of configurations at low and high densities see Fig. 4. The nature of this phase transition, however, is still being debated. Early simulations [16] always found strong first order transitions. As computational power increased the observed strength of the first order transition progressively decreased! Using sophisticated techniques Lee and Strandburg [17] and Zollweg and Chester [18] found evidence for, at best, a weak first order transition. A first order transition has also been predicted by theoretical approaches based on density functional theory [9, 19]. On the other hand, recent simulations of hard disks [20] find evidence for a KTHNY transition [21] from liquid to a hexatic phase, with orientational but no translational order, at $\rho = 0.899$. Nothing could be ascertained, however, about the expected transition from the hexatic phase to the crystalline solid at higher densities because the computations became prohibitively expensive. The solid to hexatic melting transition was estimated to occur at a density $\rho_c \geq 0.91$. A priori, it is difficult to assess why various simulations give contradicting results

concerning the order of the transition. In [7] we took an approach, complementary to Jaster's [20], and investigated the melting transition of the solid phase. We showed that the hard disk solid is unstable to perturbations which attempt to produce free dislocations leading to a solid → hexatic transition in accordance with KTHNY theory [21] and recent experiments in colloidal systems [22]. Though this has been attempted in the past [6, 23], numerical difficulties, especially with regard to equilibration of defect degrees of freedom, makes this task highly challenging. The elastic Hamiltonian for hard disks is given by, $F = -p\epsilon_+^2 + (B/2)\epsilon_+^2 + (\mu + p)(\epsilon_-^2/2 + 2\epsilon_{xy}^2)$, where B is the bulk modulus. The quantity $\mu_{eff} = \mu + p$ is the "effective" shear modulus (the slope of the shear stress vs shear strain curve) and p is the pressure.

The KTHNY-theory [21] is presented usually for a 2-d triangular solid under *zero external stress*. It is shown that the dimensionless Young's modulus of a two-dimensional solid, $K = (8/\sqrt{3}\rho)(\mu/\{1 + \mu/(\lambda + \mu)\})$, where μ and λ are the Lamé constants, depends on the fugacity of dislocation pairs, $y = \exp(-E_c)$, where E_c is the core energy of the dislocation, and the "coarse-graining" length scale l. This dependence is expressed in the form of the following coupled recursion relations for the renormalization of K and y:

$$\frac{\partial K^{-1}}{\partial l} = 3\pi y^2 e^{\frac{K}{8\pi}} \left[\frac{1}{2} I_0\left(\frac{K}{8\pi}\right) - \frac{1}{4} I_1\left(\frac{K}{8\pi}\right)\right], \tag{14}$$

$$\frac{\partial y}{\partial l} = \left(2 - \frac{K}{8\pi}\right) y + 2\pi y^2 e^{\frac{K}{16\pi}} I_0\left(\frac{K}{8\pi}\right) .$$

where I_0 and I_1 are Bessel functions. The thermodynamic value is recovered by taking the limit $l \to \infty$.

We see in Fig. 5 that the trajectories in y-K plane can be classified in two classes, namely those for which $y \to 0$ as $l \to \infty$ (ordered phase) and those for which $y \to \infty$ as $l \to \infty$ (disordered phase). These two classes of flows are separated by lines called the separatrix. The transition temperature T_c (or ρ_c) is given by the intersection of the separatrix with the line of initial conditions $K(\rho, T)$ and $y = \exp(-E_c(K))$ where $E_c \sim cK/16\pi$. The disordered phase is a phase where free dislocations proliferate. Proliferation of dislocations however *does not* produce a liquid, rather a liquid crystalline phase called a "hexatic" with quasi-long ranged (QLR) orientational order but short ranged positional order. A *second* KTHNY transition destroys QLR orientational order and takes the hexatic to the liquid phase by the proliferation of "disclinations" (scalar charges). Apart from T_c there are several universal predictions from KTHNY-theory, for example, the order parameter correlation length and susceptibility has essential singularities ($\sim e^{bt^{-\nu}}, t \equiv T/T_c - 1$) near T_c. All these predictions can, in principle, be checked in simulations [20].

One way to circumvent the problem of large finite size effects and slow relaxation due to diverging correlation lengths is to simulate a system which is constrained to remain defect (dislocation) free and, as it turns out, without a phase transition. Surprisingly, using this data it is possible to predict the

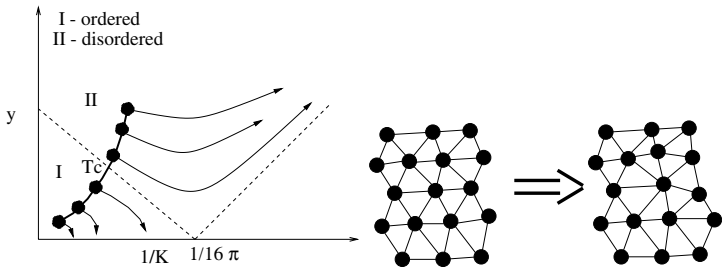


Fig. 5. *Left*: Schematic flows of the coupling constant K and the defect fugacity y under the action of the KTHNY recursion relations. The dashed line is the separatrix whose intersection with the line of initial state (solid line connecting filled circles, $y(T, l = 0), K^{-1}(T, l = 0)$) determines the transition point T_c *Right*: Typical move which attempts to change the coordination number and therefore the local connectivity around the central particle. Such moves were rejected in our simulation

expected equilibrium behavior of the unconstrained system. The simulation [7] is always started from a perfect triangular lattice which fits into our box – the size of the box determining the density. Once a regular MC move is about to be accepted, we perform a *local* Delaunay triangulation involving the moved disk and its nearest and next nearest neighbors. We compare the connectivity of this Delaunay triangulation with that of the reference lattice (a copy of the initial state) around the same particle. If any old bond is broken and a new bond formed (Fig. 5) we reject the move since one can show that this is equivalent to a dislocation – antidislocation pair separated by one lattice constant involving dislocations of the smallest Burger's vector.

Microscopic strains $\epsilon_{ij}(\mathbf{R})$ can be calculated now for every reference lattice point $\mathbf{R}$. Next, we coarse grain (average) the microscopic strains within a subbox of size L_b and calculate the (L_b dependent) quantities [3],

$$S_{++}^{L_b} = L_b^2 \langle \bar{\epsilon}_+ \bar{\epsilon}_+ \rangle, \; S_{--}^{L_b} = L_b^2 \langle \bar{\epsilon}_- \bar{\epsilon}_- \rangle, \; S_{33}^{L_b} = 4L_b^2 \langle \bar{\epsilon}_{xy} \bar{\epsilon}_{xy} \rangle \tag{15}$$

The elastic constants in the thermodynamic limit are obtained from, the set: $B = 1/(2S_{++}^{\infty})$ and $\mu_{eff} = 1/(2S_{--}^{\infty}) = 1/(2S_{33}^{\infty})$. We obtain highly accurate values of the unrenormalized coupling constant K and the defect fugacity y which can be used as inputs to the KTHNY recursion relations. Numerical solution of these recursion relations then yields the renormalized coupling K_R and hence the density and pressure of the solid to hexatic melting transition.

We can draw a few very precise conclusions from our results. Firstly, a solid without dislocations is stable against fluctuations of the amplitude of the solid order parameter and against long wavelength phonons. So any melting transition mediated by phonon or amplitude fluctuations is ruled out in our system. Secondly, the core energy $E_c > 2.7$ at the transition so KTHNY perturbation theory is valid though numerical values of nonuniversal quantities may depend on the order of the perturbation analysis. Thirdly, solution of the recursion relations shows that a KTHNY transition at $p_c = 9.39$ *preempts*

the first order transition at $p_1 = 9.2$. Since these transitions, as well as the hexatic -liquid KTHNY transition lie so close to each other, the effect of, as yet unknown, higher order corrections to the recursion relations may need to be examined in the future. Due to this caveat, our conclusion that a hexatic phase exists over some region of density exceeding $\rho = 0.899$ still must be taken as preliminary [24]. Also, in actual simulations, cross over effects near the bicritical point, where two critical lines corresponding to the liquid-hexatic and hexatic-solid transitions meet a first order liquid-solid line may complicate the analysis of the data, which may, in part, explain the confusion which persists in the literature on this subject.

4 Phase Transitions of Model Colloids in External Periodic Light Fields

4.1 Model and Method

The liquid-solid transition in two dimensional systems of particles under the influence of external modulating potentials has recently attracted a fair amount of attention from experiments [22, 25–30], density functional theory [31, 32], dislocation unbinding calculations [21, 33] and computer simulations [34–38]. This is partly due to the fact that well controlled, clean experiments can be performed using colloidal particles [39] confined between glass plates, producing essentially a two-dimensional system. These systems are subjected to a spatially periodic electromagnetic field generated by interfering two crossed laser beams. This field acts on the particles like a commensurate, one dimensional, modulating potential, see Fig. 6.

One of the more surprising results of these studies is the fact that there exist regions in the phase diagram over which one observes reentrant [28–30] freezing/melting behavior. A schematic phase diagram for hard disk systems is shown in Fig. 7.

As a function of the laser field intensity the system first freezes from a modulated liquid to a two dimensional triangular solid. This effect is described in the literature as *Laser Induced Freezing* (LIF) and is shown in Fig. 8.

A further increase of the intensity confines the particles strongly within the troughs of the external potential, suppressing fluctuations perpendicular to the troughs, which leads to an uncoupling of neighboring troughs and to remelting, see Fig. 9. This effect is described in the literature as *Laser Induced Melting* (LIM) and is shown in Fig. 9.

Based on these considerations we therefore expect an influence of the range of the particle potential on the width of the freezing and/or reentrance region. In particular, since the fluctuations of the particles perpendicular to the troughs (see the argument above) are getting less important for longer ranged potentials, we expect the reentrance region to be smaller (or even vanishing) for long ranged potentials, see Fig. 10.

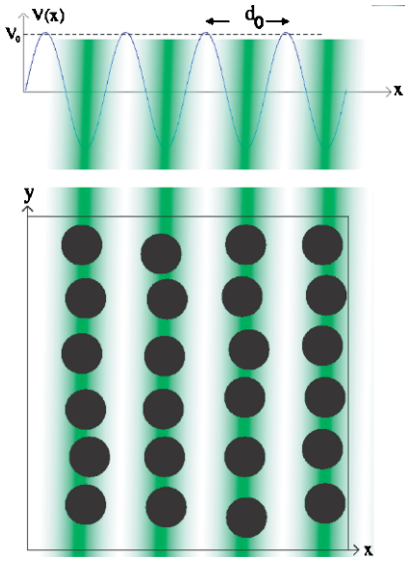


Fig. 6. Schematic picture of the (two-dimensional) system in an external periodic field, periodic in x-direction and with amplitude V_0 and wave length d_0-top view

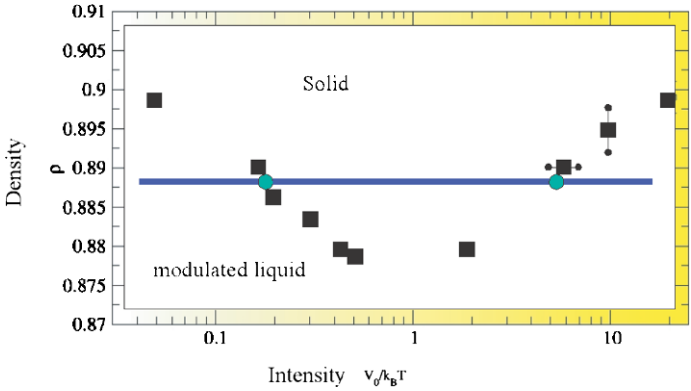


Fig. 7. Schematic phase diagram for hard disk systems in external periodic fields of amplitude V_0

To clarify the situation, a comparative study [47] of the effect of the range of the interaction potentials on the reentrance region by computer simulations for different types of particle potentials [40–43] has been done, with particular focus on the dependence of the width of the freezing/reentrance region on the particle potential. In addition, experimental results on colloidal particles [30] are compared with our data.

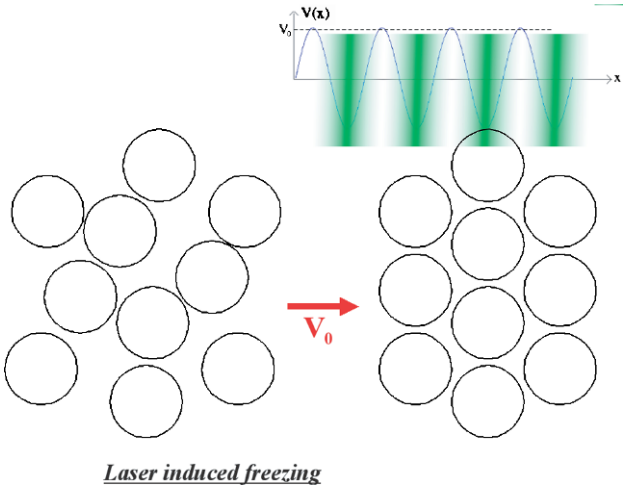


Fig. 8. Schematic picture of the *Laser Induced Freezing*-effect on a system in an external periodic field of amplitude V_0

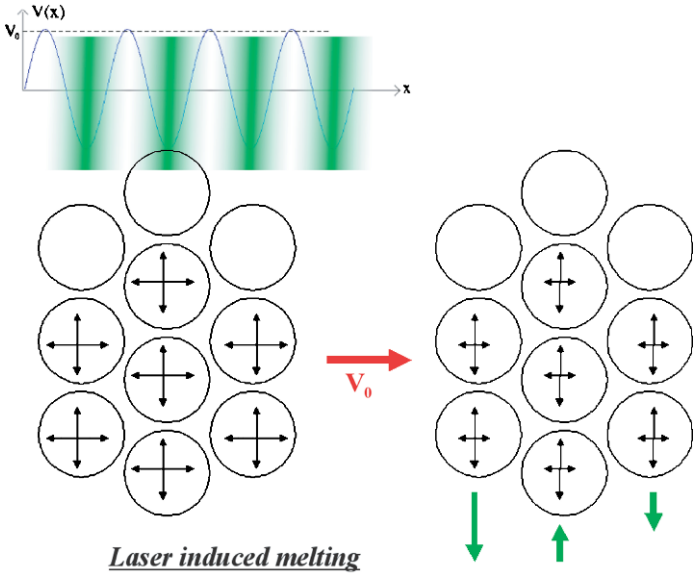


Fig. 9. Schematic picture of the *Laser Induced Melting*-effect on a system in an external periodic field of amplitude V_0

The Model

We study a system of N particles in a two dimensional rectangular box of size $S_x \times S_y$ ($S_x/S_y = \sqrt{3}/2$) and periodic boundary conditions. The particles

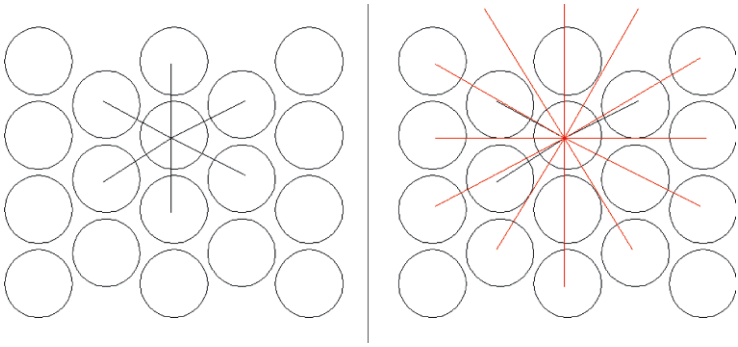


Fig. 10. Schematic picture of the particles interacting with a given particle for hard disk interaction (*left side*) and longer ranged interactions

interact by a pair potential $\phi(r_{ij})$ [39], where r_{ij} is the distance between particle i and j.

We use the following potentials:

- hard disks with diameter σ [40]:

$$\phi(r_{ij}) = \begin{cases} \infty & r_{ij} \leq \sigma \\ 0 & r_{ij} > \sigma \end{cases} \tag{16}$$

  Here we use a reduced density: $\rho^* = \rho\sigma^2$.
- DLVO (Derjaguin, Landau, Vervey and Overbeek) potential [41]:

$$\phi(r_{ij}) = \frac{(Z^*e)^2}{4\pi\epsilon_0\epsilon_r} \left(\frac{\exp(\kappa R)}{1+\kappa R} \right)^2 \frac{\exp(-\kappa r_{ij})}{r_{ij}} \tag{17}$$

  where R is the radius of the particles, $\kappa = \sqrt{\frac{e^2}{\epsilon_0\epsilon_r k_B T} \sum_i n_i z_i^2}$ the inverse Debye screening length, Z^* the effective surface charge, and ϵ_r the dielectric constant of water. We use $\epsilon_r = 78$, $2R = 1.07$ µm and $Z^* = 7800$. The temperature is chosen as in experimental setups: $T = 293.15\,\mathrm{K}$ (i.e. 20°C), and the particle density such that the particle spacing in an ideal lattice is: $a_s = 2.52578$ µm. The different values for the reduced inverse Debye screening length κa_s, which is the dominant parameter here, are obtained by varying κ as required. In addition, we use a cut-off radius r_c: $\phi(r > r_c) = 0$, where r_c obeys the condition: $\phi(r_c) = 0.001\ k_B T$.
- algebraic potentials [42, 43]:

$$\phi(r_{ij}) = k_B T / r^n \tag{18}$$

  where we examine $n = 12$ and $n = 6$. We chose $r_c = 2$ for $n = 12$, and $r_c = \sqrt{10}$ for $n = 6$.

The forces of the laser field on a particle with coordinates (x, y) is modelled as follows:

$$V(x, y) = V_0 \sin(2\pi x/d_0) \tag{19}$$

The constant d_0 in (19) is chosen such that, for a density $\rho = N/(S_x S_y)$, the modulation is commensurate to a triangular lattice of particles with nearest neighbor distance a_s: $d_0 = a_s\sqrt{3}/2$.

The main parameters which define our systems are ρ^* or ρ (or κa_s for the DLVO potential) and the reduced potential strength $V_0/k_BT = V_0^*$.

Observables

In order to obtain thermodynamic quantities for a range of system sizes, we use subsystems of size $L_x \times L_y$, where $L_y = La_s$ and $L_x = L_y\sqrt{3}/2 = Ld_0$.

In the presence of the external potential, we use the positional order parameter ψ_{G_1} to detect the phase transition between modulated liquid (at low density) and crystal (at high density):

$$\psi_{G_1} = \left| \frac{1}{N} \sum_{j=1}^{N} \exp(i\boldsymbol{G}_1 \cdot \boldsymbol{r}_j) \right| \tag{20}$$

where $\boldsymbol{r}_j$ is the position vector of the j^{th} particle. $\boldsymbol{G}_1$ is one of the six smallest reciprocal lattice vectors of the two-dimensional triangular lattice, pointing in a direction at an angle of $\pi/3$ to the x-axis.

In zero external field the crystal can rotate easily, and we therefore use an orientation-corrected version of ψ_{G_1} which we denote by $\tilde{\psi}_{G_1}$. The orientation of the crystal can be extracted from the phase information of the orientational order parameter ψ_6. For a particle j located at $\boldsymbol{r}_j$ we define the local orientational order:

$$\psi_{6,j} = \frac{1}{N_b} \sum_{l=1}^{N_b} e^{i6\theta_{lj}}$$

where N_b is the number of nearest neighbors, and θ_{lj} the angle between the axis $\boldsymbol{r}_l - \boldsymbol{r}_j$ and an arbitrary reference axis. For the total system we set:

$$\psi_6 = \left| \frac{1}{N} \sum_{j=1}^{N} \psi_{6,j} \right|$$

Properties of these order parameters are described in [41].

Based on these order parameters, we have determined phase transition points by the cumulant intersection method. The fourth order cumulant U_L of the order parameter distribution is given by [44]:

$$U_L(V_0^*, \rho^*|\rho|\kappa a_s) = 1 - \frac{\langle \psi_x^4 \rangle_L}{3\ \langle \psi_x^2 \rangle_L^2} \tag{21}$$

where ψ_x refers to one of the order parameters described above. In order to distinguish between the cumulants of ψ_6 and ψ_{G_1}, we denote them with $U_{L,6}$ and $U_{L,G}$, respectively. In the liquid (short ranged order) $U_L \to 1/3$ and in the solid (long range order) $U_L \to 2/3$ for $L \to \infty$. In case of a continuous transition close to the transition point the cumulant is only a function of the ratio of the system size $\approx La_s$ and the correlation length ξ: $U_L(La_s/\xi)$. Since ξ diverges at the critical point the cumulants for different system sizes intersect in one point: $U_{L_1}(0) = U_{L_2}(0) = U^*$. U^* is a non-trivial value, i.e. $U^* \neq 1/3$ and $U^* \neq 2/3$. Even for first order transitions these cumulants intersect [45] though the value U^* of U_L at the intersection is not universal any more. The intersection point can, therefore, be taken as the phase boundary regardless of the order of the transition. In case of a KTHNY-transition we expect a merging of the cumulants at the onset of the ordered phase, rather than an intersection point.

In all scenarios the cumulant intersection- or merging points should reliably detect the phase transition. Therefore, in order to map the phase diagram, we systematically vary the system parameters V_0^* and ρ^* or ρ (or κa_s) and identify the cumulant intersection- or merging points with the phase boundary.

Numerical Details

Monte Carlo (MC) simulations [1, 2, 46] are done in the NVT ensemble, the total system size is $N = 1024$ or $N = 400$. A typical simulation run with 10^7 Monte Carlo steps (MCS) per particle (including 3 $\times 10^6$ MCS for relaxation) took about 50 CPU hours on a PII/500 MHz PC.

4.2 Results and Discussion

Phase Diagrams

First we present the phase diagrams for systems with various interaction potentials: hard disks in Fig. 11 [40], DLVO in Fig. 12 [41], $1/r^{12}$ in Fig. 13

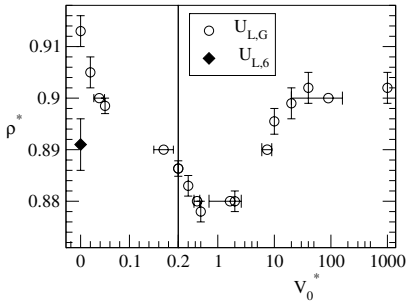


Fig. 11. Phase diagram of hard disks ($N = 1024$) [40]. At $V_0^* = 0.2$ a change from a linear to a logarithmic scale on the x-axis is indicated by a vertical line

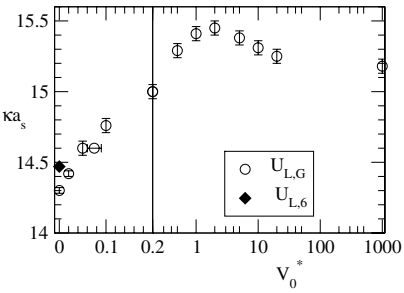


Fig. 12. Phase diagram in case of the DLVO interaction potential ($N = 1024$) [41]. At $V_0^* = 0.2$ a change from a linear to a logarithmic scale on the x-axis is indicated by a vertical line

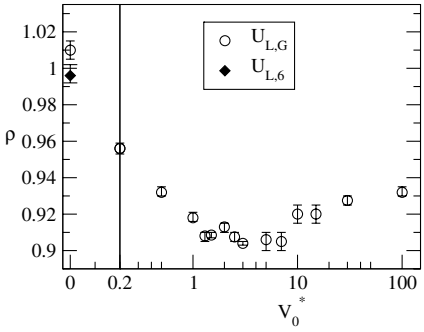


Fig. 13. Phase diagram in case of the $1/r^{12}$ potential ($N = 1024$ particles, except: $N = 400$ at $V_0^* = 2.5, 5, 7$) [42, 43]. At $V_0^* = 0.2$ a change from a linear to a logarithmic scale on the x-axis is indicated by a vertical line

[42, 43], and $1/r^6$ in Fig. 14. There are two transition points marked at zero external potential: one calculated from $\tilde{\psi}_{G_1}$ and one from ψ_6. The reason is that there are two possible types of order: orientational order, to which ψ_6 is sensitive, and positional order, which is detected by $\tilde{\psi}_{G_1}$. The transitions for $V_0^* \neq 0$, extracted from ψ_{G_1}, for $V_0^* \to 0$ converge to the transition calculated from $\tilde{\psi}_{G_1}$. We thus define the phase boundary to be spanned by the $\tilde{\psi}_{G_1}$ transition at $V_0^* = 0$, and by the ψ_{G_1} transitions for $V_0^* \neq 0$.

Generally, beginning at zero external potential, the phase boundary bends down to lower densities for growing V_0^* and reaches a minimum at $V_0^* = 1 \ldots 2$. Increasing V_0^* further, it turns back to higher densities and finally saturates at $V_0^* \geq 50$. This "turning-back" is clearly visible in all cases except for the $1/r^6$ potential, where it seems to be less pronounced.

To quantify these observations, we define two dimensionless numbers Δ_f and Δ_r for the width of the freezing- and reentrance region:

$$\Delta_{f,r} = \delta\rho_{f,r}/\rho_0 \tag{22}$$

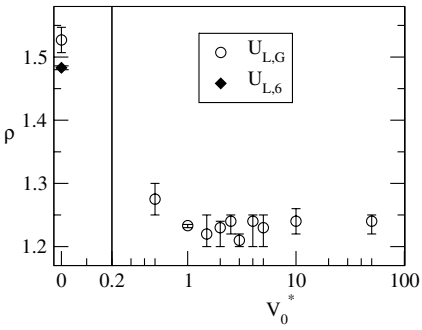


Fig. 14. Phase diagram in case of the $1/r^6$ potential ($N = 400$) [47]. At $V_0^* = 0.2$ a change from a linear to a logarithmic scale on the x-axis is indicated by a vertical line

Table 1. Relative density difference (34),(35) $\Delta_{f,r}$ of the freezing and reentrance region and ratio Δ_f/Δ_r for various interaction potentials, and experimental data for colloids

	hard disk	DLVO	$1/r^{12}$	$1/r^6$	colloids [28]
Δ_f	0.043	0.13	0.11	0.20	0.39
Δ_r	0.028	0.028	0.029	0.016	0.10
Δ_f/Δ_r	1.5	4.6	3.8	12.5	3.9

where $\delta\rho_{f,r} = \rho_{0,\infty} - \rho_{min}$, see Fig. 15. In case of the DLVO potential, using $\rho = (2\kappa^2)/((\kappa a_s)^2\sqrt{3})$ and assuming κ as constant and a_s as variable (which is justified since only the product κa_s is important, whereas the separate values of κ and a_s are not), we can transform (22) into:

$$\Delta_{f,r} = \frac{(\kappa a_s)_0^2}{(\kappa a_s)_{0,\infty}^2} - \frac{(\kappa a_s)_0^2}{(\kappa a_s)_{min}^2} \tag{23}$$

The values of Δ_f, Δ_r and the ratio Δ_f/Δ_r for the different systems are summarized in Table 1. The first four columns contain the values of our simulations, ordered from short range (left) to long range interaction (right). The last column contains the values obtained in an experiment with colloidal

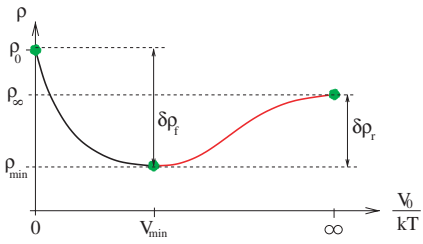


Fig. 15. Schematic picture of the phase diagram

particles [28]. Analyzing the simulations, we roughly observe an increase of the freezing region (Δ_f) when going from short to long range interactions, whereas the reentrance region (Δ_r) is nearly constant for the first three potentials. Only in case of the $1/r^6$-potential a drop of Δ_r is found due to the long range nature of the pair interaction.

The experimental data in the last column shows freezing- and reentrance regions which are about three times bigger than those of the corresponding DLVO simulation, the ratio Δ_f/Δ_r of the freezing and reentrance regions however is nearly identical. We have also done simulations with slightly altered parameters, i.e. using particles with diameter $2R = 3\mu$m, effective surface charge $Z^* = 20000$ and $a_s = 8\mu$m, to match the experiments in [30] as close as possible. Only a small shift of the phase diagram was found by $\Delta(\kappa a_s) \approx 0.35$ towards higher values of κa_s, and a negligible change in Δ_f and Δ_r. The differences between simulation and experiment may be due to the presence of many-body interactions in the experiment, which are not treated properly by the DLVO interactions used in the simulation (see [48]).

In summary, we have presented the phase diagrams of two dimensional systems of hard [40] and various soft disks [41–43] in an external sinusoidal potential. We find an increase of the freezing region with the range of particle interaction, and a decrease of the reentrance region for the most long ranged potential $1/r^6$. The relative extent of the reentrance region is closest to the experimental data for the DLVO or the $1/r^{12}$-potentials.

The corresponding problem in three dimensions has been analyzed as well by MC simulations [49]. In case of a Lennard-Jones pair interaction the phase transition density reduces with increasing potential amplitude. A reentrance phenomena however was not found in this case. Additional studies are planned to quantify the effect of the interaction potential on the phase diagram.

Incommensurate Potentials

Incommensurate potentials have been applied to hard disk systems [50]. As in the commensurate case the particles gather within the potential minima. Depending on the wavelength and potential strength different phases occur. In Fig. 16 the phase diagram for a system of hard disks of diameter σ in an incommensurate potential, $V_{ext}(x, y) = V_0 \sin(2\pi x/\lambda)$ with $\lambda = 0.65579\sigma$, is shown. The phase diagram shows a laser induced freezing transition, but no reentrant melting, in contrast to the commensurate case.

Quantum Effects

Besides these classical studies we explore the validity of our results on atomic length scales. In this context we were able to investigate the properties of quantum hard disks with a finite particle mass m and interaction diameter σ in an external periodic potential by path integral Monte Carlo simulations (PIMC) [51, 52].

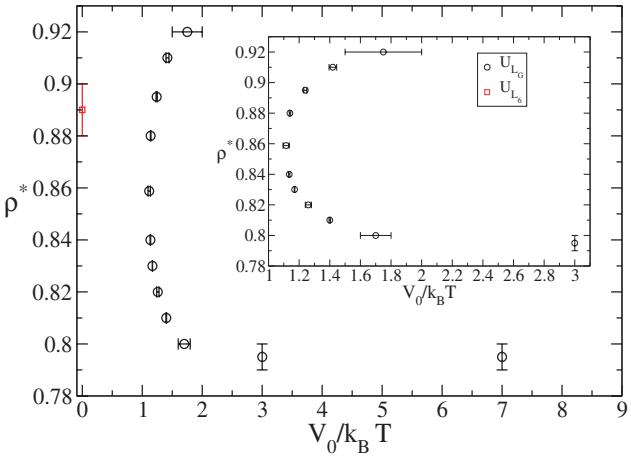


Fig. 16. Phase diagram of a hard disk system with $N = 1024$ particles under the influence on an incommensurate potential with $\lambda/\sigma = 0.65579$

Canonical averages $\langle A \rangle$ of an observable A in a N-particle system with Hamiltonian $\mathcal{H} = E_{kin} + V_{pot}$ in the volume V are given by:

$$\langle A \rangle = Z^{-1} \quad Sp \quad [\mathrm{A} \exp(-\beta \mathcal{H})] \quad . \tag{24}$$

Here $Z = Sp[\exp(-\beta\mathcal{H})]$ is the partition function and $\beta = 1/k_B T$ the inverse temperature. After the application of the Trotter-product-formula,

$$\exp(\beta\mathcal{H}) = \lim_{P\to\infty} (\exp(-\beta E_{kin}/P)\exp(-\beta V_{pot}/P))^P \ , \tag{25}$$

we obtain the path integral formulation of the partition function:

$$Z(N,V,T) = \lim_{P\to\infty} \left(\frac{mP}{2\pi\beta\hbar^2}\right)^{3NP/2} \prod_{S=1}^{P} \int d\{\mathbf{r}^{(S)}\} \exp[-\frac{\beta}{P}\sum_{J=1}^{N} \frac{mP^2}{2\hbar^2\beta^2}(\mathbf{r}_J^{(S)} - \mathbf{r}_J^{(S+1)})^2 + V_{pot}(\{\mathbf{r}^{(S)}\})] \tag{26}$$

Here m is the particle mass, P the Trotter number and $\mathbf{r}_J^{(S)}$ the coordinate of particle J at Trotter index S, and periodical boundary conditions apply, $P + 1 = 1$. This form of the partition function makes it possible to perform Monte Carlo simulations with increasing values of P in order to approximate the quantum limit ($P \to \infty$). Thermal averages in the ensemble with constant pressure p are given by the corresponding partition function $\Delta(N,p,T) = \int_0^\infty dV \exp[-\beta p V] Z(N,V,T)$.

In (26) we see, that in the path integral formalism each quantum particle J (for finite P-values) can be represented by a closed quantum chain of length

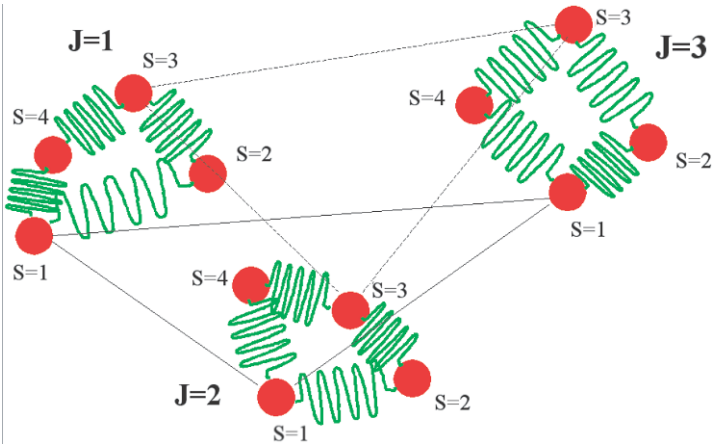


Fig. 17. Schematic picture of a system with $N = 3$ particles and Trotter number $P = 4$. All harmonic chains for each quantum particle are shown, and for the Trotter index $S = 1$ and $S = 3$ the interacting "particles"

P in coordinate space, where the classical coordinate of the point $\mathbf{r}_J^{(S)}$ on this chain at Trotter index S has a harmonic interaction to the nearest neighbors on the chain at $\mathbf{r}_J^{(S+1)}$ and $\mathbf{r}_J^{(S-1)}$. An interaction between different quantum particles only happens between particles with coordinates $\{\mathbf{r}^{(S)}\}$ at the same Trotter index S. A schematic picture of a system with $N = 3$ particles and $P = 4$ is shown in Fig. 17.

Due to the quantum delocalization effect a larger effective particle diameter results, and in the external potential this delocalization is asymmetrical: in the direction perpendicular to the potential valleys we obtain a stronger particle localization than parallel to the valleys, see Fig. 18. As a result the reentrance region in the phase diagram is significantly modified in comparison to the classical case, see Fig. 19.

Due to the larger quantum "diameter" the transition densities at small potential amplitudes are reduced in comparison to the classical values. At large amplitudes the classical and quantum transition densities merge. This effect is due to the approach of the effective quantum disk size to the classical value in the direction perpendicular to the potential valleys and leads to the surprising prediction, that the quantum crystal in a certain density region has a direct transition to the phase of the modulated liquid by an increase of the potential amplitude. This scenario is not known in the classical case. We plan to explore this interesting topic for systems with different particle masses by PIMC studies and finite-size-scaling methods.

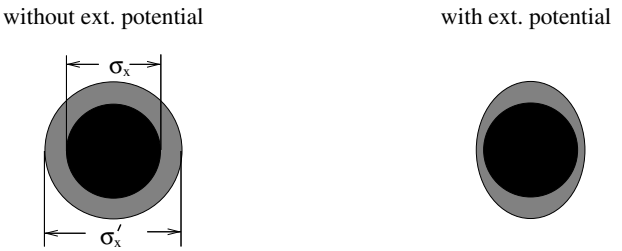


Fig. 18. Schematic picture of the effect of an external periodic potential of the form $V(x, y) = V_0 \sin(x/a)$ on the "effective" diameter of quantum hard disks

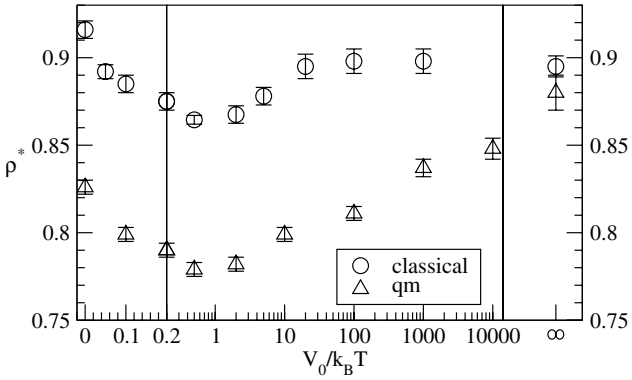


Fig. 19. Phase diagram in the density ($\rho^* = \rho\sigma^2$)-potential amplitude (V_0/k_BT)-plane for a system with $N = 400$ particles, $m^* = mT\sigma^2 = 10.000$ ("qm") and $m^* = \infty$ (classical) and Trotter order $P = 64$

Acknowledgements

We gratefully acknowledge useful discussions with C. Bechinger, K. Binder, D. Chaudhuri, M. Dreher, K. Franzrahe, P. Henseler, Chr. Kircher, M. Lohrer, W. Quester, M. Rao, A. Ricci, support by the SFB 513 and the SFB-TR6 and granting of computer time from the NIC, the HLRS and the SSC. PN thanks the institutes for theoretical physics of the FU Berlin and the U Mainz for the hospitality during his sabbatical stay.

References

1. D. P. Landau and K. Binder (2000) A Guide to Monte Carlo Simulations in Statistical Physics. Cambridge University Press
2. *Bridging Time Scales: Molecular Simulations for the Next Decade.* edited by P. Nielaba, M. Mareschal, G. Ciccotti, Springer, Berlin (2002)
3. S. Sengupta, P. Nielaba, M. Rao, and K. Binder (2000) Elastic constants from microscopic strain fluctuations. *Phys. Rev.* **E61**, pp. 1072–1080

4. P. M. Chaikin and T. C. Lubensky (1995) *Principles of condensed matter physics*. Cambridge University Press, Cambridge
5. D. C. Wallace (1958) in *Solid state physics*, eds. H. Ehrenreich, F. Seitz and D. Turnbull, Academic Press, New York; J. H. Weiner (1983), *Statistical mechanics of elasticity*. Wiley, New York
6. K. W. Wojciechowski and A. C. Brańka (1988) Elastic moduli of a perfect hard disc crystal in two dimensions. *Phys. Lett.* **134A**, pp. 314–318
7. S. Sengupta, P. Nielaba, and K. Binder (2000) Elastic moduli, dislocation core energy, and melting of hard disks in two dimensions. *Phys. Rev.* **E61**, pp. 6294–6301
8. F. Seitz (1940) *Modern Theory of Solids*, McGaraw-Hill, New York; M. Born and K. Huang (1954) *Dynamical Theory of Crystal Lattices*, Oxford University Press, Oxford; For experimental temperature dependent elastic constants of solid argon see: H. Meixner, P. Leiderer, P. Berberich and E. Lüscher (1972) The elastic constants of solid argon determined by stimulated brillouin scattering. *Phys. Lett.* **40A**, pp. 257–258
9. V. N. Ryzhov and E. E. Tareyeva (1995) Two-stage melting in two dimensions: first principles approach. *Phys. Rev.* **B51**, pp. 8789–8794
10. K. Franzrahe, Ph.D. thesis (in work)
11. P. Henseler, Ph.D. thesis (in work)
12. P. Nielaba, K. Binder, D. Chaudhuri, K. Franzrahe, P. Henseler, M. Lohrer, A. Ricci, S. Sengupta, and W. Strepp (2004) Elastic properties, structures and phase transitions in model colloids. *J. Phys.: Cond. Mat.* **16**, pp. S4115–S4136
13. K. Franzrahe, P. Henseler, A. Ricci, W. Strepp, S. Sengupta, M. Dreher, Chr. Kircher, M. Lohrer, W. Quester, K. Binder, and P. Nielaba (2005) Two-dimensional model colloids and nano-wires: phase transitions, effects of external potentials and quantum effects. *Comp. Phys. Commun.* **169**, pp. 197–202
14. A. Ricci (2006) Ph.D. thesis (U. Mainz)
15. K. Zahn, A. Wille, G. Maret, S. Sengupta, and P. Nielaba (2003) Elastic properties of 2D colloidal crystals from video microscopy. *Phys. Rev. Lett.* **90**, pp. 155506-1–155506-4
16. B. J. Alder and T. E. Wainwright (1962) Phase transition in elastic disks. *Phys. Rev.* **127**, pp. 359–361
17. J. Lee and K. Strandburg (1992) First-order melting transition of the hard-disk system. *Phys. Rev.* **B46**, pp. 11190–11193
18. J. A. Zollweg and G. V. Chester (1992) Melting in two dimensions. *Phys. Rev.* **B46**, pp. 11186–11189
19. T. V. Ramakrishnan (1982) Density-wave theory of first-order freezing in two dimensions. *Phys. Rev. Lett.* **48**, pp. 541–545; X. C. Zeng and D. W. Oxtoby (1990) Applications of modified weighted density functional theory: freezing of simple liquids. *J. Chem. Phys.* **93**, pp. 2692–2700; Y. Rosenfeld (1990) Free-energy model for the inhomogeneous hard-sphere fluid in D dimensions: structure factors for the hard-disk (D=2) mixtures in simple explicit form. *Phys. Rev.* **A42**, pp. 5978–5989
20. A. Jaster (1999) Computer simulations of the two-dimensional melting transition using hard disks. *Phys. Rev. E* **59**, pp. 2594–2602; A. Jaster (2000) Short-time behaviour of the two-dimensional hard-disk model. *Physica A* **277**, pp. 106–114

21. J. M. Kosterlitz and D. J. Thouless (1973) Ordering, metastability and phase transitions in two-dimensional systems. *J. Phys.* **C 6**, pp. 1181–1203; B. I. Halperin and D. R. Nelson (1978) Theory of two-dimensional melting, *Phys. Rev. Lett.* **41**, pp. 121–124; D. R. Nelson and B. I. Halperin (1979) Dislocation-mediated melting in two dimensions. *Phys. Rev. B* **19**, pp. 2457–2484; A. P. Young (1979) Melting and the vector Coulomb gas in two dimensions. *Phys. Rev. B* **19**, pp. 1855–1866
22. K. Zahn, R. Lenke, and G. Maret (1999) Two-stage melting of paramagnetic colloidal crystals in two dimensions. *Phys. Rev. Lett.* **82**, pp. 2721–2724
23. M. Bates and D. Frenkel (2000) Influence of vacancies on the melting transition of hard disks in two dimensions. *Phys. Rev.* **E61**, pp. 5223–5227
24. K. Binder, S. Sengupta, and P. Nielaba (2002) The liquid-solid transition of hard-discs: first order transition or Kosterlitz-Thouless-Halperin-Nelson-Young scenario? *J. Phys.: Condens. Matter* **14**, pp. 2323–2333
25. N. A. Clark, B. J. Ackerson, and A. J. Hurd (1983) Multidetector scattering as a probe of local structure in disordered phases. *PRL* **50**, pp. 1459–1462
26. A. Chowdhury, B. J. Ackerson, and N. A. Clark (1985) Laser-induced freezing. *PRL* **55**, pp. 833–836
27. K. Loudiyi and B. J. Ackerson (1992) Direct observation of laser induced freezing. Physica **A 184**, 1–25; Monte Carlo simulation of laser induced freezing. *ibid* pp. 26–41
28. Q.-H. Wei, C. Bechinger, D. Rudhardt, and P. Leiderer (1998) Experimental study of laser-induced melting in two-dimensional colloids. *PRL* **81**, pp. 2606–2609
29. C. Bechinger, Q. H. Wei, and P. Leiderer (2000) Reentrant melting of two-dimensional colloidal systems. *J. Phys.: Cond. Mat.* **12**, pp. A425–A430
30. C. Bechinger, M. Brunner, and P. Leiderer (2001) Phase behavior of two-dimensional colloidal systems in the presence of periodic light fields. *PRL* **86**, pp. 930–933
31. J. Chakrabarti, H. R. Krishnamurthy, and A. K. Sood (1994) Density functional theory of laser-induced freezing in colloidal suspensions. *PRL* **73**, pp. 2923–2926
32. L. L. Rasmussen and D. W. Oxtoby (2002) Induced freezing and re-entrant melting in the hard-disc fluid; applications of the fundamental measure functional. *J. Phys.: Cond. Mat.* **14**, pp. 12021–12030
33. E. Frey, D. R. Nelson, and L. Radzihovsky (1999) Light-induced melting of colloidal crystals in two dimensions. *PRL* **83**, pp. 2977–2980; L. Radzihovsky, E. Frey, D. R. Nelson (2001) Novel phases and reentrant melting of two-dimensional colloidal crystals. *Phys. Rev.* **E63**, pp. 031503-1–031503-34
34. J. Chakrabarti, H. R. Krishnamurthy, A. K. Sood, and S. Sengupta (1995) Reentrant melting in laser field modulated colloidal suspensions. *PRL* **75**, pp. 2232–2235
35. C. Das and H. R. Krishnamurthy (1998) Laser-induced quasicrystalline order in charge-stabilized colloidal systems. *PRB* **58**, pp. R5889–R5892
36. C. Das, A. K. Sood, and H. R. Krishnamurthy (1999) Bond-orientational ordering and shear rigidity in modulated colloidal liquids. *Physica* **A 270**, pp. 237–244
37. C. Das, P. Chaudhuri, A. Sood, and H. Krishnamurthy (2001) *Current Science* **80**(8), p. 959
38. D. Chaudhuri and S. Sengupta (2004) A numerical renormalization group study of laser-induced freezing. *Europhys. Lett.* **67**, pp. 814–819; (2006) Direct test of

defect-mediated laser-induced melting theory for two-dimensional solids. *Phys. Rev.* **E73**, pp. 011507-1–011507-12
39. For an introduction to phase transitions in colloids see, A. K. Sood (1991) in *Solid State Physics*, E. Ehrenfest and D. Turnbull Eds., Academic Press, New York; **45**, 1; P. N. Pusey in Liquids (1991) Freezing and the Glass Transition, J. P. Hansen and J. Zinn-Justin Eds. (North Holland, Amsterdam)
40. W. Strepp, S. Sengupta, and P. Nielaba (2001) Phase transitions of hard disks in external periodic potentials: a Monte Carlo study. *Phys. Rev.* **E63**, pp. 046106-1–046106-10
41. W. Strepp, S. Sengupta, and P. Nielaba (2002) Phase transitions of soft disks in external periodic potentials: a Monte Carlo study. *Phys. Rev.* **E66**, pp. 056109-1 – 056109-13
42. W. Strepp, S. Sengupta, M. Lohrer, and P. Nielaba (2002) Phase transitions of hard and soft disks in external periodic potentials: A Monte Carlo study. *Comput. Phys. Commun.* **147**, pp. 370–373
43. W. Strepp, S. Sengupta, M. Lohrer, and P. Nielaba (2003) Phase transitions in model colloids in reduced geometry. *Mathematics and Computers in Simulation* **62**, pp. 519–527
44. K. Binder (1981) Finite size scaling analysis of Ising model block distribution functions. *Z. Phys.* **B43**, pp. 119–140; K. Binder (1981) Critical properties from Monte Carlo coarse graining and renormalization. *PRL* **47**, pp. 693–696
45. K. Vollmayr, J. D. Reger, M. Scheucher, and K. Binder (1993) Finite size effects at thermally-driven first order transitions: a phenomenological theory of the order parameter distribution. *Z. Phys.* **B 91**, pp. 113–125
46. N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller and E. Teller (1953) Equation of state calculations by fast computing machines. *J. Chem. Phys.* **21**, pp. 1087–1092
47. W. Strepp. M. Lohrer, S. Sengupta, and P. Nielaba, Phase transitions of hard and soft disks in external periodic potentials: a Monte Carlo study of the effect of the interaction range, preprint
48. M. Brunner, C. Bechinger, W. Strepp, V. Lobaskin, and H. H. von Gruenberg (2002) Density-dependent pair interaction in 2D colloidal suspensions. *Europhys. Lett.* **58**, pp. 926–932
49. W. Quester (2003) Diplomarbeit (U. Konstanz)
50. Chr. Kircher (2004) Diplomarbeit (U. Konstanz)
51. W. Strepp (2003) Ph.D-thesis (U. Konstanz)
52. W. Strepp and P. Nielaba, in preparation

Computer Simulation of Liquid Crystals

M.P. Allen

Department of Physics and Centre for Scientific Computing University of Warwick, Coventry CV4 7AL, United Kingdom
m.p.allen@warwick.ac.uk



Michael P. Allen

M.P. Allen: *Computer Simulation of Liquid Crystals*, Lect. Notes Phys. **704**, 191–210 (2006)
DOI 10.1007/3-540-35284-8_9 © Springer-Verlag Berlin Heidelberg 2006

1 Introduction . . . 193

1.1 Simulation Models . . . 194
1.2 Order in Nematic Phases . . . 194
1.3 Oseen-Frank Elastic Description . . . 195
1.4 Onsager Theory . . . 195

2 Elastic Properties . . . 196

2.1 Bulk Elastic Constants . . . 196
2.2 Helical Twisting Power . . . 197

3 Surface Properties . . . 199

3.1 Anchoring and Interfaces . . . 199
3.2 Solid Surface Anchoring . . . 199
3.3 Fluctuations in Confined Geometry . . . 200
3.4 Nematic-Isotropic Interface . . . 201

4 Macroparticles in Liquid Crystals . . . 203

4.1 Colloidal Suspensions . . . 203
4.2 Single Spherical Macroparticle . . . 204
4.3 Two Spherical Macroparticles . . . 206

5 Conclusions . . . 207

References . . . 208

This chapter describes some of the simulation methods that are used to investigate the properties of liquid crystals. Orientational elasticity, one of the characteristic features of liquid crystals, may be studied through equilibrium wave-vector-dependent fluctuations, or by directly deforming the director field via imposed boundary conditions. This leads to investigations of surface anchoring coefficients and the helical twisting power of chiral dopants. Interfacial properties such as surface tension may be studied using standard methods, but the orientational ordering of liquid crystals leads to a richer behaviour than is seen for simple fluids. Finally, simulations of macroparticles suspended in liquid crystalline solvents may provide information about the defect structure on the molecular scale, and help to make contact with larger-scale modelling methods and with experiment.

1 Introduction

Liquid crystals are characterized by long-range orientational order of the molecules (a property more usually associated with the solid phase) while the positional order, in some directions at least, remains short-ranged and the system is fluid (properties inherited from the liquid state). There are many liquid crystalline phases; two of the simplest, the nematic and smectic-A phases, are illustrated in Fig. 1. In the nematic phase, there is a preferred direction of ordering called the *director*. This is usually represented as a unit vector $\boldsymbol{n}$; however it should be noted that there is macroscopic symmetry $\boldsymbol{n} \rightarrow -\boldsymbol{n}$ even if the individual molecules are not symmetric. For this reason, the vector $\boldsymbol{n}$ is sometimes depicted with two arrow-heads (or with none). Apart from this orientational ordering, the nematic phase is translationally disordered, i.e. although there are typically short-ranged correlations between the positions of nearby molecules, these correlations decay exponentially with separation. In the smectic-A phase, the molecules form layers with the director aligned along the layer-normal; this means that there is long-ranged positional ordering in

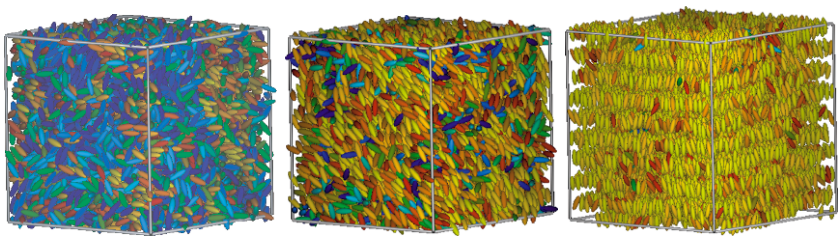


Fig. 1. Snapshots of configurations from molecular simulations of (left to right) isotropic, nematic and smectic phases. The molecules are represented by rigid ellipsoidal shapes, and are colour-coded to indicate the orientation. Cubic periodic boundary conditions apply

that direction, but the layers themselves are liquid-like and possess only short-ranged in-plane positional ordering. There are many other more complicated liquid crystal phases, including a whole range of smectics. This chapter will concentrate on the nematic phase.

1.1 Simulation Models

A characteristic feature of the molecules that form liquid crystals, so-called *mesogens*, is that they are highly non-spherical: most commonly rod-like. Many liquid crystal simulations have used very simple rigid bodies to describe these molecules, such as those illustrated in Fig. 2. Hard-particle models such as ellipsoids of revolution of length A, width B, and spherocylinders of overall length $L + D$, width D, have been studied in great detail, and (for suitable elongations) have been shown to exhibit at least the basic nematic phase. Such simple models are, of course, at best a caricature of real molecules, although they come closer to describing some colloidal suspensions of non-spherical particles which also show liquid crystalline behaviour. The effects of attractive forces have been investigated through the so-called Gay-Berne family of models [1–5]. This chapter will concentrate on models of the above type, which are sufficiently detailed to exhibit the desired phenomena while being economical on the computer. However considerable progress has been made in refining these models. Non-uniaxial shapes [6–8], flexibility [9, 10], dipolar forces [11, 12], and the effects of hydrogen bonds [13] have all been studied. Computer power has now advanced to the stage where realistic, fully atomistic, potentials may be directly useful [14, 15].

1.2 Order in Nematic Phases

Orientational ordering in nematics composed of uniaxial molecules is described by a second-rank symmetric, traceless, tensor Q

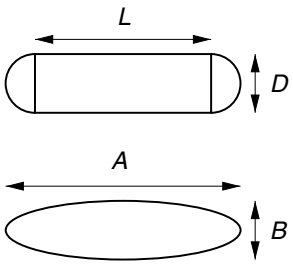


Fig. 2. Two typical hard-particle shapes used in Monte Carlo simulations of liquid crystals. The spherocylinder is characterized by the length L of the cylindrical section and the diameter D. The axially symmetric ellipsoid is characterized by the length A of the major axis and the two transverse minor axis diameters B

$$Q_{\alpha\beta} = \sum_{j=1}^{N} \left(\tfrac{3}{2} u_{j\alpha} u_{j\beta} - \tfrac{1}{2} \delta_{\alpha\beta} \right) , \quad \alpha, \beta = x, y, z$$

where the sum is over all N molecules, and the unit vector $\boldsymbol{u}_j$ points along the axis of molecule j; $\delta_{\alpha\beta}$ is the Kronecker delta. This may be diagonalized to give three eigenvalues which are conveniently written and ordered from highest to lowest: $S > -\frac{1}{2}(S - S_\mathrm{b}) \gtrsim -\frac{1}{2}(S + S_\mathrm{b})$. The largest eigenvalue S is generally called the *order parameter*, and in nematic phases it is *significantly* larger than the other two; it takes values between 0 (completely disordered) and 1 (completely ordered). The quantity S_b measures biaxial ordering, and is typically quite small: in fact, in simulations of simple nematics, it is nonzero only due to finite-size effects, and the lowest two eigenvalues are nearly degenerate. The eigenvector $\boldsymbol{n}$ corresponding to eigenvalue S is the director. In an isotropic phase, both S and S_b are small, becoming zero in the limit of infinite system size. As a consequence, the eigenvalues of Q are all close to zero, and the director effectively becomes undefined because of isotropic orientational averaging.

1.3 Oseen-Frank Elastic Description

Simulations may be compared with theoretical descriptions of the liquid crystal, and the ones of interest here are formulated as free-energy *functionals* of some position- and orientation-dependent quantity. The Oseen-Frank [16, 17] free energy $\mathcal{F}_\mathrm{OF}$ is a functional of the position-dependent liquid crystal director $\boldsymbol{n}(\boldsymbol{r})$:

$$\mathcal{F}_\mathrm{OF}\big[\boldsymbol{n}(\boldsymbol{r})\big] = \int_V \mathrm{d}^3\boldsymbol{r}\ f_\mathrm{b}(\boldsymbol{r}) + \int_{\partial V} \mathrm{d}^2\boldsymbol{s}\ f_\mathrm{s}(\boldsymbol{s}) \tag{1a}$$

$$f_\mathrm{b}(\boldsymbol{r}) = \tfrac{1}{2} K_{11} \big(\nabla \cdot \boldsymbol{n}\big)^2 + \tfrac{1}{2} K_{22} \big(\boldsymbol{n} \cdot \nabla \times \boldsymbol{n}\big)^2 + \tfrac{1}{2} K_{33} \big(\boldsymbol{n} \times \nabla \times \boldsymbol{n}\big)^2 \tag{1b}$$

$$f_\mathrm{s}(\boldsymbol{s}) = \tfrac{1}{2} W \sin^2\theta \quad \text{where} \cos\theta = \boldsymbol{n} \cdot \boldsymbol{n}_s \, . \tag{1c}$$

The director is assumed to vary smoothly with position, leading to a squared-gradient form of the bulk free energy density f_b, which is integrated over the sample volume V. The coefficients appearing here are the splay (K_{11}), twist (K_{22}), and bend (K_{33}) elastic constants. Also included in (1) is a simple form of the surface free energy density f_s, which expresses the orienting effects of the bounding surface ∂V. The coefficient here is the surface anchoring strength W, and the form chosen here depends on the angle θ between $\boldsymbol{n}$ and the orientation $\boldsymbol{n}_s$ favoured by the surface. These various elastic and anchoring constants may be calculated in simulation and compared with experiment or theory.

1.4 Onsager Theory

The second density-functional formulation that will be referred to later is the Onsager [18] free energy $\mathcal{F}_\mathrm{Ons}$. This depends on the single-particle density

$\varrho(\boldsymbol{r},\boldsymbol{u})$, a function of both position *and* orientation:

$$\mathcal{F}_{\rm Ons}[\varrho(\boldsymbol{r},\boldsymbol{u})] = \mathcal{F}^{(1)}_{\rm Ons}[\varrho(\boldsymbol{r},\boldsymbol{u})] + \mathcal{F}^{(2)}_{\rm Ons}[\varrho(\boldsymbol{r},\boldsymbol{u})] \tag{2a}$$

$$\mathcal{F}^{(1)}_{\rm Ons}[\varrho(\boldsymbol{r},\boldsymbol{u})] = k_{\rm B}T \int {\rm d}^3\boldsymbol{r}\, {\rm d}^2\boldsymbol{u}\ \varrho(\boldsymbol{r},\boldsymbol{u}) \big[\ln \varrho(\boldsymbol{r},\boldsymbol{u})\Lambda^3 - 1 + W(\boldsymbol{r},\boldsymbol{u})/k_{\rm B}T\big] \tag{2b}$$

$$\mathcal{F}^{(2)}_{\rm Ons}[\varrho(\boldsymbol{r},\boldsymbol{u})] = -\tfrac{1}{2}k_{\rm B}T \iint {\rm d}^3\boldsymbol{r}\, {\rm d}^2\boldsymbol{u}\ {\rm d}^3\boldsymbol{r}'\, {\rm d}^2\boldsymbol{u}'\ \varrho(\boldsymbol{r},\boldsymbol{u})\varrho(\boldsymbol{r}',\boldsymbol{u}')\ f(\boldsymbol{r}-\boldsymbol{r}',\boldsymbol{u},\boldsymbol{u}') \tag{2c}$$

Here $W(\boldsymbol{r},\boldsymbol{u})$ represents an external or boundary field, and the double integral involves the Mayer function

$$f(\boldsymbol{r}-\boldsymbol{r}',\boldsymbol{u},\boldsymbol{u}') = {\rm e}^{-v(\boldsymbol{r}-\boldsymbol{r}',\boldsymbol{u},\boldsymbol{u}')/k_{\rm B}T} - 1$$

which is related directly to the intermolecular pair potential $v(\boldsymbol{r}-\boldsymbol{r}',\boldsymbol{u},\boldsymbol{u}')$. This is a forerunner of modern density functional theories: it is expected to be accurate at low densities, as it is effectively a second-virial approximation. Note that, as written, there are no adjustable parameters.

2 Elastic Properties

2.1 Bulk Elastic Constants

Long-wavelength deformations $\delta\boldsymbol{n}(\boldsymbol{r})$ from uniform alignment incur a free-energy penalty through $\mathcal{F}_{\rm OF}$, and one approach to the measurement of the elastic constants K_1, K_2, K_3 in a simulation is to observe the thermally induced equilibrium fluctuations of orientation. It is easiest to compute the Fourier coefficients $\delta\hat{\boldsymbol{n}}(\boldsymbol{k})$ at low wavenumber $|\boldsymbol{k}|$. Choosing coordinates so that the undistorted director is $\boldsymbol{n}(\boldsymbol{r}) = (0,0,1)$, and $\boldsymbol{k} = (k_x, 0, k_z)$, it is easy to show that $\delta\hat{n}_x \propto \delta\hat{Q}_{xz}$, $\delta\hat{n}_y \propto \delta\hat{Q}_{yz}$, where

$$\hat{Q}_{\alpha\beta}(\boldsymbol{k}) = \sum_j \big(\tfrac{3}{2}u_{j\alpha}\,u_{j\beta} - \tfrac{1}{2}\delta_{\alpha\beta}\big)\,{\rm e}^{{\rm i}\boldsymbol{k}\cdot\boldsymbol{r}_j}\ , \quad \alpha,\beta = x,y,z\ .$$

Here $\boldsymbol{r}_j$ is the position vector of molecule j, and $\boldsymbol{u}_j$, as before, is its orientation. An argument based on equipartition of energy leads to the following expressions for the *inverse* of the mean-square fluctuations $\Phi_{\alpha\beta}(\boldsymbol{k}) \equiv \langle|\hat{Q}_{\alpha\beta}(\boldsymbol{k})|^2\rangle$:

$$\Phi^{-1}_{xz} \propto K_1 k_x^2 + K_3 k_z^2\ , \qquad \Phi^{-1}_{yz} \propto K_2 k_x^2 + K_3 k_z^2\ . \tag{3}$$

Typical results for simulations of a Gay-Berne model [19] (essentially soft, attractive, ellipsoids with $A/B = 3$) are shown in Fig. 3. The essential point here is that it is possible to reach sufficiently low wavenumbers to observe the asymptotic behaviour predicted by (3), and hence extract the elastic constants. In this case, quite modest system sizes ($N = 8000$) were sufficient, but runs of sufficient length to equilibrate all the long-wavelength modes were essential.

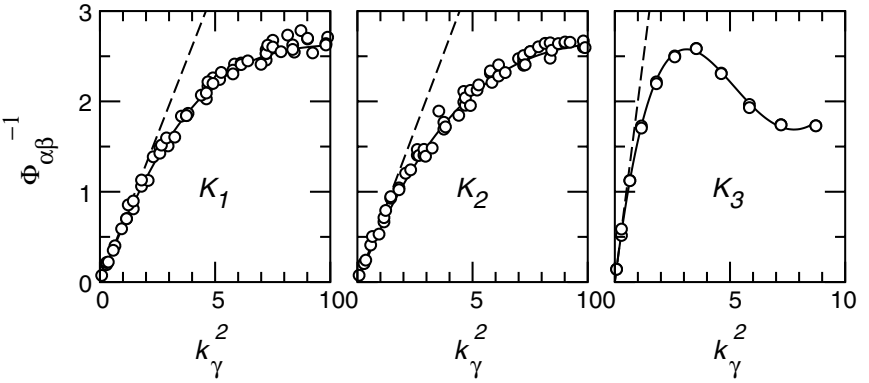


Fig. 3. Inverse orientational order fluctuations as functions of k_γ^2. Components α, β, γ are chosen to select the appropriate elastic constant according to (3). *Dashed lines* indicate the low-k fit; *solid lines* are sections through a global polynomial fit to the entire dataset; error bars are approximately indicated by the symbol size. After [19]

2.2 Helical Twisting Power

The director may be twisted by external forces or dissolved *chiral* molecules, as illustrated schematically in Fig. 4. At low concentrations of chiral dopants, the effect is linear in concentration. This effect is important in display devices. However it is very subtle: the twisted structures have dimensions hundreds of times larger than the molecules. How can the *twisting power* be measured? One method, exemplified in Fig. 5, is to measure the free energy difference between right- and left-handed molecules in a twisted liquid crystal [20]. In this early work, a trick was used to improve the statistics: the chiral mole-

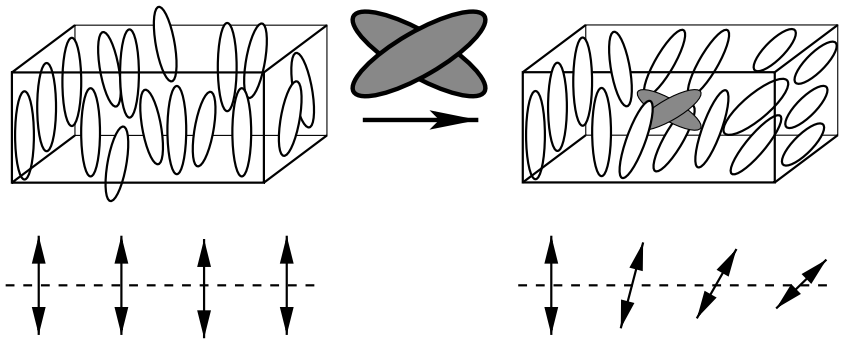


Fig. 4. Schematic diagram of the effect of a chiral molecule on the director in a nematic liquid crystal. An untwisted configuration develops a twist with a well-defined pitch: one quarter of a turn is shown here. The inverse pitch of the twist is proportional to the concentration of chiral dopants

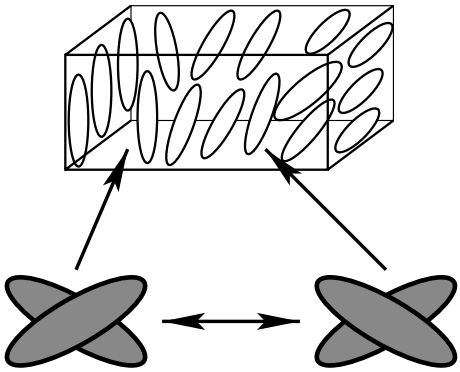


Fig. 5. Schematic diagram illustrating possible ways of measuring the helical twisting power. The chemical potential of left- and right-handed chiral dopants may be estimated by the Widom test-particle insertion method. More efficient methods are: (**a**) to compute the difference in chemical potentials by directly transforming one species into the other; (**b**) to measure the torsional stresses

cule was taken to be a dimer of the nematic solvent arranged in a "scissors" conformation, characterized by an internal twist angle.

Typical results for hard ellipsoids are shown in Fig. 6 as a function of twist angle. The dimer is most effective at generating a helical structure when the two monomers have a relative angle of about 60° or 120°, although this is somewhat density dependent.

Without the dimer trick, which cannot be applied to most realistic models of liquid crystals, the evaluation of the necessary free-energy difference requires a lengthy thermodynamic integration [21]. Another approach is based on measuring the torsional stress on a chiral molecule in an *untwisted* liquid

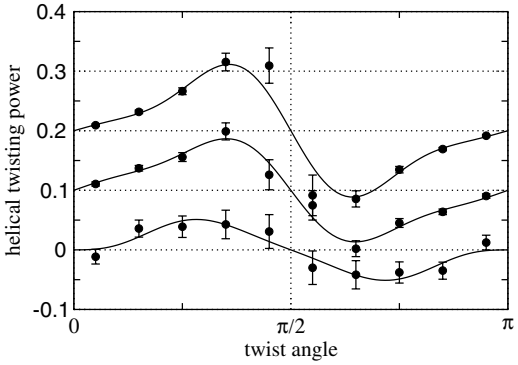


Fig. 6. Helical twisting power for hard ellipsoid dimers with $A/B = 5$ at reduced densities $\rho AB^2 = 0.50$ (*bottom*), 0.55 (*middle*), and 0.60 (*top*); successive curves are shifted vertically for clarity. After [20]

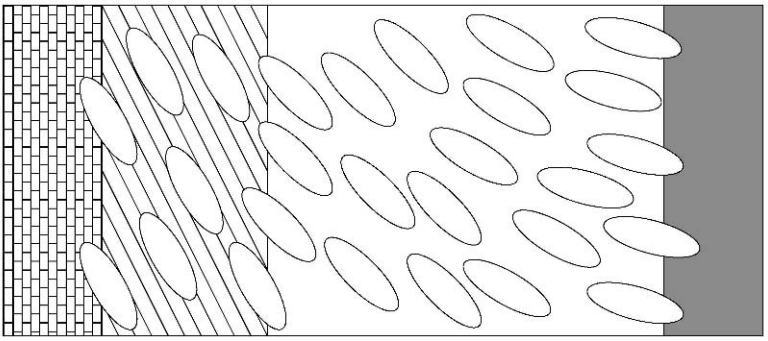


Fig. 7. Schematic diagram of the simulation geometry used to investigate anchoring at an interface. At the left (*brick pattern*) is a wall, which excludes the centres of the molecules. Next to this (*slanted lines*) is a strong field which orients the liquid crystal molecules. In the middle of the cell is a wide region of liquid crystal, with no applied field. At the right (*grey*) is the interface of interest, which may be another solid wall, or a fluid-fluid interface such as that between the nematic and isotropic phases

crystal [22]. This approach has been applied to realistic systems with some success [23].

3 Surface Properties

3.1 Anchoring and Interfaces

The simplest approach to the investigation of surface anchoring is to measure the bulk elastic constants as described in the previous section, and then simulate a system in which the director field is deformed by the boundary conditions, fitting the resulting profiles to those predicted by variationally minimizing (1).

A simple slab geometry for this is sketched in Fig. 7. This has been used to study anchoring at: (a) the nematic-solid interface; and (b) the equilibrium nematic-isotropic interface. In both cases, reasonable sized systems (of order tens of thousands of molecules) are needed to allow sufficient separation between interfaces, as well as reasonable transverse dimensions; and as always in inhomogeneous systems, equilibration needs to be monitored carefully.

3.2 Solid Surface Anchoring

As an example, consider the case where the liquid crystal is *homeotropically anchored* at the interface of interest, i.e. the preferred orientation is normal to the surface. In this case, if the orientation is forced to adopt a different orientation at the opposite wall, and assuming that the deformation is small,

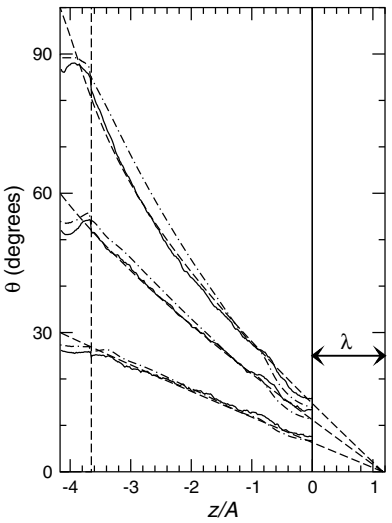


Fig. 8. Anchoring in a slab bounded by hard walls: director angle θ as a function of position, for hard ellipsoids with $A/B = 15$. The interface of interest is at $z = 0$. The range of the orienting field at the other interface is indicated by the *vertical dashed line*; three different orientation directions are shown. Simulation results: *solid lines*. Onsager theory predictions: *dot-dash lines*. Elastic theory fits: *dashed lines*. The extrapolation of the elastic theory defining the length λ is shown. After [25]

the director profile is largely determined by competition between the bend elastic constant K_{33} and the surface anchoring strength W. An analytical treatment [24] of (1) shows that the key physical quantity is the ratio of these two, the so-called *extrapolation length* $\lambda = K_{33}/W$. Typical results [25] for hard ellipsoids with $A/B = 15$ are shown in Fig. 8. They are compared with the predictions of Onsager theory, i.e. the profiles obtained by numerically minimizing the free energy of (2) with respect to variations of the single-particle density. The essential conclusions are: (a) elastic theory fits quite well, except near the walls; (b) the extrapolation length λ for this system is of order one molecular length; and (c) Onsager theory, with essentially no adjustable parameters, reproduces these director profiles very well, *even in the wall regions where the elastic theory is inaccurate.*

3.3 Fluctuations in Confined Geometry

The surface anchoring coefficient also plays a role in determining the spectrum of director fluctuations in slab geometry, and this can help resolve an interesting question: where are the elastic boundaries relative to the physical walls of the system? In slab geometry, with the walls normal to the z-direction, and homeotropic anchoring, it turns out that the elastic theory of (1) leads to a Fourier-like-decomposition of director fluctuations, like that of Sect. 2.1, but

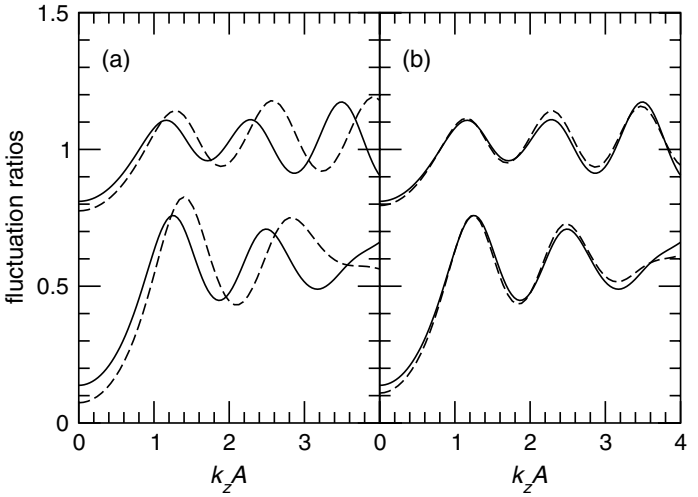


Fig. 9. Ratios of mean-squared fluctuations for box lengths $L_z/A = 3.29$ (shifted vertically down by 0.5 for clarity) and $L_z/A = 4.11$, all relative to $L_z/A = 4.93$. Simulation results: *solid lines*. Elastic theory predictions: *dashed lines* (**a**) assuming that $\Delta/A = 0$; (**b**) assuming $\Delta/A = 0.3$. After [26]

with unevenly-spaced wavenumbers k_z. Suppose for simplicity $k_x = k_y = 0$. If the elastic boundary planes are separated by L, the fluctuation spectrum $\Phi_{xz}(k_z; L)$ (see (3)) depends on both k_z and L; in fact it is parametrized by the dimensionless quantities $\chi = k_z L$ and $\xi = WL/K_{33} = L/\lambda$. Of course, from the simulation, L is unknown, but it may be assumed that $L = L_z + 2\Delta$ where L_z is the physical separation of the simulation box walls and Δ is a constant shift characteristic of each wall. Ratios of mean-square fluctuations for different box lengths $\Phi_{xz}(k_z; L_1)/\Phi_{xz}(k_z; L_2)$ turn out to be very sensitive to the value of Δ. This is shown in Fig. 9 for hard ellipsoids with $A/B = 15$, homeotropically anchored between two hard parallel confining walls separated by $L_z/A = 3.29, 4.11, 4.93$ [26]. As well as demonstrating that the relevant quantities can be estimated in this way, this study revealed that changes in the surface layer structure with density resulted in a non-monotonic variation of W.

3.4 Nematic-Isotropic Interface

When nematic and isotropic phases coexist, one may observe a thick film of nematic between a solid wall and the bulk isotropic phase. The interest here lies in studying the structure of the nematic-isotropic (N-I) interface, and any director distortion within the film, caused by competing anchoring conditions at the two interfaces. Another way of looking at the anchoring effect at the N-I interface is to regard it as an orientation-dependent surface

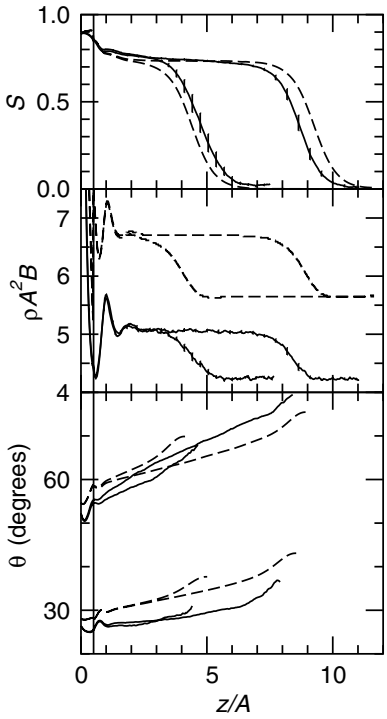


Fig. 10. Profiles of order parameter S, reduced density ρAB^2, and director angle θ measured from the wall normal, for hard ellipsoids with $A/B = 15$ in the vicinity of a hard wall with homeotropic anchoring, at the nematic-isotropic coexistence point. A nematic film is adsorbed at the wall, $z = 0$, and an orienting field is applied in the immediate vicinity, for $z/A < 0.5$. The isotropic phase appears at the right; the diagrams show just half of the simulation box. Solid lines with error bars represent simulation results for two different film thicknesses; only a subset of the data point are shown for clarity. Dashed lines represent the predictions of Onsager theory at two comparable film thicknesses. For the director angle, the results of two different orienting fields are shown in each case. After [27]

tension, and the first question to answer is: what is the preferred director orientation? Results for hard ellipsoids with $A/B = 15$ are shown in Fig. 10 [27]. The simulation results show that the Onsager theory, again obtained by numerically minimizing the free energy of (2), is reasonably accurate for this system. When the director at the wall is tilted relative to the normal, it tilts further as it approaches the N-I interface, confirming the long-held belief that the anchoring at this surface is in-plane ($\theta = 90°$), rather than homeotropic ($\theta = 0°$). Further evidence comes from a direct calculation of the surface tension γ [28] for the cases where the director is forced to lie perpendicular to, and in the plane of, the interface. This may be done through the microscopic expression

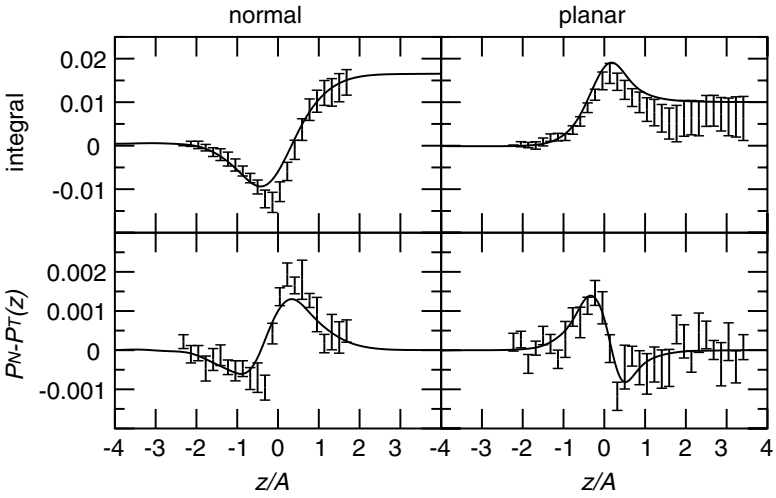


Fig. 11. Surface tension integrand $P_{\mathrm{T}}(z) - P_{\mathrm{N}}$ and running integral, for hard ellipsoids with $A/B = 15$ at the isotropic-nematic interface. Two director orientations are shown: normal to the interface and in the plane of the interface. Error bars are simulation results: only 25% of the points are shown for clarity. The solid line is the Onsager theory prediction. After [28]

$$\gamma = \int_{-\infty}^{\infty} \mathrm{d}z \, \left[P_{\mathrm{N}} - P_{\mathrm{T}}(z) \right]$$

where $P_{\mathrm{N}} = P_{zz} = P$ is the normal component, and $P_{\mathrm{T}}(z) = P_{xx}(z) = P_{yy}(z)$ is the transverse component, of the pressure tensor P. Far from the interface, $P_{\mathrm{T}}(z) = P_{\mathrm{N}} = P$. The integrand, and the running integral, in the above expression, are shown in Fig. 11 for the two director orientations mentioned. The integral, and hence the value of γ, is smaller for the in-plane orientation, as expected. Interestingly, the Onsager theory again gives an accurate prediction of the curves in both cases.

4 Macroparticles in Liquid Crystals

4.1 Colloidal Suspensions

Macroparticles (solid or liquid) introduced into liquid crystals distort the director distribution, through the surface anchoring condition. Considering a single macroparticle, for homeotropic boundaries, topological mismatch between the local director on the surface and a uniform director at large distances creates a hedgehog configuration around the particle with a complementary defect in the surrounding solvent. In the defect region, the nematic order parameter is low and the director becomes undefined. This gives rise to an effective

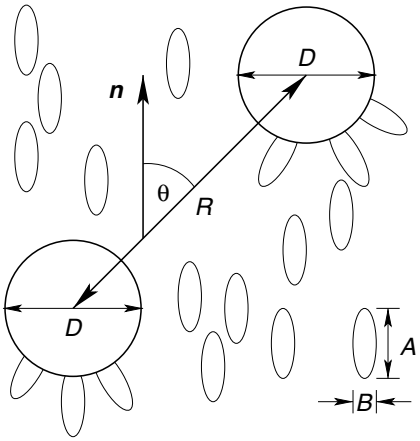


Fig. 12. Schematic diagram showing two macroparticles of diameter D, separated by distance R, in a nematic liquid solvent composed of molecules of length A and width B. The interparticle vector makes an angle θ with the director. In the studies reported here, the molecules in the solvent tend to be aligned homeotropically at the macroparticle surface

long-range interaction between macroparticles, simply because the free energy of the director field depends on these disortions, and hence on the macroparticle positions. The interaction is not, generally, a simple pairwise-additive one, although the effective force between an isolated pair of macroparticles can be studied [29]. Figure 12 shows schematically a two-particle system geometry for studying this by simulation. The aggregation of macroparticles through this mechanism creates, for example, e.g. threadlike structures [30]. For this case, the structure around a single spherical macroparticle typically adopts one of two forms, depending on the particle radius. For small spheres, a defect ring forms in the solvent around the "equator"; for large spheres this is replaced by a single point-like defect located near the north or south "pole". In the first case, the effective interaction between a pair of particles has a similar form to that between electrostatic quadrupoles, while in the latter case it is analogous to a dipolar interaction.

4.2 Single Spherical Macroparticle

These structures can be studied by molecular simulation [31]. To study, especially, the polar or "satellite" defect requires quite large system sizes, in order to cope with the long range of the structural effects. Results have been reported [32] for up to 10^6 Gay-Berne-like ellipsoids of length $A = 3$, width $B = 1$, and macroparticle diameters in the range $6 \leq D \leq 30$. For this system, the equatorial ring defect is seen for all diameters, and is certainly the stable state for $D \leq 20$. For $D = 30$ the polar satellite defect is at least

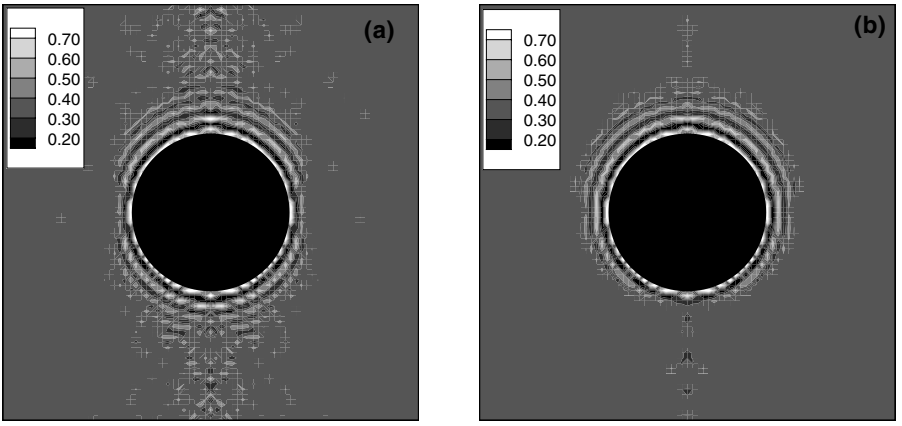


Fig. 13. Slices through the 3D density map ρ, represented on a grey scale, for macroparticle diameter $D = 30$: (**a**) for the equatorial ring defect; (**b**) for the polar satellite defect. The systems are symmetric to rotation about a vertical line through the centre of the macroparticle (*in black*). After [32]

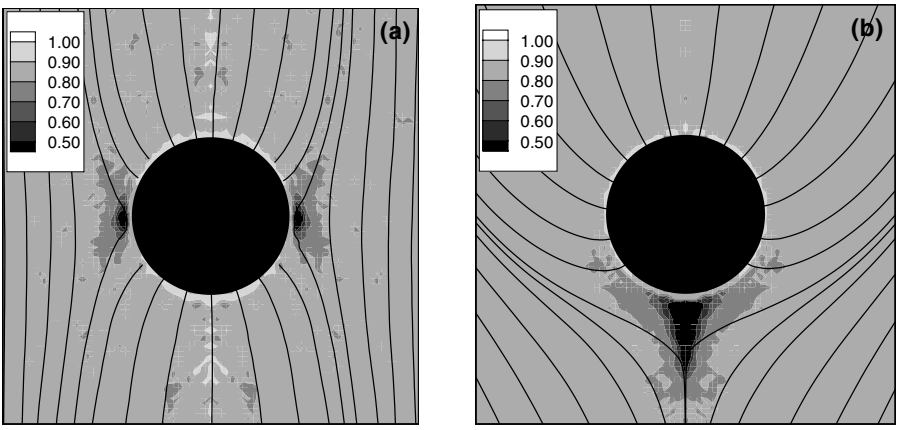


Fig. 14. Slices through the 3D order map S, represented on a grey scale, with superimposed director streamlines for $D = 30$: (**a**) for the equatorial ring defect; (**b**) for the polar satellite defect. The systems are symmetric to rotation about a vertical line through the centre of the macroparticle (*in black*). After [32]

metastable on the timescale of the simulations. The solvent density in both cases is shown in Fig. 13. The density oscillations typical of a liquid near a rigid wall are softened in the defect regions. Corresponding slices through the order parameter field, with streamlines indicating the spatial variation of the director, are shown in Fig. 14. The defect structures appear more clearly in these plots. Especially interesting is the size of the polar defect: it is clearly not point-like, but extends over distances comparable with the size of the

macroparticle itself. Closer study gives more information about the sizes and positions of these defects, and compares with elastic theory predictions [32].

4.3 Two Spherical Macroparticles

It is particularly interesting to study the distortions in the director field that occur when two macroparticles approach one another, and to measure the effective force between them. Consider again ellipsoidal molecules with length $A = 3$, width $B = 1$, and two spherical macroparticles of diameter $D = 6$. Results for this system have been reported for a range of separations $9 \leq R \leq 15$ [33]. Typical order parameter maps and director streamlines are shown in Fig. 15. The effect of the softened defect regions on the effective interparticle forces is particularly strong when they approach along a vector normal to the director. The elastic theory prediction [34] is that at long range the effective forces are derived from a potential of the form

$$\mathcal{U} \propto R^{-5}\left[9 - 20\cos 2\psi + 35\cos 4\psi\right] \tag{4}$$

where $\psi = \pi/2 - \theta$. The proportionality constant depends on the elastic constants, but the angular dependence is independent of them. The potential

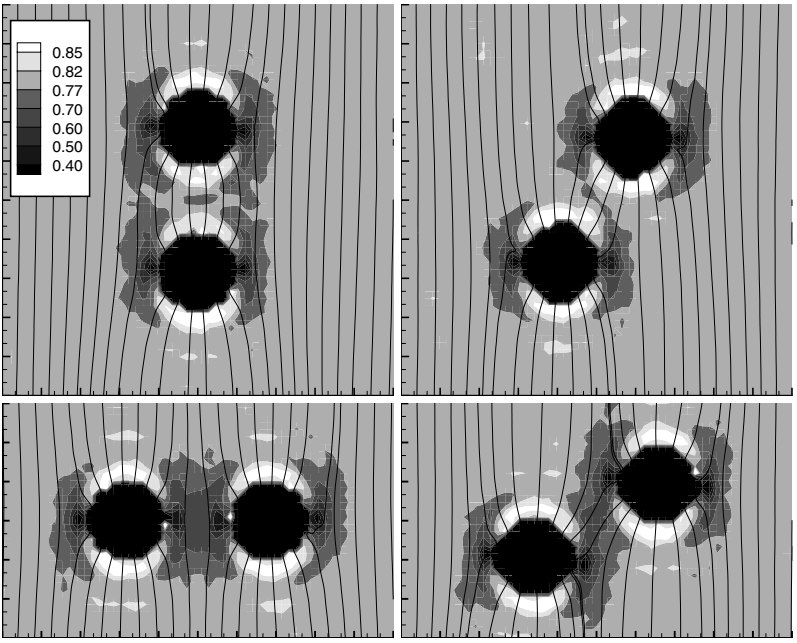


Fig. 15. Slices through the 3D order parameter maps for two macroparticles of diameter $D = 6$ at a separation $R = 11$ in a solvent of ellipsoids with axes $A = 3$, $B = 1$. The interparticle vector is at angles of $\theta = 0°, 30°, 60°$ and $90°$ relative to the bulk nematic director. The lines are streamlines of the director field

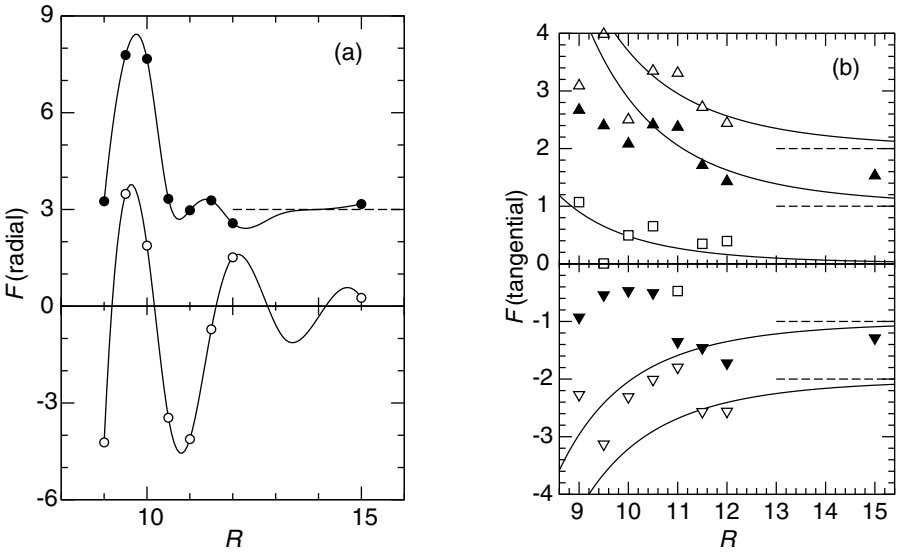


Fig. 16. Effective force between spherical macroparticles in a nematic solvent, for various angles of approach (**a**) parallel and (**b**) perpendicular to the line of centres. The results for each angle are shifted vertically with respect to each other for clarity, with the zero axis indicated by a dashed line in each case. Open circles: $\theta = 0°$; open up triangles: $\theta = 15°$; filled up triangles: $\theta = 30°$; open squares: $\theta = 45°$; filled down triangles: $\theta = 60°$; open down triangles: $\theta = 75°$; filled circles: $\theta = 90°$. The curves on the left are simply to guide the eye; on the right they are the predictions of (4). After [33]

has minima with respect to angular variation at $\psi = \frac{1}{2}\arctan(4\sqrt{3}) \approx \pm 41°$, that is at $\theta \approx 49°, 131°$. Now, it is unreasonable to expect this result to apply at the relatively short distances used in the molecular simulation, and the parallel components of the force are certainly dominated by other effects. However, Fig. 16 shows that the transverse force (which would be identically zero if the solvent were not a liquid crystal) is less affected by the short-range structure: indeed, with the eye of faith one can see some agreement with the curves predicted by (4), and the sign of the transverse force does seem to switch at about the right critical angle of $\theta \approx 49°$.

5 Conclusions

This chapter has covered some aspects of liquid crystal simulations that are unique to this fascinating state of matter. Naturally, a wide variety of other simulation techniques are used to map out the rich phase diagrams of these systems: descriptions of these have been omitted here, to minimize overlap with material covered in other chapters. Some discussion of these techniques,

as well as methods used to study liquid crystal dynamical properties, may be found elsewhere [35]. Limitations of space have also meant that the substantial literature on lattice spin models of orientational ordering in liquid crystals has not been referenced [36–41]. Many of the effects in liquid crystals cannot be properly tackled by molecular simulations. Mesoscopic simulation techniques [42–44] are especially promising, as they can reach length- and time-scales which are out of reach of molecular simulation methods. However, molecular calculations are still useful to relate the coefficients which are needed as input parameters for these methods, to molecular structure and interactions.

Acknowledgements

It is a pleasure to acknowledge the contributions of all my collaborators on various pieces of work referenced herein. This research has been supported by the Engineering and Physical Sciences Research Council, the Alexander von Humboldt foundation, the Overseas Research Students Award Scheme, the British Council, and the Leverhulme Trust. Some of the simulations reported here were performed on the computing facilities of the Centre for Scientific Computing, University of Warwick.

References

1. J. G. Gay and B. J. Berne (1981) Modification of the overlap potential to mimic a linear site-site potential. *J. Chem. Phys.* **74**, pp. 3316–3319
2. E. de Miguel, L. F. Rull, M. K. Chalam, and K. E. Gubbins (1990) Liquid-vapor coexistence of the Gay-Berne fluid by Gibbs ensemble simulation. *Molec. Phys.* **71**, pp. 1223–1231
3. R. Berardi, A. P. J. Emerson, and C. Zannoni (1993) Monte Carlo investigations of a Gay-Berne liquid crystal. *J. Chem. Soc. Faraday Trans.* **89**, pp. 4069–4078
4. E. de Miguel, E. Martín del Río, J. T. Brown, and M. P. Allen (1996) Effect of the attractive interactions on the phase behavior of the Gay-Berne liquid crystal model. *J. Chem. Phys.* **105**, pp. 4234–4249
5. J. T. Brown, M. P. Allen, E. Martín del Río, and E. de Miguel (1998) Effects of elongation on the phase behavior of the Gay-Berne fluid. *Phys. Rev. E* **57**, pp. 6685–6699
6. R. Berardi, C. Fava, and C. Zannoni (1995) A generalized Gay-Berne intermolecular potential for biaxial particles. *Chem. Phys. Lett.* **236**, pp. 462–468
7. P. J. Camp and M. P. Allen (1997) Phase diagram of the hard biaxial ellipsoid fluid. *J. Chem. Phys.* **106**, pp. 6681–6688
8. P. J. Camp, M. P. Allen, and A. J. Masters (1999) Theory and computer simulation of bent-core molecules. *J. Chem. Phys.* **111**, pp. 9871–9881
9. J. S. van Duijneveldt and M. P. Allen (1997) Computer simulation study of a flexible-rigid-flexible model for liquid crystals. *Molec. Phys.* **92**, pp. 855–870
10. A. V. Lyulin, M. S. Al-Barwani, M. P. Allen, M. R. Wilson, I. Neelov, and N. K. Allsopp (1998) Molecular dynamics simulation of main chain liquid crystalline polymers. *Macromolecules* **31**, pp. 4626–4634

11. J. S. van Duijneveldt, A. Gilvillegas, G. Jackson, and M. P. Allen (2000) Simulation study of the phase behavior of a primitive model for thermotropic liquid crystals: Rodlike molecules with terminal dipoles and flexible tails. *J. Chem. Phys.* **112**, pp. 9092–9104
12. R. Berardi, S. Orlandi, and C. Zannoni (2003) Molecular dipoles and tilted smectic formation: a Monte Carlo study. *Phys. Rev. E* **67**, p. 041708
13. R. Berardi, M. Fehervari, and C. Zannoni (1999) A Monte Carlo simulation study of associated liquid crystals. *Molec. Phys.* **97**, pp. 1173–1184
14. D. L. Cheung, S. J. Clark, and M. R. Wilson (2002) Parametrization and validation of a force field for liquid-crystal forming molecules. *Phys. Rev. E* **65**, p. 051709
15. R. Berardi, L. Muccioli, and C. Zannoni (2004) Can nematic transitions be predicted by atomistic simulations? A computational study of the odd-even effect. *Chem. Phys. Chem.* **5**, pp. 104–111
16. C. Oseen (1933) Theory of liquid crystals. *Trans. Faraday Soc.* **29**, pp. 883–899
17. F. C. Frank (1958) On the theory of liquid crystals. *Discuss. Faraday Soc.* **25**, pp. 19–28
18. L. Onsager (1949) The effects of shape on the interaction of colloidal particles. *Ann. N. Y. Acad. Sci.* **51**, p. 627
19. M. P. Allen, M. A. Warren, M. R. Wilson, A. Sauron, and W. Smith (1996) Molecular dynamics calculation of elastic constants in Gay-Berne nematic liquid crystals. *J. Chem. Phys.* **105**, pp. 2850–2858
20. M. P. Allen (1993) Calculating the helical twisting power of dopants in a liquid crystal by computer simulation. *Phys. Rev. E* **47**, pp. 4611–4614
21. M. J. Cook and M. R. Wilson (2000) Calculation of helical twisting power for liquid crystal chiral dopants. *J. Chem. Phys.* **112**, pp. 1560–1564
22. G. Germano, M. P. Allen, and A. J. Masters (2002) Simultaneous calculation of the helical pitch and the twist elastic constant in chiral liquid crystals from intermolecular torques. *J. Chem. Phys.* **116**, pp. 9422–9430
23. D. J. Earl and M. R. Wilson (2004) Calculations of helical twisting powers from intermolecular torques. *J. Chem. Phys.* **120**, pp. 9679–9683
24. P. G. de Gennes and J. Prost (1995) *The Physics of Liquid Crystals* (Oxford: Clarendon Press, second, paperback ed)
25. M. P. Allen (1999) Molecular simulation and theory of liquid crystal surface anchoring. *Molec. Phys.* **96**, pp. 1391–1397
26. D. Andrienko, G. Germano, and M. P. Allen (2000) Liquid crystal director fluctuations and surface anchoring by molecular simulation. *Phys. Rev. E* **62**, pp. 6688–6693
27. M. P. Allen (2000) Molecular simulation and theory of the isotropic-nematic interface. *J. Chem. Phys.* **112**, pp. 5447–5453
28. A. J. McDonald, M. P. Allen, and F. Schmid (2000) Surface tension of the isotropic-nematic interface. *Phys. Rev. E* **63**, pp. 010701(R)/1–4
29. P. Poulin, V. Cabuil, and D. A. Weitz (1997) Direct measurement of colloidal forces in an anisotropic solvent. *Phys. Rev. Lett.* **79**, pp. 4862–4865
30. P. Poulin, H. Stark, T. C. Lubensky, and D. A. Weitz (1997) Novel colloidal interactions in anisotropic fluids. *Science* **275**, pp. 1770–1773
31. J. L. Billeter and R. A. Pelcovits (2000) Defect configurations and dynamical behavior in a Gay-Berne nematic emulsion. *Phys. Rev. E* **62**, pp. 711–717

32. D. Andrienko, G. Germano, and M. P. Allen (2001) Computer simulation of topological defects around a colloidal particle or droplet dispersed in a nematic host. *Phys. Rev. E* **63**, pp. 041701/1–8
33. M. S. Al-Barwani, G. Sutcliffe, and M. P. Allen (2004) Forces between two colloidal particles in a nematic solvent. *J. Phys. Chem. B* **108**, pp. 6663–6666
34. R. W. Ruhwandl and E. M. Terentjev (1997) Long-range forces and aggregation of colloidal particles in a nematic liquid crystal. *Phys. Rev. E* **55**, pp. 2958–2961
35. M. P. Allen (2004) Liquid crystal systems. In *Computational soft matter: from synthetic polymers to proteins* (N. Attig, K. Binder, H. Grubmüller, and K. Kremer, eds.) **23** of *NIC Series*, (Jülich), pp. 289–320, John von Neumann Institute for Computing, NIC-Directors
36. G. R. Luckhurst and P. Simpson (1982) Computer simulation studies of anisotropic systems. VIII. The Lebwohl-Lasher model of nematogens revisited. *Molec. Phys.* **47**, pp. 251–265
37. U. Fabbri and C. Zannoni (1986) A Monte-Carlo investigation of the Lebwohl-Lasher lattice model in the vicinity of its orientational phase transition. *Molec. Phys.* **58**, pp. 763–788
38. Z. Zhang, M. J. Zuckermann, and O. G. Mouritsen (1993) Phase transition and director fluctuations in the 3-dimensional Lebwohl-Lasher model of liquid crystals. *Molec. Phys.* **80**, pp. 1195–1221
39. R. Hashim and S. Romano (1999) Computer simulation study of a nematogenic lattice model based on the Nehring-Saupe interaction potential. *Int. J. Mod. Phys. B* **13**, pp. 3879–3902
40. C. Chiccoli, Y. Lansac, P. Pasini, J. Stelzer, and C. Zannoni (2002) Effect of surface orientation on director configurations in a nematic droplet. A Monte Carlo simulation. *Mol. Cryst. Liq. Cryst.* **372**, pp. 157–165
41. M. P. Allen (2005) Spin dynamics for the Lebwohl-Lasher model. *Phys. Rev. E* **72**, p. 036703
42. D. Svenšek and S. Žumer (2002) Hydrodynamics of pair-annihilating disclination lines in nematic liquid crystals. *Phys. Rev. E* **66**, p. 021712
43. R. Yamamoto, Y. Nakayama, and K. Kim (2004) A smooth interface method for simulating liquid crystal colloid dispersions. *J. Phys. Cond. Mat.* **16**, pp. S1945–S1955
44. C. Denniston, D. Marenduzzo, E. Orlandini, and J. M. Yeomans (2004) Lattice Boltzmann algorithm for three-dimensional liquid-crystal hydrodynamics. *Phil. Trans. Roy. Soc. A* **362**, pp. 1745–1754

Coarse-Grained Models of Complex Fluids at Equilibrium and Under Shear

F. Schmid

Fakultät für Physik, Universität Bielefeld, Universitätsstraße 25, 33615 Bielefeld
schmid@physik.uni-bielefeld.de



Friederike Schmid

F. Schmid: *Coarse-Grained Models of Complex Fluids at Equilibrium and Under Shear*, Lect. Notes Phys. **704**, 211–258 (2006)
DOI 10.1007/3-540-35284-8_10
© Springer-Verlag Berlin Heidelberg 2006

1 Introduction 213

2 Coarse-Grained Models for Surfactant Layers 214

2.1 Introduction 214
2.2 Coarse-Grained Molecular Models 217
2.3 Mesoscopic Membrane Models 228
2.4 Summary 232

3 Liquid Crystals and Surfactant Layers Under Shear 233

3.1 Introduction 233
3.2 Simulating Shear on the Particle Level: NEMD 239
3.3 Simulations at the Mesoscopic Level 248
3.4 Summary 250

4 Conclusions 250

References 251

Complex fluids exhibit structure on a wide range of length and time scales, and hierarchical approaches are necessary to investigate all facets of their often unusual properties. The study of idealized coarse-grained models at different levels of coarse-graining can provide insight into generic structures and basic dynamical processes at equilibrium and non-equilibrium.

In the first part of this chapter, some popular coarse-grained models for membranes and membrane systems are reviewed. Special focus is given to bead-spring models with different solvent representations, and to random-interface models. Selected examples of simulations at the molecular and the mesoscopic level are presented, and it is shown how simulations of molecular coarse-grained models can bridge between different levels.

The second part addresses simulation methods for complex fluids under shear. After a brief introduction into the phenomenology (in particular for liquid crystals), different non-equilibrium molecular dynamics (NEMD) methods are introduced and compared to one another. Application examples include the behavior of liquid crystal interfaces and lamellar surfactant phases under shear. Finally, mesoscopic simulation approaches for liquid crystals under shear are briefly discussed.

1 Introduction

The term "Complex fluid" or "soft condensed matter" – these two are often used interchangeably – refers to a broad class of materials, which are usually made of large organic molecules and have a number of common features [1, 2]: They display structure on a nanoscopic or mesoscopic scale; characteristic energies are of the order of k_BT at room temperature, hence the properties are to a large extent dominated by entropic effects, and the materials respond strongly to weak external forces [3]; the characteristic response times span one or several orders of magnitude, and the rheological properties of the fluids are typically non-Newtonian [2, 4]. Some examples are polymer melts and solutions [5–8], emulsions, colloids [9–11], amphiphilic systems [12, 13], and liquid crystals [14–16].

Computer simulations of complex fluids are particularly challenging, due to the hierarchy of length and time scales that contributes to the material properties. With current computer resources, it is practically impossible to describe all aspects of such a material within one single theoretical model. Therefore, one commonly uses different models for different length and time scales. The idea is to eliminate the small-scale degrees of freedom in the large-scale models, and to incorporate the details of the small-scale properties in the parameters of a new, simpler, model. This is called coarse-graining.

There are two aspects to coarse-graining. First, systematic coarse-graining procedures must be developed and applied in order to study materials quantitatively on several length and time scales. This is the challenge of multiscale

modeling, one major growth area in materials science. It is, however, not the subject of the present chapter. Some recent reviews can be found in [17, 18]

Second, coarse-grained idealized models are used to study generic properties of materials on a given scale, and physical phenomena which are characteristic for a whole class of materials. Here, microscopic details are disregarded because they are deemed irrelevant for the physical properties under consideration. This approach to materials science has a long tradition, going back to Ising's famous lattice model for magnetism.

In the present chapter, we shall concentrate on coarse-grained models of the second type. We shall discuss how such models can be constructed, how they can be studied by computer simulations, and how such studies help to improve our understanding of equilibrium and nonequilibrium properties of complex materials. We will not attempt to present an overview – this would require a separate book – but rather focus on selected case studies. In the first part, we will consider coarse-grained model systems for amphiphilic membranes and their use for studies of equilibrium properties. In the second part, we will turn to complex fluids under shear, i.e., nonequilibrium systems, and discuss coarse-grained approaches for membrane systems and liquid crystals.

2 Coarse-Grained Models for Surfactant Layers

In this section, we shall discuss coarse-grained simulation models and methods for amphiphilic membranes and membrane stacks. After a brief introduction into the phenomenology and some general remarks on coarse-graining, we will focus on two particular types of coarse-grained models: Particle-based bead-spring models and mesoscopic membrane models. We shall present some typical examples, and illustrate their use with applications from the literature.

2.1 Introduction

Amphiphilic membranes are a particular important class of soft material structures, because of their significance for biology [19–21]. Biomembranes are omnipresent in all living beings, they delimit cells and create compartments, and participate in almost all biological functions. Figure 1 shows an artist's view on such a biomembrane. It is mainly made of two coupled layers of lipids, which serve as a matrix for embedded proteins and other molecules.

This structure rests on a propensity of lipids to self-assemble into bilayers when exposed to water. The crucial property of lipids which is responsible for this behavior is their amphiphilic character, i.e., the fact that they contain both hydrophilic (water loving) and hydrophobic (water hating) parts. Most lipid molecules have two non polar (hydrophobic) hydrocarbon chain (tails), which are attached to one polar (hydrophilic) head group. Figure 2 shows some typical examples. The tails vary in their length, and in the number and position of double bonds between carbon atoms. Whereas single bonds

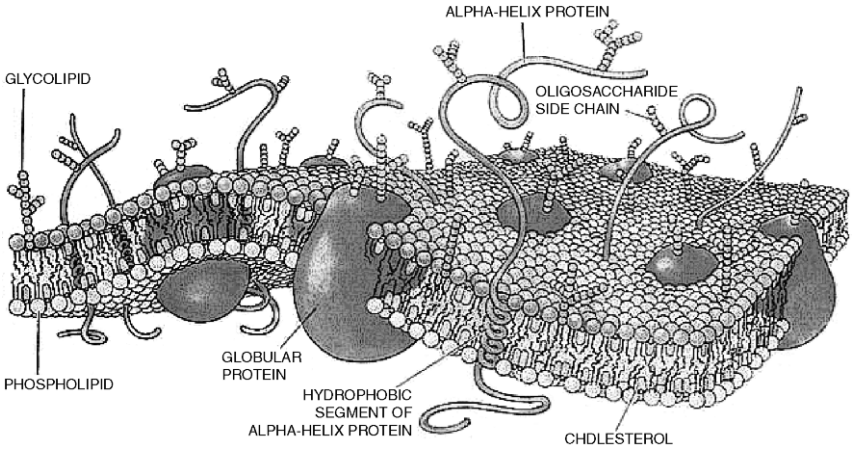


Fig. 1. Artist's view of a biomembrane. Source: NIST

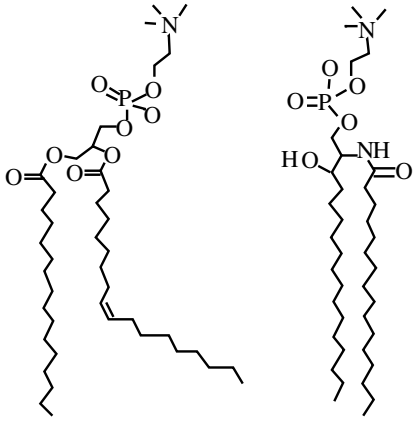


Fig. 2. Examples of membrane lipids. *Left*: Phosphatidylcholine (unsaturated); *Right*: Sphingomyelin

in a hydrocarbon chain are highly flexible, in the sense that chains can rotate easily around such a bond, double bonds are stiff and may enforce kinks in the chain. Chains with no double bonds are called saturated. The variety of polar head groups is even larger. Overviews can be found in [19].

By assembling into bilayers, the lipids can arrange themselves such that the head groups shield the tails from the water environment. Therefore, lipids in lipid-water mixtures often tend to form lamellar stacks of membranes separated by thin water layers (Fig. 3). Such stacks can typically hold up to 20% water. In the more dilute regime, membranes may close up to form vesicles or other more disordered structures. Vesicles play an important role in biological systems, since they provide closed compartments which can be used to store or

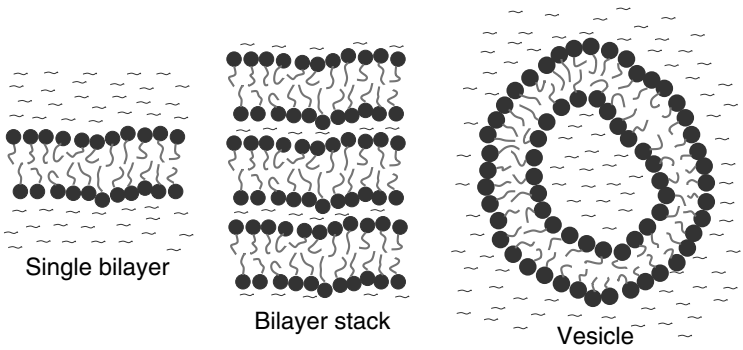


Fig. 3. Self-assembled bilayer structures

transport substances. Single, isolated membranes are metastable with respect to dissociation for entropic reasons, but they have very long lifetimes.

Depending on the particular choice of lipids and the properties of the surrounding aqueous fluid (pH, salt content etc.), several other structures can also be observed in lipid-water systems – spherical and cylindrical micelles, ordered structures with hexagonal or cubic symmetry etc. [22]. In this chapter, we shall only regard bilayer structures.

Membranes also undergo internal phase transitions. Figure 4 shows schematically some characteristic phases of one-component membranes. The most prominent phase transition is the "main" transition, the transition from the liquid state into one of the gel states, where the layer thickness, the lipid mobility, and the conformational order of the chains, jump discontinuously as a function of temperature. In living organisms, most membranes are maintained in the liquid state. Nevertheless, the main transition presumably has some biological relevance, since it is very close to the body temperature for some of the most common lipids (e.g., saturated phospholipids) [19].

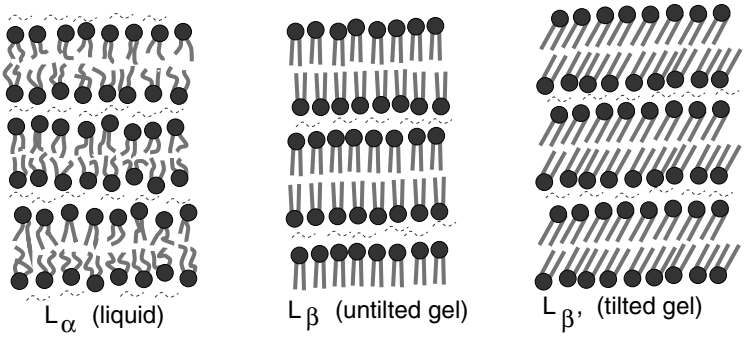


Fig. 4. Some prominent membrane phases

In sum, several aspects of membranes must be studied in order to understand their structure and function, in biological as well as in artificial systems:

- Self-assembling mechanisms
- The inner structure and internal phase transitions, and the resulting membrane properties (permeability, viscosity, stiffness)
- The organization of membranes on a mesoscopic scale (vesicles, stacks, etc.)
- The interaction of membranes with macromolecules.

These topics relate to the physics and chemistry of amphiphiles on many length and time scales, and several driving forces contribute to the unique properties of membranes: On the scale of atoms and small molecules, one has the forces which are responsible for the self-assembly: The hydrophobic effect, and the interactions between water and the hydrophilic head groups. On the molecular scale, one has the interplay of local chain conformations and chain packing, which is responsible for the main transition and determines the local structure and fluidity of the membrane. On the supramolecular scale, one has the forces which govern the mesoscopic structure and dynamics of membrane systems: The membrane elasticity, the thermal fluctuations, the hydrodynamic interactions with the surrounding fluid, the thermodynamic forces that drive phase separation in mixed membranes etc.

Studying all these aspects within one single computer simulation model is clearly impossible, and different models have been devised to study membranes at different levels of coarse-graining. Among the oldest approaches of this kind are the Doniach [23] and the Pink model [24, 25], generalized Ising models, which have been used to investigate almost all aspects of membrane physics on the supramolecular scale [21,26]. With increasing computer power, numerous other simulation models have become treatable in the last decades, which range from atomistic models, molecular coarse-grained models, up to mesoscopic models that are based on continuum theories for membranes.

In the following, we shall focus on two important examples: Coarse-grained chain models designed to describe membranes on the molecular scale, and random interface models that describe membranes on the mesoscopic scale. We stress that this our overview is not complete, and there exist other successful approaches, which are not included in our presentation. A relatively recent review can be found in [27].

2.2 Coarse-Grained Molecular Models

In coarse-grained molecular models, molecules are still treated as individual particles, but their structure is highly simplified. Only the ingredients deemed essential for the material properties are kept. For amphiphiles, these are the amphiphilic character (for the self-assembly), and the molecular flexibility (for the internal phase transitions).

Spring-Bead Models

The first coarse grained molecular models have been formulated on a lattice, relatively shortly after the introduction of the above-mentioned Pink model [24, 25]. One particularly popular model is the Larson model [28, 29]: Water molecules (w) occupy single sites of a cubic lattice, and amphiphile molecules are represented by chains of "tail" (t) and "head" (h) monomers. Only particles on neighbor lattice sites interact with each another. The lattice is entirely filled with w, h, or t particles. If one further takes all hydrophilic particles, h and w, to be identical, the interaction energy is determined by a single interaction parameter, which describes the relative repulsion between hydrophobic and hydrophilic particles. The model reproduces self-assembly and exhibits many experimentally observed phases, i.e., the lamellar phase, the hexagonal phase, the cubic micellar phase, and even the bicontinuous gyroid phase [29, 30].

Lattice models can be simulated efficiently by Monte Carlo methods. However, they have the obvious drawback of imposing an a priori anisotropy on space, which restricts their versatility. For example, internal membrane phase transitions and tilted phases cannot be studied easily. In dynamical studies, one is restricted to using Monte Carlo dynamics, which is unrealistic. Therefore, most more recent investigations rely on off-lattice models.

Off-lattice molecular approaches to study self-assembling amphiphilic systems have been applied for roughly 15 years. Presumably the first model was introduced by B. Smit et al. [31] in 1990. A schematic sketch is shown in Fig. 5 a). As in the Larson model, the amphiphiles are represented by chains made of very simple h- or t-units, which are in this case spherical beads. "Water" molecules are represented by free beads. Beads in a chain are joined together by harmonic springs, where a cutoff is sometimes imposed in order to ensure that the beads cannot move arbitrarily far apart from one another. Non-bonded beads interact via simple short-range pairwise potentials, e.g., a truncated Lennard-Jones potential. The parameters of the potentials are chosen such that ht pairs and hw pairs effectively repel each other. The Smit model reproduces self-assembly, micelle formation and membrane formation [32–36].

Most molecular off-lattice models are modifications of the Smit model [27, 37–39]. They differ from one another in the specific choice of the interaction potentials between non-bonded beads, in the bond potentials between adjacent beads on the chain, in the chain architecture (number of head and tail beads, number of tails), and sometimes contain additional potentials such as bending potentials. Nowadays, Smit models are often studied with dissipative particle (DPD) dynamics, and the interaction potentials are the typical soft DPD potentials [37, 39]. A "water particle" is then taken to represent a lump of several water molecules. An introduction into the DPD method is given in B. Dünweg's chapter (this book). It works very well for membrane simulations as long as the time step is not too large [40].

This approach is presented in more detail in B. Smit's chapter (this book).

Implicit Solvent and Phantom Solvent

The primary requirement for a lipid membrane model is to ensure that the lipids maintain a bilayer structure. In the models discussed in the previous Sect. 2.2, the bilayers are stabilized by the surrounding solvent particles. If the interactions between w, h, and t particles are chosen suitably, the solvent particles and the amphiphiles arrange themselves such that the hydrophilic and hydrophobic particles are separated (microphase separation).

This "explicit solvent" approach is in some sense the most natural one and has several advantages: The bilayers are fully flexible, and they are surrounded by a fluid. The fact that the fluid is composed of large beads as opposed to small water molecules is slightly problematic, since the solvent structure may introduce correlations, e.g., in simulations with periodic boundary conditions. However, the correlations disappear if the membranes or their periodic images are separated by a large amount of solvent.

On the other hand, explicit solvent models also have a disadvantage: The simulation time depends linearly on the *total* number of beads. A large amount of computing time is therefore spent on the uninteresting solvent particles. Therefore, effort has been spent on designing membrane models without explicit solvent particles.

One early idea has been to force the amphiphiles into sheets by tethering the head groups to two-dimensional opposing surfaces, e.g., by a harmonic potential, or by rigid constraints [41–45] (Fig. 5 b)). This model, which we call "sandwich model", is simple, extremely efficient, and the configurations are easy to analyze, since the bilayer is always well-defined. On the other hand, the bilayer has no flexibility, i.e., undulations and protrusions [46] are supressed and important physics is lost.

A second approach is to eliminate the solvent, and represent its effect on the amphiphiles by appropriate effective interactions between monomers (Fig. 5 c)). This way of modeling solvent is common in polymer simulations. Nevertheless, producing something as complex as membrane self-assembly is a non-trivial task. The first implicit solvent model was proposed only rather recently, in 2001, by Noguchi and Takasu [47]. They mimick the effect of the

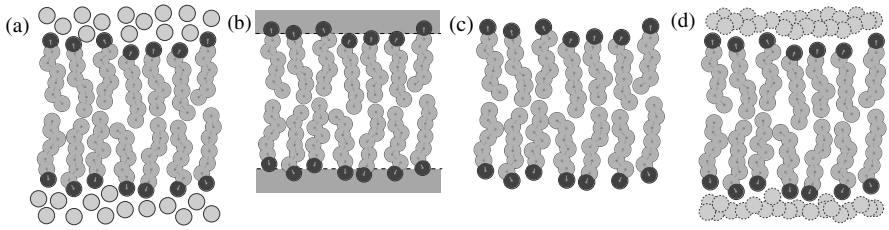


Fig. 5. Schematic picture of bead-spring models for self-assembling membranes. (**a**) Explicit solvent model (**b**) Sandwich model (**c**) Solvent free model (**d**) Phantom solvent model

solvent by a multibody "hydrophobic potential", which depends on the local density of hydrophobic particles. From a physical point of view, the introduction of multibody potentials in an implicit solvent model is reasonable – multibody potentials emerge automatically whereever degrees of freedom are integrated out systematically. Nevertheless, people usually tend to favor pair potentials. Farago [48] and Cooke et al. [49] have shown that it is possible to stabilize fluid bilayers with purely pairwise interactions in a solvent-free model. The crucial requirement seems to be that the interactions between hydrophobic beads have a smooth attractive part of relatively long range [49].

The clear benefit of implicit solvent models is their high efficiency. They are very well suited to study large scale problems like vesicle formation [47,49], vesicle fusion [50], vesicles subject to external forces [51], phase separation on vesicles [49] etc. On the other hand, they also have limitations: They require tricky potentials; the pressure is essentially zero (since the membranes are surrounded by empty space); also, they cannot be used for dynamical studies where the hydrodynamical coupling with the solvent becomes important.

A third approach was recently put forward by ourselves: the "phantom solvent" model [45] (Fig. 5d)). We introduce explicit solvent particles, which do not interact with one another, only with the amphiphiles. The amphiphiles perceive them as repulsive soft beads. In the bulk, they simply form an ideal gas. They have the simple physical interpretation that they probe the free volume which is still accessible to the solvent, once the membrane has self-assembled. Thus the self-assembly is to a large extent driven entropically.

In Monte Carlo simulations, this model is just as efficient as solvent-free models, because Monte Carlo moves of phantom beads that are not in contact with a membrane are practically free of (computational) charge. It is robust, we do not need to tune the lipid potentials in order to achieve self-assembly. Pressure can be applied if necessary, and with a suitable dynamical model for the solvent (e.g., dissipative particle dynamics), we can also study (hydro)dynamical phenomena. Compared with Smit-type models, the phantom solvent model has the advantage that the solvent is structureless and cannot induce artificial correlations. On the other hand, the fact that it is compressible like a gas, may cause problems in certain dynamical studies.

To summarize, we have presented several possible ways to model the solvent environment of a self-assembled membrane, and discussed the advantages and disadvantages. The different models are sketched schematically in Fig. 5.

We shall now illustrate the potential of molecular coarse-grained models with two examples.

First Application Example: The Ripple Phase

Our first example shows how coarse-grained simulations can help to understand the inner structure of membranes. We have introduced earlier the liquid and gel phases in membranes (Fig. 4). However, one rather mysterious phase was missing in our list: The ripple phase ($P_{\beta'}$). It emerges as an intermediate

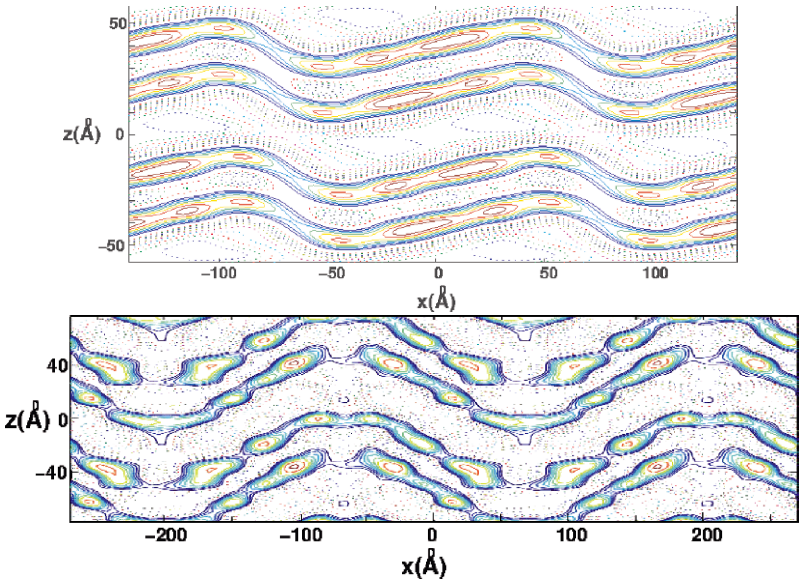


Fig. 6. Electron density maps of (*top*) an asymmetric rippled state (DMPC at 18.2 degrees) and (*bottom*) a symmetric rippled state (DPPC at 39.2°C). Reprinted with permission from Sengupta, Raghunathan, Katsaras [52] (rescaled). Copyright 2003 by the American Physical Society

state between the tilted gel phase ($L_{\beta'}$) and the liquid phase (L_α), and according to freeze-etch micrographs, AFM pictures, and X-ray diffraction studies, it is characterized by periodic wavy membrane undulations. More precisely, one observes two different rippled structures: Figure 6 shows electron density maps of these two states, as calculated from X-ray data by Sengupta et al [52]. The more commonly reported phase is the asymmetric ripple phase, which is characterized by sawtooth profile and has a wavelength of about 150 Angstrom. The symmetric rippled phase tends to form upon cooling from the liquid L_α phase and has a period twice as long. It is believed to be metastable. The molecular structure of both rippled states has remained unclear until very recently.

Atomistic simulations of de Vries et al, have finally shed light on the microscopic structure of the asymmetric ripple phase [53]. These authors observed an asymmetric ripple in a Lecithin bilayer, which had a structure that was very different from the numerous structures proposed earlier. Most strikingly, the rippled membrane is not a continuous bilayer. The ripple is a membrane defect where a single layer spans the membrane, connecting the two opposing sides (Fig. 7).

The simulations of de Vries et al seem to clarify the structure of the asymmetric ripple phase, at least for the case of Lecithin. Still the question remains whether their result can be generalized to other lipid layers, or whether it is

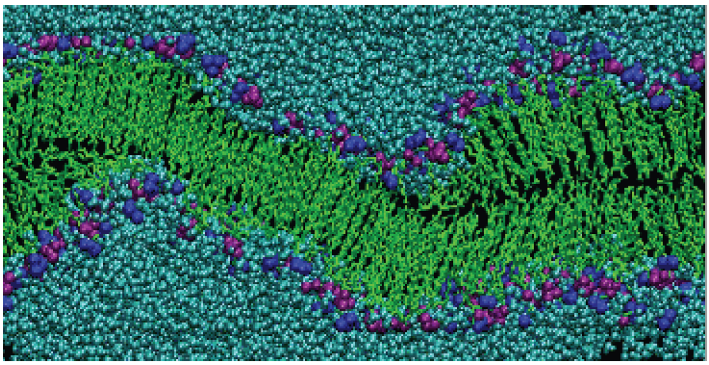


Fig. 7. Asymmetric ripple in a Lecithin Bilayer, as observed in an atomistic simulation. From de Vries et al. [53]. Copyright 2005 National Academy of Sciences, U.S.A.

related to the specific properties of Lecithin head groups. This question can be addressed with idealized coarse-grained models.

Rippled phases had also been observed in simulations of coarse-grained models. Kranenburg et al have reported the existence of a rippled state in a Smit model with soft DPD interactions [54, 55]. The structure of their ripple is different from that of Fig. 7. Judging from this result, one is tempted to conclude that the structure of de Vries et al may not be generic.

However, we recently found ripples with a structure very similar to that of Fig. 7 in simulations of a simple bead-spring model by (Lenz and Schmid [56, 57]). Our lipids are represented by linear chains with six tail beads (size σ) and one head bead (size 1.1σ), which are connected by springs of size 0.7σ. Tail beads attract one another with a truncated and shifted Lennard-Jones potential. Head beads are slightly larger and purely repulsive. The solvent environment is modelled by a fluid of phantom beads of size σ. The system was studied with Monte Carlo simulations at constant pressure, using a simulation box of fluctuating size and shape in order to ensure that the pressure tensor is diagonal.

Membranes were found to self-assemble spontaneously at sufficiently low temperatures, they exhibit a fluid L_α-phase as well as a tilted gel $L_{\beta'}$-phase. The two phases are separated by a temperature region where ripples form spontaneously. In small systems, the ripples are always asymmetric. In larger systems, asymmetric ripples form upon heating from the $L_{\beta'}$-phase. Upon cooling from the L_α-phase, a second, symmetric, type of ripple forms, which has twice the period than the asymmetric ripple and some structural similarity, except that the bilayer remains continuous (see Fig. 8). The sideview of this second ripple has striking similarity with the density maps of Fig. 6 b).

The coarse-grained simulations of Lenz thus not only demonstrate that the asymmetric ripple structure found by de Vries is generic, and can be observed

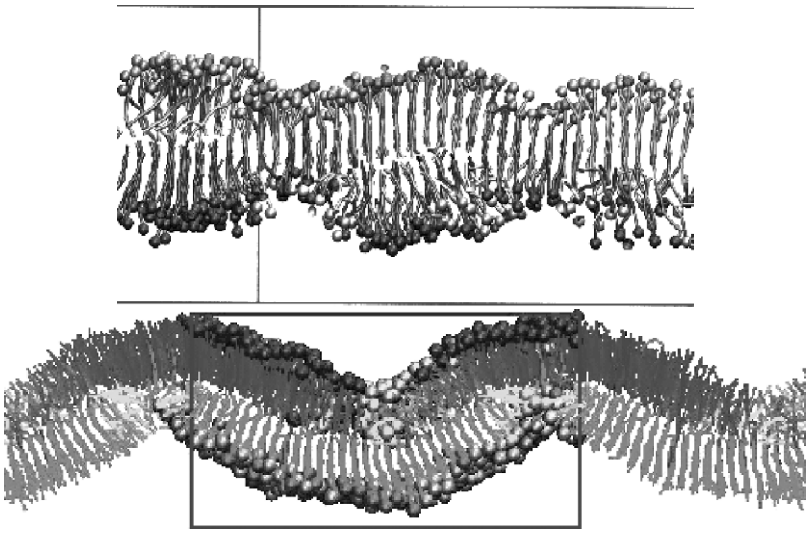


Fig. 8. Ripples from simulations of a coarse-grained model. (**a**) Asymmetric ripple (**b**) Symmetric ripple

in highly idealized membrane models as well. They also suggest a structural model for the yet unresolved symmetric ripple.

Second Application Example: Membrane Stacks

In our second example, we consider the fluctuations and defects in stacks of membranes. We shall show how simulations of coarse-grained membrane models can be used to test and verify mesoscopic models, and how they can bridge between coarse-graining levels.

At the mesoscopic level, membranes are often represented by random interfaces (see also Sect. 2.3). One of the simplest theories for thermal fluctuations in membrane stacks is the "discrete harmonic model" [58]. It describes membranes without surface tension, and assumes that the fluctuations are governed by two factors: The bending stiffness K_c of single membranes, and a penalty for compressing or swelling the stack, which is characterized by a compressibility modulus B. The free energy is given in harmonic approximation

$$\mathcal{F}_{\text{DH}} = \sum_n \int_A dx\, dy \left\{ \frac{K_c}{2} \left(\frac{\partial^2 h_n}{\partial x^2} + \frac{\partial^2 h_n}{\partial y^2} \right)^2 + \frac{B}{2} (h_n - h_{n+1} + \bar{d})^2 \right\}, \quad (1)$$

where $\bar{d}$ is the average distance between layers. The first part describes the elasticity of single membranes, and the second part accounts for the interactions between membranes. Being quadratic, the free energy functional (1) is simple enough that statistical averages can be calculated exactly.

We have tested this theory in detail with coarse-grained molecular simulations (C. Loison et al. [59]) of a binary mixture of amphiphiles and solvent, within a Smit-type model originally introduced by Soddemann et al. [60]. The elementary units are spheres with a hard core radius σ, which may have two types: "hydrophilic" or "hydrophobic". Beads of the same type attract each other at distances less than 1.5σ. "Amphiphiles" are tetramers made of two hydrophilic and two hydrophobic beads, and "solvent" particles are single hydrophilic beads. We studied systems of up to 153600 beads with constant pressure molecular dynamics, using a simulation box of fluctuating size and shape, and a Langevin thermostat to maintain constant temperature.

At suitable pressures and temperatures, the system assumes a fluid lamellar phase. Our configurations contained up to 15 stacked bilayers, which contained 20 volume percent solvent. A configuration snapshot is shown in Fig. 9 (left). In order to test the theory (1), one must first determine the local positions of the membranes in the stack. This was done as follows:

1. The simulation box was divided into $N_x N_y N_z$ cells. ($N_x = N_y = 32$). The size of the cells fluctuates because the dimensions of the box fluctuate.
2. In each cell, the relative density of tail beads $\rho_{\text{tail}}(x, y, z)$ was calculated. It is defined as the ratio of the number of tail beads and the total number of beads.
3. The hydrophobic space was defined as the space where the relative density of tail beads is higher than a given threshold ρ_0. (The value of the threshold was roughly 0.7, i.e., 80% of the maximum value of ρ_{tail}).
4. Cells belonging to the hydrophobic space are connected to clusters. Two hydrophobic cells that share at least one vortex are attributed to the same cluster. Each cluster defines a membrane.
5. For each membrane n and each position (x, y), the two heights $h_n^{\min}(x, y)$ and $h_n^{\max}(x, y)$, where the density $\rho_{\text{tail}}(x, y, z)$ passes through the threshold ρ_0, are determined. The mean position is defined as the average $h_n(x, y) = (h_n^{\min} + h_n^{\max})/2$.

The algorithm identifies membranes even if they have pores. At the presence of other defects, such as necks or passages (connections between membranes), additional steps must be taken, which are not described here. A typical membrane conformation $h_n(x, y)$ is shown in Fig. 9 (right).

The statistical distribution of $h_n(x, y)$ can be analyzed and compared with theoretical predictions. For example, the transmembrane structure factor, which describes correlations between membrane positions in different membranes is given by [59]

$$s_n(q) = \langle h_m(\mathbf{q})^* h_{m+n}(\mathbf{q}) \rangle = s_0(q) \left[1 + \frac{X}{2} - \frac{1}{2}\sqrt{X(X+4)} \right]^n , \qquad (2)$$

where $h_n(\mathbf{q})$ is the Fourier transform of $h_n(x, y)$ in the (x, y)-plane and $X = q^4 K_c / B$ is a dimensionless parameter. The function $s_0(q)$ can also be calculated explicitly [59]. The prediction (2) can be tested by simply plotting

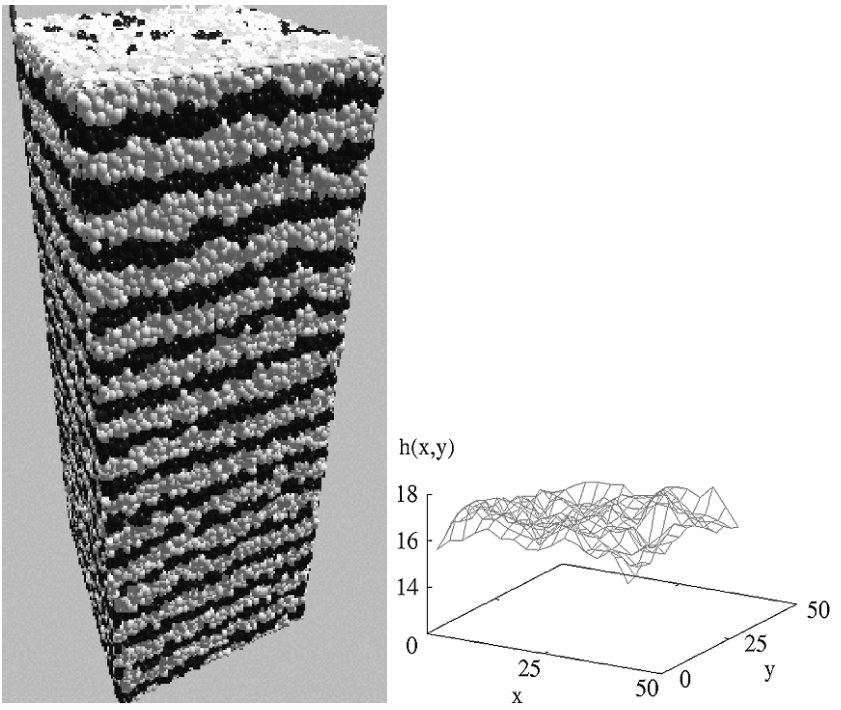


Fig. 9. *Left*: Snapshot of a bilayer stack (30720 amphiphiles and 30720 solvent beads). The "hydrophobic" tail beads are dark, the "hydrophilic" head beads and the solvent beads are light; *Right*: Snapshot of a single membrane position $h_n(x, y)$

the ratio s_n/s_0 vs. q for different n. The functional form of the curves should be given by the expression in the square brackets, with only one fit parameter K_c/B. Figure 10 shows the simulation data. The agreement with the theory is very good over the whole range of wave vectors q.

Thus the molecular simulations confirm the validity of the mesoscopic model (1). Moreover, the analysis described above yields a value for the phenomenological parameter K_c/B. By analyzing other quantities, one can also determine K_c and B separately. This gives the elastic parameters of membranes and their effective interactions, and establishes a link between the molecular and the mesoscopic description.

Many important characteristics of membranes not only depend on their elasticity, but also on their defects. For example, the permeability is crucially influenced by the number and properties of membrane pores. A number of atomistic and coarse grained simulation studies have therefore addressed pore formation [61–65], mostly in membranes under tension. In contrast, the membranes in the simulations of Loison et al. have almost zero surface tension. This turns out to affect the characteristics of the pores quite dramatically [66, 67].

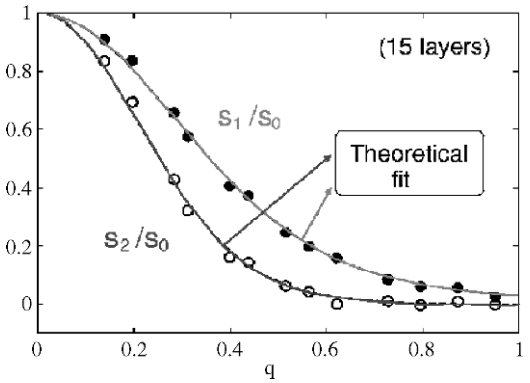


Fig. 10. Ratios of transmembrane structure factors s_1/s_0 and s_2/s_0 vs. in-plane wavevector q in units of σ^{-1}. The *solid lines* correspond to a theoretical fit to (2) with one (common) fit parameter K_c/B. After [59]

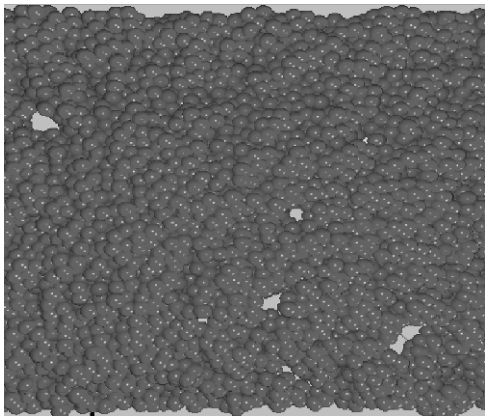


Fig. 11. Snapshot of a single bilayer (top view. Only hydrophobic beads are shown)

Figure 11 shows a snapshot of a hydrophobic layer which contains a number of pores. The pores have nucleated spontaneously. They "live" for a while, grow and shrink without diffusing too much, until they finally disappear. Most pores close very quickly, but a few large ones stay open for a long time. A closer analysis shows that pores are hydrophilic, i.e., the amphiphiles rearrange themselves at the pore edge, such that solvent beads in the pore center are mainly exposed to head beads. The total number of pores is rather high in our system. This is because the amphiphiles are short and the membranes are thin.

The analysis algorithm presented above not only localizes membranes, but also identifies pores, their position and their shape. Therefore one can again test the appropriate mesoscopic theories for pore formation. The simplest

Ansatz for the free energy of a pore with the area A and the contour length c has the form [68]

$$E = E_0 + \lambda\, c - \gamma\, A\ , \tag{3}$$

where E_0 is a core energy, λ a material parameter called line tension, and γ the surface tension. The second term describes the energy penalty at the pore rim. The last term accounts for the reduction of energy due to the release of surface tension in a stretched membrane. In our case, the surface tension is close to zero ($\gamma \approx 0$), because the simulation was conducted such that the pressure tensor is isotropic. Thus the last term vanishes.

In this simple free energy model, the pore shapes should be distributed according to a Boltzmann distribution, $P(c) \propto \exp(-\lambda c)$. Figure 12 (left) shows a histogram of contour lengths $P(c)$. The bare data do not reflect the expected exponential behavior. Something is missing. Indeed, a closer look reveals that the naive exponential Ansatz disregards an important effect: The "free energy" (3) gives only *local* free energy contributions, i.e., those stemming from local interactions and local amphiphile rearrangements. In addition, one must also account for the *global* entropy of possible contour length conformations. Therefore, we have to evaluate the "degeneracy" of contour lengths $g(c)$, and test the relation

$$P(c) \propto g(c) \exp(-\lambda c)\ . \tag{4}$$

Figure 12 (right) demonstrates that this second Ansatz describes the data very well. From the linear fit to the data, one can extract a value for the line tension λ.

The model (3) makes a second important prediction: Since the free energy only depends on the contour length, pores with the same contour length are equivalent, and the shapes of these pores should be distributed like those of two-dimensional self-avoiding ring polymers. In other words, they are not round, but have a fractal structure. From polymer theory, one knows that the size R_g of a two dimensional self-avoiding polymer scales roughly like

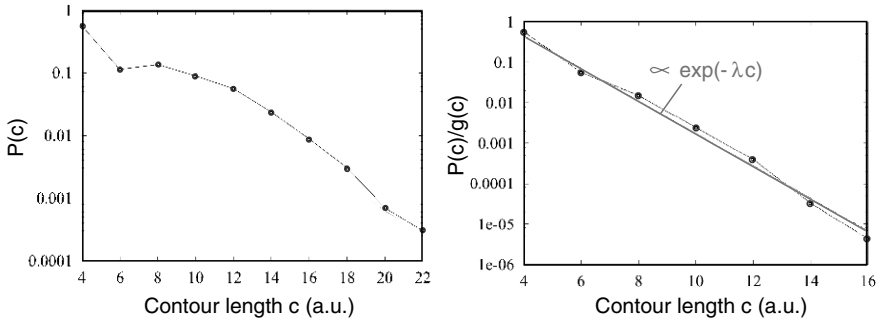


Fig. 12. Distribution of pore contours in a semi-logarithmic plot. *Left*: Raw data, *Right*: Divided by the degeneracy function $g(c)$. After [66]

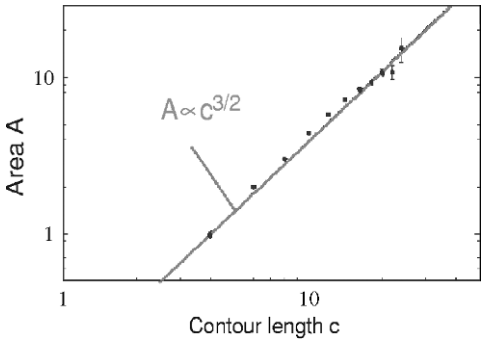


Fig. 13. Pore area vs. contour length (arbitrary units). After [66]

$R_g \propto N^{3/4}$ with the polymer length N. In our case, the "polymer length" is the contour length c. Thus the area A of a pore should scale like

$$A \propto R_g^2 \propto (C^{3/4})^2 = C^{3/2} \; . \tag{5}$$

Figure 13 shows that this is indeed the case.

Other properties of pores have been investigated, e.g., dynamical properties, pore lifetimes, pore correlations etc. [66]. All results were in good agreement with the line tension model (3). In sum, we find that the fluctuations and (pore) defects in membrane stacks can be described very well by a combination of two simple mesoscopic theories: An effective interface model for membrane undulations (1), and a line tension model for the pores (3) [67].

The simulations of Loison et al. thus demonstrate nicely that simulations of coarse-grained molecular models can be used to test mesoscopic theories, and to bridge between the microscopic and the mesoscopic level.

2.3 Mesoscopic Membrane Models

When it comes to large scale structures and long time dynamics, simulations of molecular models become too cumbersome. It then proves useful to drop the notion of particles altogether and work directly with continuous mesoscopic theories. In general, continuum theories are mostly formulated in terms of continuous fields, which stand for local, coarse-grained averages of some appropriate microscopic quantities. In the case of membranes systems, it is often more adequate to keep the membranes as separate objects, despite their microscopic thickness, and treat them as interfaces in space. This leads to random-interface theories.

Random Interface Models

In random-interface models, membranes are represented by two-dimensional sheets in space, whose conformations are distributed according to an effective

interface free energy functional. An example of such a functional has already been given in (1). However, the free energy (1) can only be used to describe slightly fluctuating planar membranes in the (xy)-plane. In order to treat membranes of arbitrary shape (vesicles etc.), one must resort to the formalism of differential geometry [12].

One of the simplest approaches of this kind is the spontaneous curvature model. The elastic free energy of a membrane is given by [70]

$$\mathcal{F}_{\text{el}} = \int dA \left[\frac{\kappa}{2}(2H - H_0)^2 + \bar{\kappa}K),\right] \tag{6}$$

where $\int dS \cdot$ integrates over the surface of the membrane, and H and K are the mean and Gaussian curvature, which are related to the local curvature radii R_1 and R_2 via $H = (1/R_1 + 1/R_2)/2$, $K = 1/(R_1 R_2)$ [69]. The parameters κ, $\bar{\kappa}$ (the bending stiffnesses) and H_0 (the spontaneous curvature) depend on the specific material properties of the membrane. The total membrane surface $A = \int dS$ is taken to be constant.

In the case $H_0 = 0$ and $\kappa > 0, \bar{\kappa} > 0$, the free energy (6) describes membranes that want to be flat, and imposes a bending penalty on all local curvatures. $H_0 \neq 0$ implies that the membrane has a favorite mean curvature. It is worth noting that the integral over the Gaussian curvature, $\int K \, dS$, depends solely on the topology of the interfaces. For membranes without rims, the Gauss-Bonnet theorem states that

$$\int dS \, K = 2\pi\chi_E \, , \tag{7}$$

where the Euler characteristics $\chi_E = 2(c - g)$ counts the number of closed surfaces c (including cavities) minus the number of handles g.

Other free energy expressions have been proposed, which allow for an asymmetric distribution of lipids between both sides of the membrane, and/or for fluctuations of A [71, 72]. Here, we shall only discuss the spontaneous curvature model. Our membranes shall be self-avoiding, i.e., they have excluded volume interactions and cannot cross one another.

Our task is to formulate a simulation model that mimicks the continuous space theory as closely as possible. More precisely, we first need a method that samples random self-avoiding surfaces, and second a way to discretize the free energy functional (6) on those surfaces.

A widely used method to generate self-avoiding surfaces is the dynamic triangulation algorithm [73–78]. Surfaces are modeled by triangular networks of spherical beads, which are linked by tethers of some given maximum length l_0 (tethered-bead model). The tethers define the neighbor relations on the surface. In order to enforce self-avoidance, the beads are equipped with hard cores (diameter σ), and the maximum tether length is chosen in the range $\sigma < l_0 < \sqrt{3}\sigma$. The surfaces are sampled with Monte Carlo, with two basic types of moves: Moves that change the *position* of beads, and moves that

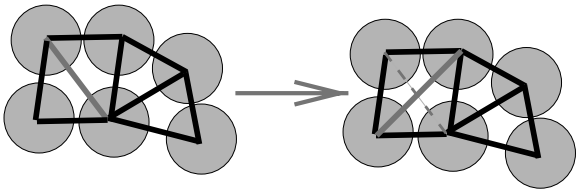


Fig. 14. Connectivity update in the dynamic triangulation algorithm

change the *connectivity* of the network, i.e., redistribute the tethers between the chains. The latter is done by randomly cutting tethers and flipping them between two possible diagonals of neighboring triangles (Fig. 14). Updates of bead positions are subject to the constraint that the maximum tether length must not exceed l_0, and connectivity updates have to comply in addition with the requirement, that at least three tethers are attached to one bead, and that two beads cannot be connected to one another by more than one tether. Otherwise, attempted updates are accepted or rejected according to the standard Metropolis prescription. Of course, additional, more sophisticated Monte Carlo moves can be introduced as well.

Next, the free energy functional (6) must be adapted for the triangulated surface. If no rims are present, we only need an expression for the first term. Since the integral over the Gaussian curvature K is then essentially given by the Euler-characteristic, it is much more convenient and accurate to evaluate the latter directly (i.e., count closed surfaces minus handles) than to attempt a discretization of K.

The question how to discretize the bending free energy in an optimal way is highly non-trivial, and several approximations have been proposed over the years [78]. Here we give a relatively recent expression due to Kumar et al. in 2001 [79]: The surface integral $\int dS$ is replaced by a sum over all triangles t with area A_t. For each triangle t, one considers the three neighbor triangles t' and determines the length $L_{tt'}$ of the common side as well as the angle $\theta_{tt'}$ between the two triangle normal vectors. The free energy is then approximated by

$$\mathcal{F}_{\text{el}} = \frac{\kappa}{2} \int dA(2H - H_0)^2 \approx \frac{\kappa}{2} \sum_t A_t \left(\frac{\sum_{t'} L_{tt'} \theta_{tt'}}{2A_t} - H_0 \right)^2 . \qquad (8)$$

This completes the tethered-bead representation for random interfaces. We have presented one of the most simple variants. The models can be extended in various directions, e.g., to represent mixed membranes [80], membranes that undergo liquid/gel transitions [81], membranes with fluctuating geometry [82] or pores [83] etc. They have even been used to study vesicle dynamics, e.g., vesicles in shear flow [84, 85]. The dynamics can be made more realistic by using hybrid algorithms, where the beads are moved according to molecular dynamics or Brownian dynamics, and the tethers are flipped with Monte Carlo. Unfortunately, the tether flips cannot easily be integrated in a pure

molecular dynamics simulation. This is a slight drawback of the tethered-bead approach.

An interesting alternative has recently been proposed by Noguchi et al. [86]. He developed a specific type of potential, which forces (single) beads to self-assemble into a two-dimensional membrane without being connected by tethers. The membrane is coarse-grained as a curved surface as in tethered-bead models, but tethers are no longer necessary. This opens the way for a fully consistent treatment of dynamics in membrane systems with frequent topology changes.

Application Example: Passage of Vesicles Through Pores

As an example for an application of a tethered-bead model, we shall discuss a problem from the physics of vesicles. Linke, Lipowsky and Gruhn have recently investigated the question, whether a vesicle can be forced to cross a narrow pore by osmotic pressure gradients [87,88]. Vesicle passage through pores plays a key role in many strategies for drug delivery. On passing through the pore, the vesicle undergoes a conformational change, which is expensive and creates a barrier. On the other hand, the vesicle can reduce its energy, if it crosses into a region with lower osmotic pressure. Linke et al. studied the question, whether the barrier height can be reduced by a sufficient amount that the pore is crossed spontaneously.

To this end, they considered a vesicle filled with N osmotically active particles. The concentration of these particles outside the vesicle is given by $c_{1,2}$ on the two sides of the pore. For a vesicle in the process of crossing the pore from side 1 to side 2, the vesicle area on both sides of the membrane is denoted $A_{1,2}$, and the volumes are $V_{1,2}$. Hence the osmotic contribution to the free energy is

$$\mathcal{F}_{\text{osm}} = (-N \ln((V_1 + V_2)/V_0) + c_1 V_1 + c_2 V_2) \,, \tag{9}$$

where V_0 is some reference volume. In the Monte Carlo simulations, a series of different crossing stages was sampled (see Fig. 15). This was done by

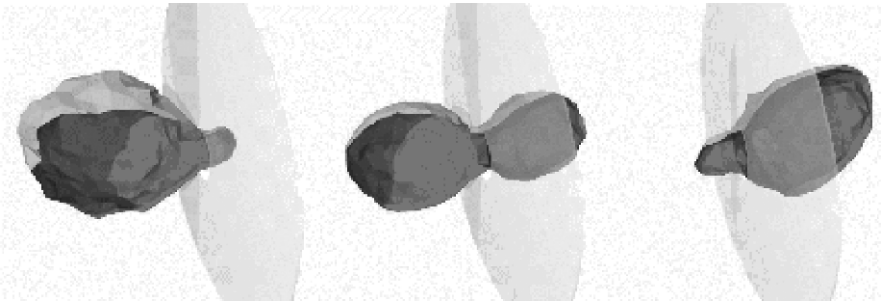


Fig. 15. Osmotically driven vesicle translocation through a pore. Configuration snapshots. From G. T. Linke [87]

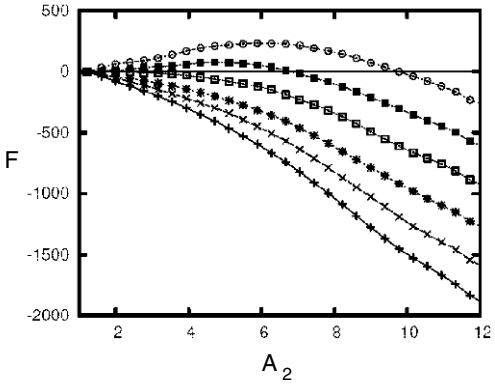


Fig. 16. Free energy landscape on crossing the pore for different osmolarities c_1 starting from $c_1 = 2700$ (*open circles*) to $c_1 = 3100$ (*plus*) in steps of 100. All units are arbitrary. The osmolarity c_2 is kept fixed. See text for more explanation. From G. Linke et al. [87, 88]

introducing an appropriate reweighting factor, which kept the area A_2 within a certain, pre-determined range. Several simulations were done with different windows for A_2. The procedure has similarity with umbrella sampling, except that the windows did not overlap. Nevertheless, the extrapolation of the histograms allowed to evaluate the free energy quite accurately, as a function of A_2.

The resulting free energy landscape is shown in Fig. 16, for different osmolarities c_1 at the side where the vesicle starts. As expected, one finds a free energy barrier for low c_1. On increasing c_1, the height of the barrier decreases, until it finally disappears. Hence the vesicle can be forced to cross the pore spontaneously, if c_1 is chosen sufficiently high. In contrast, changing the bending rigidity has hardly any effect on the height of the potential barrier.

This application shows how mesoscopic models of membrane systems can be used to study the physical basis of a potentially technologically relevant process. A comparable study with a particle-based model would have been much more expensive, and both the setup and the analysis would have been more difficult: One has no well-defined interface, one has only indirect information on the bending stiffness etc. The simplifications in the model clearly help not only to get results more quickly, but also to understand them more systematically.

2.4 Summary

In this section, we have introduced and discussed two important classes of coarse-grained models for membrane systems: Bead-spring models for simulations on the coarse-grained molecular level, and random interface models for the coarse-grained mesoscopic level. We have shown that such studies can give

insight into generic properties of membranes and into basic physical mechanisms on different scales. We have also shown that it is possible to bridge between different levels, and that simulations of particle-based models can be used to test mesoscopic theories.

3 Liquid Crystals and Surfactant Layers Under Shear

So far, we have talked about equilibrium simulations. The special properties of complex fluids lead to particularly peculiar behavior at nonequilibrium. In this section, we will discuss coarse-grained simulations of complex fluids under shear. As in the previous section, we will move from the coarse-grained molecular level to the mesoscopic level. However, we will now focus on the simulation methods rather than on the construction of simulation models. As example systems, we will consider surfactant layers and liquid crystals.

3.1 Introduction

Strain Rate and Shear Stress

We begin with recalling some basic fluid mechanics [89]. In the continuum description, the fluid is thought to be divided into fluid elements, which are big enough that microscopic details are washed out, but small enough that each element can still be considered to be homogoneous. On a macroscopic scale, the fluid is described by spatially varying fields, e.g., the velocity or flow field $\boldsymbol{u}(\boldsymbol{r}, t)$. The velocity gradient in the direction perpendicular to the flow is called the local strain rate.

One of the central quantities in fluid mechanics is the stress tensor σ. It describes the forces that the surrounding fluid exerts on the surfaces of a fluid element: For a surface with orientation $\boldsymbol{N}$ ($|\boldsymbol{N}| = 1$), the force per area is $F_i/A = \sum_j \sigma_{ij} N_j$. The diagonal components of σ give the longitudinal stress and correspond to normal forces. The off-diagonal components give the shear stress and correspond to tangential forces.

A fluid element with the velocity $\boldsymbol{u}$ is thus accelerated according to

$$\rho \frac{d}{dt} u_j = \sum_i \partial_i \sigma_{ij} \ , \tag{10}$$

where ρ is the local mass density. The stress tensor σ_{ij} is often divided in three parts.

$$\sigma_{ij} = -p\, \delta_{ij} + \sigma_{ij}^{\text{elastic}} + \sigma'_{ij} \tag{11}$$

The first term describes the effect of the local pressure. It is always present, even in an equilibrium fluid at rest. The second term, $\sigma_{ij}^{\text{elastic}}$, accounts for elastic restoring forces, if applicable. The last term, σ'_{ij}, gives the contribution of viscous forces and is commonly called the viscous stress tensor. It vanishes

in the absence of velocity gradients, i.e., if $\partial_i v_j \equiv 0$. The concrete relation between the stress tensor and the local fields is called constitutive equation and characterizes the specific fluid material.

In standard hydrodynamics, one considers so-called Newtonian fluids, which have no elasticity ($\sigma_{ij}^{\text{elastic}} \equiv 0$) and a linear, instantaneous relation between the viscous stress tensor and the velocity gradient,

$$\sigma'_{ij} = \sum_{kl} L_{ijkl}\, \partial_k u_l \ . \tag{12}$$

In the case of an incompressible, isotropic, Newtonian fluid, one has only one independent material parameter, the kinematic viscosity η, and the constitutive equation reads

$$\sigma_{ij} = -p\, \delta_{ij} + \eta(\partial_i u_j + \partial_j u_i) \ , \tag{13}$$

where the value of the local pressure $p(\boldsymbol{r})$ is determined by the requirement that the incompressibility condition

$$\nabla \boldsymbol{u} \equiv 0 \ . \tag{14}$$

is always fulfilled. Equations (10) in combination with (13) are the famous Navier-stokes equations (without external forces).

Complex fluids have internal degrees of freedom, and large characteristic length and time scales. Therefore, they become non-Newtonian already at low shear stress, and standard hydrodynamics does not apply. The relation between the stress tensor and the velocity gradient is commonly not linear, sometimes one has no unique relation at all. If the characteristic time scales are large, one encounters memory effects, e.g., the viscosity may decrease with time under shear stress [2, 4, 89].

These phenomena are often considered in planar Couette geometry: The flow $\boldsymbol{u}$ points in the x-direction, a velocity gradient in the y-direction is imposed, and the relation between the strain rate $\dot{\gamma} = \partial u_x/\partial y$ and the shear stress σ_{xy} is investigated. For Newtonian fluids, the two are proportional. In complex fluids, the differential viscosity $\partial\sigma_{xy}/\partial\dot{\gamma}$ may decrease with increasing strain rate (shear thinning), or increase (shear thickening). In some materials, the stress-strain flow curve may even be nonmonotonic. Since such flow curves are mechanically unstable in macroscopic systems, this leads to phase separation, and the system becomes inhomogeneous (Fig. 17) [9, 21, 90, 93–99, 103, 104]. The separation of a fluid into two states with distinctly different strain rates is called shear banding. As in equilibrium phase transitions, the signature of the coexistence region is a plateau region in the resulting stress-strain flow curve of the inhomogeneous systems. However, one has two important differences to the equilibrium case: First, the location of the plateau cannot be determined by a Maxwell construction, but has to be calculated by solving the full dynamical equations. Second, the plateau is

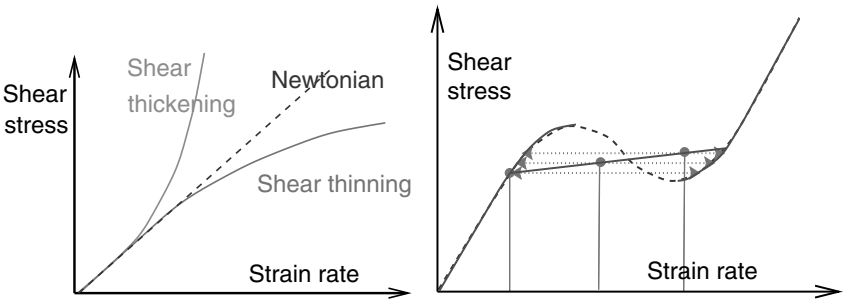


Fig. 17. Sketches of nonlinear stress-strain flow curves *Left*: Shear thickening and shear thinning; *Right*: Nonmonotonic stress-strain relation leading to phase separation. The horizontal dotted lines connect the two coexisting phases for three selected global strain rates

not necessarily horizontal. Even though mechanical stability requires that two coexisting phases must have the same shear stress, the "plateau" may have a slope or even be curved because different coexisting phases may be selected for different shear rates [100, 105] (see Fig. 17).

Nematic Liquid Crystals

Next we recapitulate some basic facts on liquid crystals. The building blocks of these materials are anisotropic particles, e.g., elongated molecules, associated structures such as wormlike micelles, or even rodlike virusses. A liquid crystal phase is an intermediate state between isotropic liquid and crystalline solid, which is anisotropic (particles are orientationally ordered), but in some sense still fluid (particles are spatially disordered in at least one direction). Some such structures are sketched in Fig. 18. The phase transitions are in some cases triggered by the temperature (thermotropic liquid crystals), and in other cases by the concentration of the anisotropic particles (lyotropic liquid crystals). Lamellar membrane stacks (see previous section) are examples of smectic liquid crystalline structures. In this section, we shall mainly be concerned with nematics, where the particles are aligned along one common direction, but otherwise fluid.

For symmetry reasons, the transition between the isotropic fluid and the nematic fluid must be first order. The nematic order is commonly characterized by the symmetric traceless order tensor

$$Q_{ij} = \left\langle \frac{1}{2}(3d_i d_j - 1) \right\rangle \ , \tag{15}$$

where $\boldsymbol{d}$ is a unit vector characterizing the orientation of a particle. In an isotropic fluid, $\mathbf{Q}$ is zero. Otherwise, the largest eigenvalue gives the nematic

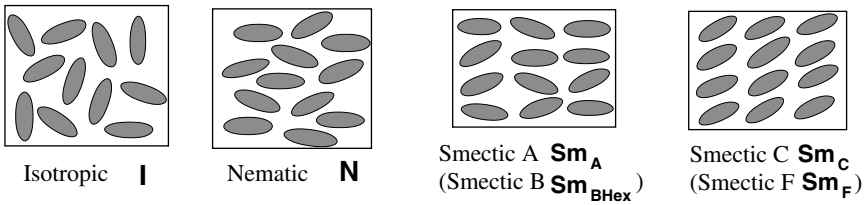


Fig. 18. Examples of liquid crystalline phases. In the nematic phase (N), the particles are oriented, but translationally disordered. In the smectic phases, they form layers in one direction, but remain fluid within the layers. In the smectic B_{Hex} and smectic F phase, the structure within the layers is not entirely disordered, but has a type of order called hexatic. See [14, 15] for explanation. The isotropic phase (I) is entirely disordered (regular isotropic liquid)

order parameter S, and the corresponding eigenvector is the direction of preferred alignment, the "director" $\boldsymbol{n}$. Note that only the direction of $\boldsymbol{n}$ matters, not the orientation, i.e., the vectors $\boldsymbol{n}$ and $-\boldsymbol{n}$ are equivalent.

Since the ordered nematic state is degenerate with respect to a continuous symmetry (the direction of the director), it exhibits elasticity: The free energy costs of spatial director variations must vanish if the wavelength of the modulations tends to infinity. At finite wavelength, the system responds to such distortions with elastic restoring forces. Symmetry arguments show that these depend only on three independent material constants [16], the Frank constants K_{11}, K_{22}, and K_{33}. The elastic free energy is given by [14]

$$\mathcal{F}_{\text{elastic}} = \int d\boldsymbol{r} \left\{ K_{11}(\nabla \boldsymbol{n})^2 + K_{22}(\boldsymbol{n} \cdot (\nabla \times \boldsymbol{n}))^2 + K_{33}(\boldsymbol{n} \times (\nabla \times \boldsymbol{n}))^2 \right\} . \quad (16)$$

Figure 19 illustrates the corresponding fundamental distortions, the splay (K_{11}), twist (K_{22}), and bend mode (K_{33}).

How does shear affect a liquid crystalline fluid? In the isotropic phase and at low strain rate , the situation is still comparably simple. The velocity gradient imposes a background angular velocity

$$\boldsymbol{\Omega} = \frac{1}{2}(\nabla \times \boldsymbol{u}) \quad (17)$$

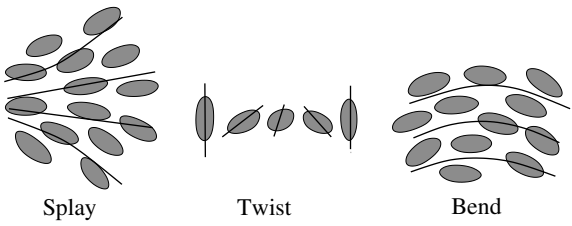


Fig. 19. Elastic modes in nematic liquid crystals

on the fluid, which the particles pick up [106]. Since they rotate faster while perpendicular to the flow, they also acquire a very slight effective orientation at the angle $\pi/4$ relative to the flow direction. At higher strain rate, the order parameter increases, and the alignment angle decreases [107]. This sometimes even leads to a shear-induced transition into the nematic phase [108–114].

Deep in the nematic phase, the nematic fluid can be described by the flow field $\boldsymbol{u}(\boldsymbol{r},t)$ and the director field $\boldsymbol{n}(\boldsymbol{r},t)$ ($|\boldsymbol{n}| \equiv 1$). We shall now assume incompressibility ($\nabla \boldsymbol{u} = 0$) and a linear, instantaneous relation stress-strain relation as in (12). Compared to simple fluids, there are several complications.

First, nematic fluids are elastic, hence the stress tensor σ_{ij} (11) has an elastic contribution

$$\sigma_{ij}^{\text{elastic}} = -\frac{\delta \mathcal{F}_{\text{elastic}}}{\delta(\partial_i n_k)}\, \partial_j n_k \,. \tag{18}$$

The free energy $\mathcal{F}_{\text{elastic}}$ is given by (16).

Second, the fluid is locally anisotropic, therefore the constitutive equation for the viscous stress tensor contains more independent terms than in simple fluids. The structure of σ'_{ij} which is compatible with all symmetry requirements reads [15]

$$\begin{aligned}\sigma'_{ij} = \alpha_1 \sum_{kl} n_i n_j n_k n_l A_{kl} + \alpha_2\, n_i N_j + \alpha_3\, N_i n_j \\ + \alpha_4\, A_{ij} + \alpha_5 \sum_k n_i n_k A_{jk} + \alpha_6 \sum_k n_j n_k A_{ik},\end{aligned} \tag{19}$$

where $\mathbf{A}$ is the symmetric flow deformation tensor, $A_{ij} = \frac{1}{2}(\partial_i u_j + \partial_j u_i)$, and the vector $\boldsymbol{N}$ gives the rate of change of the director relative to the angular velocity of the fluid, $\boldsymbol{N} = \frac{d\boldsymbol{n}}{dt} - \boldsymbol{\Omega} \times \boldsymbol{n}$. The parameters α_i are the so-called Leslie coefficients. They are linked by one relation [15],

$$\alpha_2 + \alpha_3 = \alpha_6 - \alpha_5 \tag{20}$$

hence one has five independent material constants.

Third, the flow field and the director field are coupled. Thus one must take into account the dynamical evolution of the director field,

$$\rho_1 \ddot{n}_j = g_j + \partial_i \pi_{ij} \,, \tag{21}$$

which is driven by the intrinsic body force $\boldsymbol{g}$, and the director stress tensor π_{ij}. Here ρ_1 is the moment of inertia per unit volume. As for the stress tensor, one can derive the general form of the constitutive equations for $\boldsymbol{g}$ and π with symmetry arguments. The individual expressions can be found in the literature [15]. Here, we only give the final total dynamical equation for $\boldsymbol{n}$,

$$\rho_1 \ddot{n}_j = \lambda n_j + \sum_i \partial_i \left(\frac{\delta \mathcal{F}_{\text{el}}}{\delta(\partial_i n_j)} \right) - \frac{\partial \mathcal{F}_{\text{el}}}{\partial n_j} - \gamma_1 N_j - \gamma_2 \sum_i n_i A_{ij} \,, \tag{22}$$

where $\gamma_1 = \alpha_3 - \alpha_2$, $\gamma_2 = \alpha_6 - \alpha_5$, and λ is a Lagrange parameter which keeps $|\boldsymbol{n}|$ constant.

To summarize, the dynamical equations of nematohydrodynamics for incompressible nematic fluids are given by (10) and (22) together with (11), (16), (18), and (19). These are the Leslie-Ericksen equations of nematohydrodynamics [15]. The theory depends on five independent viscosity coefficients α_i and on three independent elastic constants K_{ii}. Not surprisingly, it predicts a rich spectrum of phenomena. For example, in the absence of director gradients, steady shear flow is found to have two opposing effects on the director: It rotates the molecules, and it aligns them. The relative strength of these two tendencies depends on the material parameters. As a result, one may either encounter stable flow alignment or a rotating "tumbling" state [115–118].

The Ericksen-Leslie theory describes incompressible and fully ordered nematic fluids. The basic version presented here does not account for the possibility of disclination lines and other topological defects, and it does not include the coupling with a variable density and/or order parameter field. The situation is even more complicated in the vicinity of stable or metastable (hidden) thermodynamic phase transitions. Thermodynamic forces may contribute to a nonmonotonic stress-strain relationship, which eventually lead to mechanical phase separation [98].

Figure 20 shows an experimental example of a nonequilibrium phase diagram for a system of rodlike virusses [102].

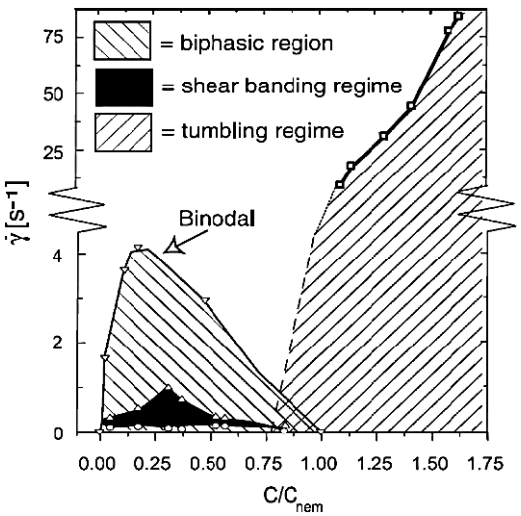


Fig. 20. Nonequilibrium phase diagram of a lyotropic liquid crystal under shear as a function of strain rate $\dot{\gamma}$ and particle density C. ($C_{\rm nem}$ is the density of the equilibrium isotropic/nematic transition). The system is a mixture of rodlike virusses (fd virus) and polymers (dextran). From [102]

3.2 Simulating Shear on the Particle Level: NEMD

In the linear response regime, i.e., in the low-shear limit, many rheological properties of materials can be determined from equilibrium simulations. For example, the viscosity coefficients introduced in the previous section can be related to equilibrium fluctuations with Green-Kubo relations [119, 120].

However, the non-Newtonian character of most complex fluids and their resulting unique properties manifest themselves only beyond the linear response regime. To study these, non-equilibrium simulations are necessary. We will now briefly review non-equilibrium molecular dynamics (NEMD) methods for the simulation of molecular model systems under shear. Some of the methods are covered in much more detail in [121–123].

In NEMD simulations, one faces two challenges. First, one must mechanically impose shear. Second, most methods that enforce shear constantly pump energy into the system. Hence one must get rid of that heat, i.e., apply an appropriate thermostat. We will address these two issues separately.

The most direct way of imposing shear is to confine the system between two rough walls, and either move one of them (for Couette flow), or apply a pressure gradient (for Poiseuille flow). Varnik and Binder have shown that this approach can be used to measure the shear viscosity in polymer melts [124]. It has the merit of being physical: Strain is enforced physically, and the heat can be removed in a physical way by coupling a thermostat to the walls. On the other hand, the system contains two surfaces, and depending on the material under consideration, one may encounter strong surface effects. This can cause problems.

An alternative straightforward way to generate planar Couette flow is to use moving periodic boundary conditions as illustrated in Fig. 21 (Lees-Edwards boundary conditions [121, 125]): In order to enforce shear flow $\boldsymbol{u} = (u_x, 0, 0)$ with an average strain rate $\dot{\gamma} = \partial u_x / \partial y$, one proceeds as follows: One replicates the particles in the x and z direction like in regular periodic boundary conditions,

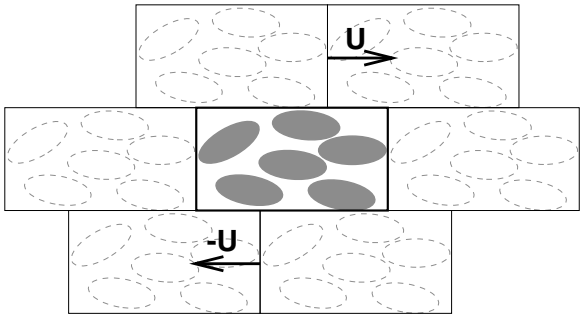


Fig. 21. Lees-Edwards boundary conditions

$$\begin{array}{ll} r_x \to r_x \pm L_x & r_{x,y} \to r_{x,y} \\ r_{y,z} \to r_{y,z} & r_z \to r_z \pm L_z \ . \\ \boldsymbol{v} \to \boldsymbol{v} & \boldsymbol{v} \to \boldsymbol{v} \end{array} \tag{23}$$

In the y-direction, the replicated particles acquire an additional velocity $U = \dot{\gamma} L_y$ and an offset Ut in the x-direction,

$$\begin{array}{l} r_x \to r_x \pm (Ut \,\mathrm{mod} L_x) \\ r_y \to r_y \pm L_y \\ r_z \to r_z \\ \boldsymbol{v} \to \boldsymbol{v} \pm U \boldsymbol{e}_x \end{array} \tag{24}$$

Here $L_{x,y,z}$ are the dimensions of the simulation box, and $\boldsymbol{e}_x$ the unit vector in the x-direction.

Lees-Edwards boundary conditions are perfectly sufficient to enforce planar Couette flow. In some applications, it has nevertheless proven useful to supplement them with fictitious bulk forces that favor a linear flow profile. One particularly popular algorithm of this kind is the SLLOD algorithm [122,126]. For the imposed flow $\boldsymbol{u}(\boldsymbol{r}) = \dot{\gamma} y \boldsymbol{e}_x$, the SLLOD equations of motion for particles i read

$$\frac{d\boldsymbol{r}_i}{dt} = \frac{\boldsymbol{p}'_i}{m_i} + \dot{\gamma} y_i \boldsymbol{e}_x \tag{25}$$

$$\frac{d\boldsymbol{p}'_i}{dt} = \boldsymbol{F}_i - \dot{\gamma} p'_{yi} \boldsymbol{e}_x \ , \tag{26}$$

where $\boldsymbol{p}'_i = m_i(\boldsymbol{v}_i - \boldsymbol{u}(\boldsymbol{r_i}))$ is the momentum of particle i in a reference frame moving with the local flow velocity $\boldsymbol{u}(\boldsymbol{r}_i)$, and $\boldsymbol{F}_i$ is the regular force acting on the particle i. The equations of motion can be integrated with standard techniques, or with specially devised operator-split algorithms [128,129]. It is also possible to formulate SLLOD variants for arbitrary steady flow [127]. The SLLOD algorithm is very useful for the determination of viscosities in complex fluids [119,130–134]. It has the advantage that the system relaxes faster towards a steady state, and that the analysis of the data is simplified [122].

A fourth method to enforce shear has recently been proposed by Müller-Plathe [135]. One uses regular periodic boundary conditions, and divides the system into slabs in the desired direction of flow gradient. In periodic intervals, the particle in the central slab ($y = 0$) with the largest velocity component in the x direction, and the particle in the topmost slab ($y = L_y/2$) with the largest velocity component in the $(-x)$-direction is determined, and the momentum components p_x of the two particles are swapped. This leads to a zigzag-shaped shear profile [135].

The Müller-Plathe algorithm can be implemented very easily, see Fig. 22 Just one small change turns a molecular dynamics program for thermal equilibrium into a NEMD program that produces shear flow. Furthermore, the algorithm conserves momentum and energy: Contrary to the algorithms discussed before, it does not pump heat into the system. Instead, the heat is constantly redistributed. It is drained out of the system at $y = 0$ and $y = L_y/2$ by

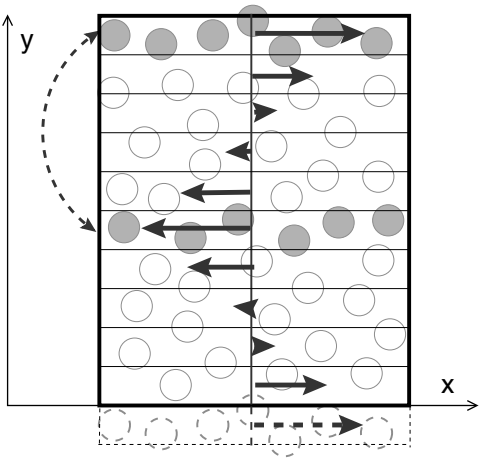


Fig. 22. Schematic sketch of the Müller-Plathe method for imposing shear. In periodical intervals, the particle in the middle slab that moves fastest in the x-direction and the particle in the top slab that moves fastest in the $(-x)$-direction exchange their momentum components p_x. After [135]

the momentum swaps, but it is also produced in the bulk through the energy dissipation associated with the shear flow. As a result, the zigzag shear profile is accompanied with a W-shaped temperature profile. Since this is not always wanted, the algorithm is sometimes used in combination with a thermostat. A slight drawback of the algorithm is the fact that it produces intrinsically inhomogeneous profiles, due to the presence of the two special slabs at $y = 0$ and $y = L_y/2$. This may cause problems in small systems.

Having introduced these four methods to impose shear, we now turn to the problem of thermostatting. All shearing methods except for the Müller-Plathe method produce heat, therefore the thermostat is an essential part of the simulation algorithm.

A thermostat is a piece of algorithm that manipulates the particle velocities such that the temperature is adjusted to its desired value. This is commonly done by adjusting the kinetic energy. In a flowing fluid with flow velocity $\boldsymbol{u}(\boldsymbol{r})$, the kinetic energy has a flow contribution and a thermal contribution. Therefore, it is important to define the thermostat in a frame that moves with the fluid, i.e., to couple it to the "peculiar" velocities $\boldsymbol{v}'_i = \boldsymbol{v}_i - \boldsymbol{u}(\boldsymbol{r}_i)$, rather than to the absolute velocities $\boldsymbol{v}_i$. Unfortunately, most thermostats don't account for this automatically, and one has to put in the flow profile by hand. The only exception is the dissipative particle dynamics (DPD) thermostat (see below).

This raises the question how to determine the flow $\boldsymbol{u}(\boldsymbol{r})$. In a homogeneous system, at low strain rates, it can be acceptable to use the pre-imposed profile ($\boldsymbol{u}(\boldsymbol{r}) = \dot{\gamma} y \boldsymbol{e}_x$ in our earlier examples). Thermostats that use the pre-imposed

flow profile are called "profile-biased thermostats". At high strain rates, and in inhomogeneous systems, the use of profile-biased thermostats is dangerous and may twist the results [122, 133]. In that case, the profiles $\boldsymbol{u}(\boldsymbol{r})$ must be calculated directly from the simulations, "on the fly". Such thermostats are called "profile-unbiased thermostats".

One can distinguish between deterministic and stochastic thermostatting methods. The simplest deterministic thermostat simply rescales the peculiar velocities of all particles after every time step such that the total thermal kinetic energy remains constant. Among the more sophisticated deterministic thermostats, the ones that have been used most widely in NEMD simulations are the so-called Gaussian thermostats [122], and the Nosé Hoover thermostats [121, 136].

The idea of Gaussian thermostats is very simple: For a system of particles j subject to forces $\boldsymbol{f}_j$, one minimizes the quantity $C = \sum_j (\dot{\boldsymbol{p}}'_j - \boldsymbol{f}_j)^2/(2m_j)$ with respect to $\dot{\boldsymbol{p}}_j$, with the constraint that either the time derivative of the total energy, or the time derivative of the total kinetic energy be zero (Gaussian isoenergetic and Gaussian isothermal thermostat, respectively). In the absence of constraints, the minimization procedure would yield Newton's law. The constraints modify the equations of motion in a way that can be considered in some sense minimal. One obtains

$$\dot{\boldsymbol{p}}'_j = \boldsymbol{f}_j - \alpha \boldsymbol{p}_j \ , \tag{27}$$

where α is a Lagrange multiplier, which has to be chosen such that the desired constraint – constant energy or constant kinetic energy – is fulfilled.

The Nosé Hoover thermostats are well-known from equilibrium molecular dynamics and shall not be discussed in detail here. We only note that in NEMD simulations, one uses not only the standard isothermal Nosé-Hoover thermostats, but also isoenergetic variants.

Deterministic thermostats have the slightly disturbing feature that they change the dynamical equations in a rather unphysical way. It is sometimes hard to say how that might affect the system. At equilibrium, Nosé-Hoover thermostats rest on a well-established theoretical basis: They can be derived from extended Hamiltonians, they produce the correct thermal averages for static quantities etc. Outside of equilibrium, the situation is much more vague. For the case of steady states, some results on the equivalence of thermostats are fortunately available. Evans and Sarman have shown that steady state averages and time correlation functions are identical for Gaussian isothermal and isoenergetic thermostats and for Nosé-Hoover thermostats [123, 137]. Ruelle has recently proved that the Gaussian isothermal and the Gaussian isoenergetic thermostat are equivalent in an infinite system which is ergodic under space translations [138].

An alternative approach to thermostatting are the stochastic methods, such as the Langevin thermostat [139] and the DPD thermostat [140]. They maintain the desired temperature by introducing friction and stochastic forces. This approach has the virtue of being physically motivated. The dissipative

and stochastic terms stand for microscopic degrees of freedom, which have been integrated out in the coarse-grained model, but can still absorb and carry heat. From the theory of coarse-graining, it is well-known that integrating out degrees of freedom effectively turns deterministic equations of motion into stochastic equations of motion [141–143].

The Langevin thermostat is the simplest stochastic dynamical model. One modifies the equation of motion for the peculiar momenta by introducing a dissipative friction ζ and a random force $\boldsymbol{\eta}_j$,

$$\dot{\boldsymbol{p}}'_j = \boldsymbol{f}_j - \zeta \boldsymbol{v}'_j + \boldsymbol{\eta}_j \ , \tag{28}$$

where $\boldsymbol{\eta}_j$ fulfills

$$\langle \boldsymbol{\eta}_k \rangle = 0 \tag{29}$$

$$\langle \eta_{i\alpha}(t)\eta_{j\beta}(t') \rangle = 2\zeta k_B T \, \delta_{ij} \, \delta_{\alpha\beta} \, \delta(t-t') \ , \tag{30}$$

i.e., the noise is delta-correlated and Gaussian distributed with mean zero and variance $2\zeta k_B T$ (fluctuation-dissipation theorem). In practice, this can be done by adding $(-\zeta v'_{j\alpha})$ and a random number to each force component $f_{i\alpha}$ in each time step Δt. The random number must be distributed with mean zero and variance $\sqrt{2k_B T\zeta/\Delta t}$. Because of the central limit theorem, the distribution does not necessarily have to be Gaussian, if the time step is sufficiently small [144].

The DPD thermostat is an application of the dissipative particle dynamics algorithm [140]. The idea is similar to that of the Langevin algorithm, but instead of coupling to absolute velocities, the thermostat couples to velocity differences between neighbor particles. Therefore, the algorithm is Galilean invariant, and accounts automatically for the difference between flow velocity and thermal velocity. We refer to B. Dünweg's chapter (this book) for an introduction into the DPD method (see also [136]).

This completes our brief introduction into NEMD methods for systems under shear. We close with a general comment: The strain rate $\dot{\gamma}$ has the dimension of inverse time. It introduces a time scale $1/\dot{\gamma}$ in the system, which diverges in the limit $\dot{\gamma} \to 0$. Therefore, systems at low strain rates converge very slowly towards their final steady state, and the study of systems at low strain rates is very time-consuming. In fact, even in coarse-grained molecular models, simulated strain rates are usually much higher than experimentally accessible strain rates.

We now give two examples of recent large-scale NEMD studies of inhomogeneous complex systems under shear.

First Application Example: Nematic-Isotropic Interfaces Under Shear

Our first example is a study of the behavior of a Nematic-Isotropic interface under shear (Germano and Schmid [145, 146]). Interfaces play a central role

for shear banding, and understanding their structure should provide a key to the understanding of phase separation under shear. Our study is a first step in this direction. We will discuss the questions whether the structure of an equilibrium interface is affected by shear, whether the phase transition is shifted, and whether shear banding can be observed.

The model system was a fluid of soft repulsive ellipsoids with the aspect ratio 15:1, in a simulation box of size (150:300:150) (in units of the ellipsoid diameter σ). The density was chosen in the coexistence region between the nematic and isotropic phase, such that the system phase separates at equilibrium. The initial configuration was a relaxed equilibrium configuration with a nematic slab and an isotropic slab, separated by two interfaces. The latter align the director of the nematic phase such that it is parallel to the interface [147–149].

Shear was imposed with Lees-Edwards boundary conditions in combination with a profile unbiased Nosé Hoover thermostat, in the direction normal to the interface. The two coexisting phases thus share the same shear stress. Two setups are possible within this "common stress" geometry. In the flow-aligned setup, the director points in the direction of flow $\boldsymbol{u}$ (the x direction); in the log-rolling setup, it points in the direction of "vorticity" $\nabla \times \boldsymbol{u}$ (the z direction). Only the flow-aligned setup turned out to be stable in our system. We considered mainly the strain rate $\dot{\gamma} = 0.001\tau^{-1}$. Here τ is the natural time unit, $\tau = \sigma\sqrt{m/k_BT}$, with the particle mass m, the particle diameter σ, and the temperature T. This strain rate is small enough that the interface is still stable. A configuration snapshot is shown in Fig. 23.

To analyze the system, it was split into columns of size $B \times L_y \times B$. The columns were further split into bins, which contained enough particles that a local order parameter could be determined. In this manner, one obtains local order parameter profiles for each column in each configuration. From the profile $S(y)$, we determined the local positions of the two interfaces as follows: We computed at least two coarse-grained profiles $\bar{S}(y;\delta_i)$ by convoluting the profile $S(y)$ with a symmetrical box-like coarse-graining function of width δ_i. The coarse-graining width must be chosen such that the coarse-grained profiles still reflect the interfaces, but short ranged fluctuations are averaged out. The intersection of two averaged profiles locates a "dividing surface", where the negative order parameter excess on the nematic side balances the positive order parameter excess on the paranematic side. We define this to be the "local position" of the interface. Due to fluctuations, it depends slightly on the particular choice of the δ_i. Nevertheless, the procedure gives remarkably unambiguous values even for strongly fluctuating profiles [149]. The procedure is illustrated in Fig. 24.

Once the positions h_{NI}, h_{IN} of the two interfaces have been determined, we calculate profiles for all quantities of interest and shift them by the amount h_{NI} or h_{IN}, respectively. This allows to perform averages over *local* profiles. Furthermore, the interface positions h_{IN} and h_{NI} themselves can be used to analyze the fluctuation spectrum of the interface positions.

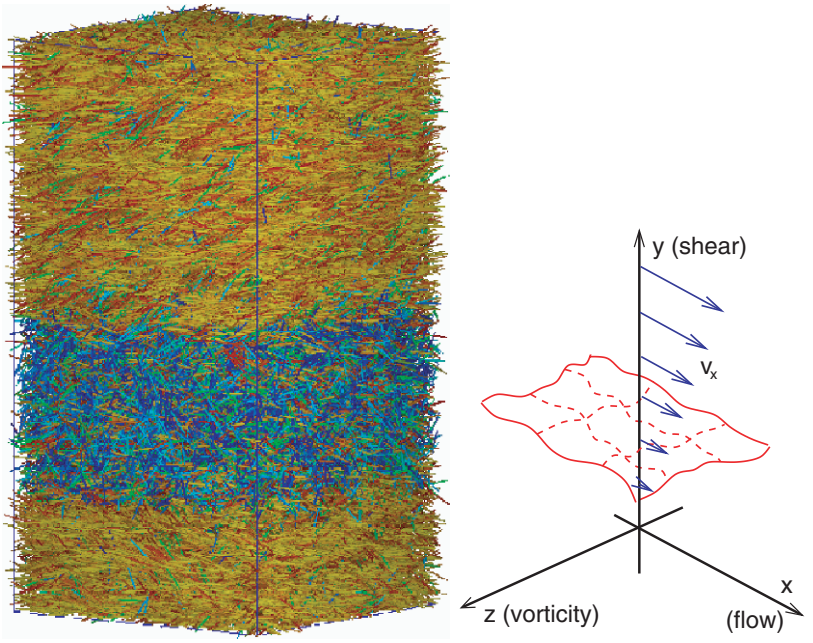


Fig. 23. Snapshot of a sheared nematic-isotropic interface (115200 particles) at strain rate $\dot{\gamma} = 0.001\tau^{-1}$. Also shown is the coordinate system used throughout this section. From [145]

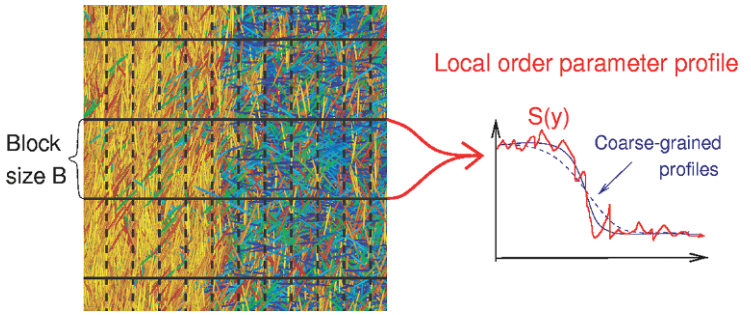


Fig. 24. Illustration of block analysis to obtain local profiles

Unexpectedly, the shear turned out to have almost no influence at all on the interfacial structure. Under shear, the interfaces are slightly broadened, indicating that the interfacial tension might be slightly reduced (Fig. 25). But the effect is barely noticeable. Likewise, the density profile and the capillary wave spectrum are almost undistinguishable from those of the equilibrium interface, within the error. The velocity profiles, however, are much more interesting. The flow profile exhibits a distinct kink at the interface position (Fig. 26). Hence we clearly observe shear banding – the total strain is

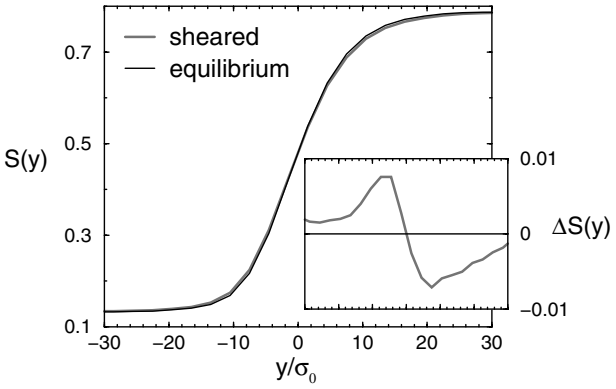


Fig. 25. Interfacial order parameter profile at strain rate $\dot{\gamma} = 0.001\tau^{-1}$ (flow aligned setup). Inset shows the difference between profiles with and without shear. From [146]

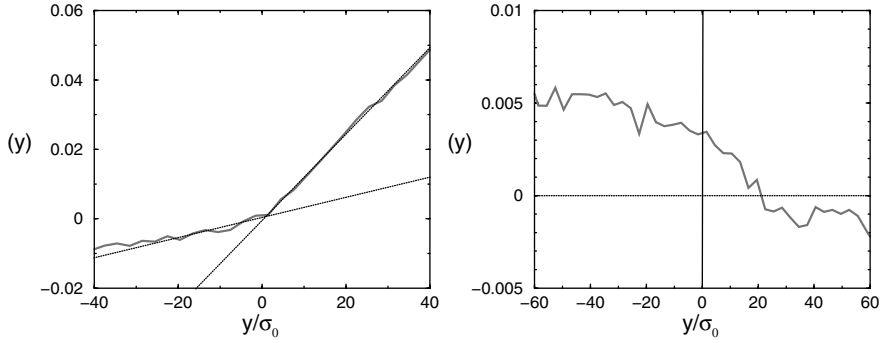


Fig. 26. Velocity profile in the flow direction (*left*), and in the vorticity direction (*right*), at strain rate $\dot{\gamma} = 0.001\tau^{-1}$. From [146]

distributed nonuniformly between the two phases. At the same shear stress, the nematic phase supports a higher strain rate than the isotropic phase. Second and surprisingly, the interface apparently also induces a streaming velocity gradient in the vorticity direction (the z-direction). As a consequence, vorticity flow is generated in opposite directions in the isotropic and in the nematic phase. This flow is symmetry breaking, and as yet, we have no explanation for it. We hope that future theoretical and simulation studies will clarify this phenomenon.

At high strain rates, the interface gets destroyed. Figure 27 shows a nonequilibrium phase diagram, which has been obtained from interface simulations of small systems. The local strain rate is always higher in the nematic phase than in the isotropic phase. It is remarkable that the coexistence region does not close up. Instead, the interface disappears abruptly at average strain rates above $\dot{\gamma} \approx 0.006/\tau$.

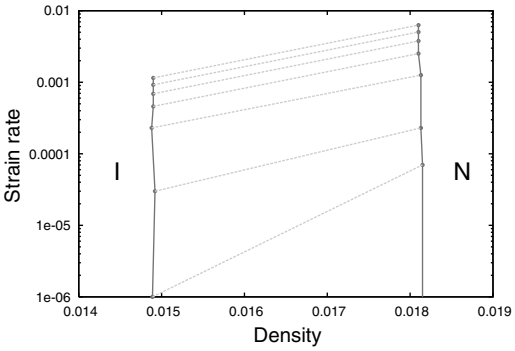


Fig. 27. Phase diagram of the flow-aligned system. The *dashed tie lines* connect coexisting states. I denotes the isotropic region, and N the nematic region. From [146]

Second Application Example: Shear-Induced Phenomena in Surfactant Layers

As a second example, we discuss the effect of shear on lamellar stacks of surfactants. This has been studied in coarse-grained molecular simulations by Soddemann, Guo, Kremer et al. [150,151]. The model system was very similar to that used in Sect. 2.2, except that it does not contain solvent particles [60]. Shear was induced with the Müller-Plathe algorithm [135] in combination with a DPD thermostat [60].

The first series of simulations by Guo et al. [150] addressed the question, whether the surfactant layers can be forced to reorient under shear. Three different orientations of the layers with respect to the shear geometry were considered: In the transverse orientation, the layers are normal to the flow direction, in the parallel orientation, they are normal to the direction of the velocity gradient, and in the perpendicular orientation, they lie in the plane of the flow and the velocity gradient. The transverse orientation is unstable under shear flow. Experimentally, both a transition from the transverse to the parallel orientation and from the transverse to the perpendicular orientation have been observed [152, 153].

Guo et al. have studied this by simulations of a system of dimer amphiphiles. Figures 28 and 29 show a series of snapshots for different times at two different strain rates. Both systems were set up in the transverse state. At low strain rate, the shear generates defects in the transverse lamellae. Mediated by these defects, the lamellae gradually reorient into the parallel state. At high strain rate, the lamellae are first completely destroyed and then reorganize in the perpendicular state.

The final state is not necessarily the most favorable steady state. In fact, Guo et al. found that the strain energy dissipation was always smallest in the perpendicular state, regardless of the strain rate. When starting from the transverse state at low strain rates, the parallel state formed nevertheless for

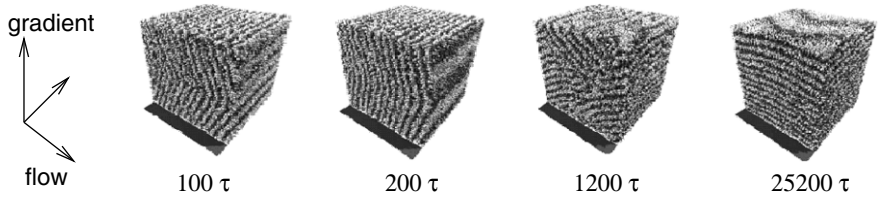


Fig. 28. Reorientation of surfactant layers under shear from an initial transverse state at low strain rate, $\dot{\gamma} = 0.001/\tau$. Reprinted with permission from [150]. Copyright 2002 by the American Physical Society

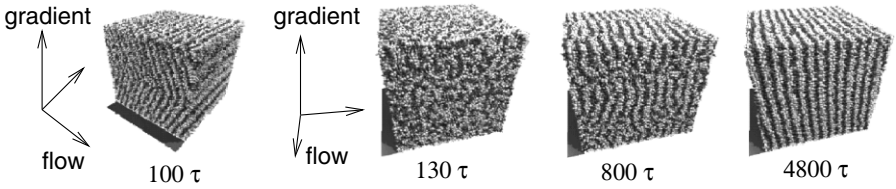


Fig. 29. Reorientation of surfactant layers under shear from an initial transverse state at high strain rate, $\dot{\gamma} = 0.03/\tau$. Reprinted with permission from [150]. Copyright 2002 by the American Physical Society

kinetic reasons. At intermediate strain rates, shear may induce undulations in the parallel state. This has been predicted by Auernhammer et al. [154, 155] and studied by Soddemann et al. by computer simulation of a system of layered tetramers [151]. The phenomenon results from the coupling between the layer normal, the tilt of the molecules, and the shear flow. It is triggered by the fact that shear flow induces tilt. The tilted surfactant layer dilates and uses more area, which eventually leads to an undulation instability. This could indeed be observed in the simulations. Figure 30 shows a snapshot of an undulated configuration, and a plot of the undulation amplitude as a function of strain rate. The undulations set in at a well-defined strain-rate.

Our two examples show that coarse-grained molecular simulations of complex, inhomogeneous fluids under shear are now becoming possible. A wealth of new, intriguing physics can be expected from this field in the future.

3.3 Simulations at the Mesoscopic Level

We will only touch very briefly on the wide field of mesoscopic simulations for complex fluids under shear.

The challenge of mesoscopic simulations is to find and/or formulate the appropriate mesoscopic model, and then put it on the computer. For liquid crystals, we have already discussed one candidate, the Leslie-Ericksen theory. However, this is by far not the only available mesoscopic theory for liquid crystals. A popular alternative for the description of lyotropic liquid crystals is

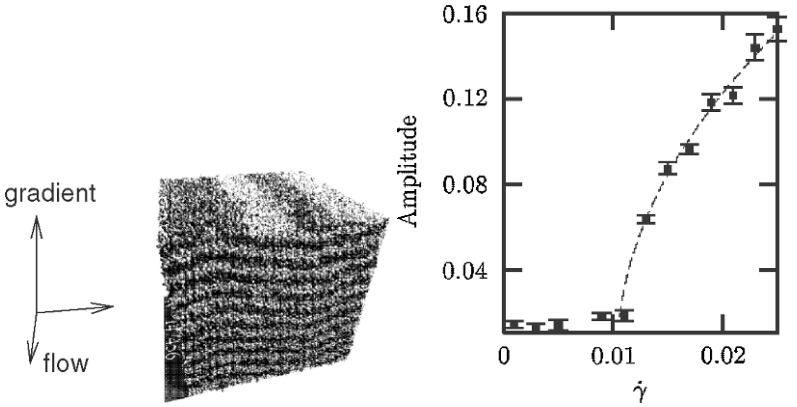


Fig. 30. Shear induced undulations. *Left*: configuration snapshot at strain rate $\dot{\gamma} = 0.015/\tau$. *Right*: Undulation amplitude as a function of strain rate. Undulations set in at a strain rate of $\dot{\gamma} \approx 0.01$. From [151]

the Doi theory, which is based on a Smoluchowski equation for the distribution function for rods. Numerous variants and extensions of the Doi model have been proposed. A relatively recent review on both the Leslie-Ericksen and the Doi theory can be found in [156].

The phenomenological equations can be solved with traditional methods of computational fluid dynamics. An alternative approach which is particularly well suited to simulate flows in complex and fluctuating geometries is the Lattice Boltzmann method, which has been discussed in B. Dünweg's chapter. Yeomans and coworkers have recently proposed a Lattice Boltzmann algorithm for liquid-crystal hydrodynamics, which allows to simulate liquid crystal flow [157,158]. The starting point is the Beris-Edwards theory [159], a dynamical model for liquid crystals that consists of coupled dynamical equations for the velocity profiles and the order tensor $\mathbf{Q}$ (15). The Lattice-Boltzmann scheme works with the usual discrete velocities i and distributions $f_i(\boldsymbol{r}_n)$ for the fluid flow field on the lattice site $\boldsymbol{r}_n$. In addition, a second distribution $G_i(\boldsymbol{r}_n)$ is introduced, which describes the flow of the order tensor field. The time evolution includes free streaming and collision steps. The exact equations are quite complicated and shall not be given here, they can be found in [158].

As an application example of this method, we discuss a recent simulation by Marenduzzo et al. of a sheared hybrid aligned nematic (HAN) cell [160,161]. The surfaces in a HAN cell impose conflicting orientations on the director of an adjacent nematic fluid, i.e., one surface orients ("anchors") the fluid in a parallel way ("planar"), and the other in a perpendicular way ("homeotropic"). In the system under consideration, the parameters were chosen such that the fluid is in the isotropic phase, but very close to the nematic phase. The fluid is at rest at time $t > 0$. At times $t > 0$, the surfaces are moved relative to each other, and drag the fluid with them (no-slip boundary conditions). Figure 31

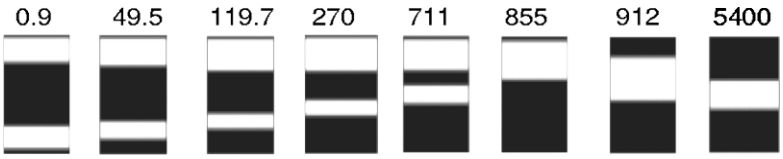


Fig. 31. Snapshots from a Lattice-Boltzmann simulation of a sheared hybrid aligned nematic cell. White regions are ordered, and black regions are disordered. The numbers give time in milliseconds. From [160]

shows snapshots of the cell for different times. The evolution of the structure inside the cell is quite complex. First, two ordered bands form close to the surface, due to the fact that the flow field still builds up and strain rates are quite high in the vicinity of the surfaces. The director in these bands is flow-aligned, and in the lower band, its direction competes with the anchoring energy of the homeotropic surface. As a result, the lower band unbinds and crosses the system to join the other band. Finally, the top band also unbinds and moves towards the center of the cell. This simulation shows nicely how a complex dynamical process, which results from an interplay of shear flow and elastic deformations, can be studied with a mesoscopic simulation method. We note that the time scale of this simulation is seconds, i.e., inaccessible for molecular simulations.

3.4 Summary

Under shear, complex fluids exhibit a wealth of new, interesting, and practically relevant phenomena. To study these, various simulation approaches for different levels of coarse-graining have been developed. In this section, we have presented and discussed several variants of non-equilibrium molecular dynamics (NEMD), and illustrated their use with examples of large scale NEMD simulations of inhomogeneous liquid crystals and surfactant layers. Furthermore, we have briefly touched on mesoscopic simulation methods for liquid crystals under shear.

4 Conclusions

Due to the rapidly improving performance of modern computers, complex fluids can be studied on a much higher level today than just ten years ago. A large number of coarse-grained models and methods have been designed, which allow to investigate different aspects of these materials, at equilibrium as well as far from equilibrium. The new computational possibilities have vitally contributed to the current boost of soft matter science.

In the two central sections of this chapter, we have given an introduction into two important chapters of computational soft matter: In Sect. 2, we have

given a rough overview over idealized coarse-grained simulation models for membranes, and hopefully conveyed an idea of the potential of such models. In Sect. 3, we have discussed simulation methods for complex fluids under shear. Generally, nonequilibrium studies of complex fluids are still much more scarce than equilibrium studies. Much remains to be done, both from the point of view of method development and applications, and many exciting developments can be expected for the near future.

Acknowledgments

I would like to thank Olaf Lenz, Claire Loison, Michel Mareschal, Kurt Kremer, and Guido Germano, for enjoyable collaborations that have lead to some of the results discussed in this paper. I am grateful to Alex de Vries, Siewert-Jan Marrink, Gunnar Linke, and Thomas Gruhn, for allowing to use their configuration snapshots. This work was supported by the John von Neumann institute for computing, and by the Deutsche Forschungsgemeinschaft.

References

1. M. Daoud and C. E. Williams, Eds. (1995) *Soft Matter Physics*. Springer-Verlag, Berlin
2. R. G. Larson (1999) *The Structure and Rheology of Complex Fluids*. Oxford University Press, New York
3. J. N. Israelachvili (1992) *Intermolecular and Surface Forces*. Academic Press, London
4. M. Kröger (2003) Simple models for complex nonequilibrium fluids. *Physics Reports* **390**, p. 453
5. P. J. Flory (1969) *Statistical Mechanics of Chain Molecules*. Interscience Publishers, New York
6. P.-G. de Gennes (1979) *Scaling Concepts in Polymer Physics*. Cornell University Press, Ithaca
7. M. Doi (1992) *Introduction to Polymer Physics*. Clarendon Press, Oxford
8. M. Doi and S. Edwards (1986) *The Theory of Polymer Dynamics*. Clarendon Press, Oxford
9. W. B. Russel, D. A. Saville, and W. R. Schowalter (1989) *Colloidal Dispersions*. Cambridge University Press, Cambridge
10. R. J. Hunter (1989) *Foundations of Colloid Science*. Clarendon Press, Oxford
11. M. Borowko, Ed. (2000) *Computational Methods in Surface and Colloid Science*. Marcel Dekker Inc., New York
12. S. A. Safran (1994) *Statistical Thermodynamics of Surfaces, Interfaces, and Membranes*. Addison-Wesley, Reading, MA
13. G. Gompper and M. Schick (1994) *Self-Assembling Amphiphilic Systems*, in *Phase Transitions and Critical Phenomena*, C. Domb and J. Lebowitz Eds., Academic Press, London **16**, p. 1
14. P. G. de Gennes and J. Prost (1995) *The Physics of Liquid Crystals*. Oxford University Press, Oxford

15. S. Chandrasekhar (1992) *Liquid Crystals*. Cambridge University Press, Cambridge
16. P. M. Chaikin and T. C. Lubensky (1995) *Principles of Condensed Matter Physics*. Cambridge University Press, Cambridge
17. F. Müller-Plathe (2002) Coarse-graining in polymer simulation: From the atomistic to the mesoscopic scale and back. *Chem. Phys. Chem.* **3**, p. 754
18. S. O. Nielsen, C. F. Lopez, G. Srinivas, and M. L. Klein (2004) Coarse grain models and the computer simulation of soft materials. *J. Phys.: Cond. Mat.* **16**, p. R481
19. R. B. Gennis (1989) *Biomembranes*. Springer-Verlag, New York
20. R. Lipowsky and E. Sackmann Eds. (1995) *Structure and Dynamics of Membranes – from Cells to Vesicles,* vol. 1 of *Handbook of Biological Physics*. Elsevier, Amsterdam
21. O. G. Mouritsen (2005) *Life – As a Matter of Fat*. Springer, Berlin Heidelberg
22. K. V. Schubert and E. W. Kaler (1996) Nonionic microemulsions. *Ber. Bunseng. Phys. Chemie* **100**, p. 190
23. S. Doniach (1978) Thermodynamic fluctuations in phospholipid bilayers. *J. Chem. Phys.* **68**, p. 4912
24. D. A. Pink, T. J. Green, and D. Chapman (1980) Raman-scattering in bilayers of saturated phosphatidylcholines – experiment and theory. *Biochemistry* **19**, p. 349
25. A. Caillé, D. A. Pink, F. de Verteuil, and M. Zuckermann (1980) Theoretical models for quasi-two-dimensional mesomorphic monolayers and membrane bilayers. *Can. J. Physique* **58**, p. 581
26. B. Dammann, H. C. Fogedby, J. H. Ipsen, C. Jeppesen, K. Jorgensen, O. G. Mouritsen, J. Risbo, M. C. Sabra, M. M. Sperooto, and M. J. Zuckermann (1995) In *Handbook of Nonmedical Applications of Liposomes* vol. 1, Ed. D. D. Lasic and Y. Barenholz, CRC press, p. 85
27. F. Schmid (2000) Systems involving surfactants. In *Computational Methods in Surface and Colloid Science*, *Surfactant Science Series* vol. 89, Ed. M. Borowko, Marcel Dekker Inc., New York, p. 631
28. R. G. Larson, L. E. Scriven, and H. T. Davis (1985) Monte-Carlo simulation of model amphiphilic oil-water systems. *J. Chem. Phys.* **83**, p. 2411
29. T. B. Liverpool (1996) Larson models of Amphiphiles in Complex Fluids. In *Ann. Rev. Comp. Phys. IV*, Ed. D. Stauffer, World Scientific, Singapore, p. 317
30. R. G. Larson (1996) Monte Carlo simulations of the phase behavior of surfactant solutions. *J. Physique II* **6**, p. 1441
31. B. Smit, A. G. Schlijper, L. A. M. Rupert, and N. M. van Os (1990) Effects of chainlength of surfactants on the interfacial tension – molecular-dynamics simulations and experiments. *J. Phys. Chem.* **94**, p. 6933
32. B. Smit, K. Esselink, P. A. J. Hilbers, N. M. van Os, L. A. M. Rupert, and I. Szleifer (1993) Computer simulations of surfactant self-assembly. *Langmuir* **9**, p. 9
33. S. Karaborni, K. Esselink, P. A. J. Hilbers, B. Smit, J. Karthäuser, N. M. van Os, and R. Zana (1994) Simulating the self-assembly of gemini (dimeric) surfactants. *Science* **266**, p. 254
34. B. J. Palmer and J. Liu (1996) Simulations of micelle self-assembly in surfactant solutions. *Langmuir* **12**, p. 746
35. B. J. Palmer and J. Liu (1996) Effects of solute-surfactant interactions on micelle formation in surfactant solutions. *Langmuir* **12**, p. 6015

36. R. Göetz and R. Lipowsky (1998) Computer simulations of bilayer membranes: Self-assembly and interfacial tension. *J. Chem. Phys.* **108**, p. 7397
37. J. C. Shillcock and R. Lipowsky (2002) Equilibrium structure and lateral stress distribution of amphiphilic bilayers from dissipative particle dynamics simulations. *J. Chem. Phys.* **117**, p. 5048
38. M. J. Stevens (2004) Coarse-grained simulations of lipid bilayers. *J. Chem. Phys.* **121**, p. 11942
39. M. Kranenburg, J. P. Nicolas, and B. Smit (2004) Comparison of mesoscopic phospholipid-water models. *Phys. Chem. Chem. Phys.* **6**, p. 4142
40. A. F. Jakobsen, O. G. Mouritsen, and G. Besold (2005) Artifacts in dynamical simulations of coarse-grained model lipid bilayers. *J. Chem. Phys.* **122**, p. 204901
41. D. Harries and A. Ben-Shaul (1997) Conformational chain statistics in a model lipid bilayer: Comparison between mean field and Monte Carlo calculations. *J. Chem. Phys.* **106**, p. 1609
42. A. Baumgärtner (1995) Asymmetric partitioning of a polymer into a curved membrane. *J. Chem. Phys.* **103**, p. 10669
43. A. Baumgärtner (1996) Insertion and hairpin formation of membrane proteins: A Monte Carlo study. *Biophys. J.* **71**, p. 1248
44. T. Sintes and A. Baumgärtner (1997) Short-range attractions between two colloids in a lipid monolayer. *Biophys. J.* **73**, p. 2251
45. O. Lenz and F. Schmid (2004) A simple computer model for liquid lipid bilayers. *J. Mol. Liquids* **117**, p. 147
46. R. Goetz, G. Gompper, and R. Lipowsky (1999) Mobility and elasticity of self-assembled membranes. *Phys. Rev. Lett.* **82**, p. 221
47. H. Noguchi and M. Takasu (2001) Self-assembly of amphiphiles into vesicles: A Brownian dynamics simulation. *Phys. Rev. E* **64**, 041913
48. O. Farago (2003) "Water-free" computer model for fluid bilayer membranes. *J. Chem. Phys.* **119**, p. 596
49. I. R. Cooke, K. Kremer, and M. Deserno (2005) Tunable generic model for fluid bilayer membranes. *Phys. Rev. E* **72**, 011506
50. H. Noguchi and M. Takasu (2001) Fusion pathways of vesicles: A Brownian dynamics simulation. *J. Chem. Phys.* **115**, p. 9547
51. H. Noguchi and M. Takasu (2002) Structural changes of pulled vesicles: A Brownian dynamics simulation. *Phys. Rev. E* **65**, 051907
52. K. Sengupta, V. A. Raghunathan, and J. Katsaras (2003) Structure of the ripple phase of phospholipid multibilayers. *Phys. Rev. E* **68**, 031710
53. A. H. de Vries, S. Yefimov, A. E. Mark, and S. J. Marrink (2005) Molecular structure of the lecithin ripple phase. *PNAS* **102**, p. 5392
54. M. Kranenburg, C. Laforge, and B. Smit (2004) Mesoscopic simulations of phase transitions in lipid bilayers. *Phys. Chem. Chem. Phys.* **6**, p. 4531
55. M. Kranenburg and B. Smit (2005) Phase behavior of model lipid bilayers. *J. Phys. Chem. B* **109**, p. 6553
56. O. Lenz and F. Schmid (2006) Structure of symmetric and asymmetric 'ripple' phases in lipid bilayers. submitted, preprint www.arxiv.org/abs/physics/0608146
57. F. Schmid, D. Düchs, O. Lenz, and B. West (2006) A generic model for lipid monolayers, bilayers, and membranes. submitted, preprint www.arxiv.org/abs/physics/0608226; O. Lenz and F. Schmid, in preparation

58. N. Lei, C. R. Safinya, and R. F. Bruinsma (1995) Discrete harmonic model for stacked membranes – theory and experiment. *J. Phys. II* **5**, p. 1155
59. C. Loison, M. Mareschal, K. Kremer, and F. Schmid (2003) Thermal fluctuations in a lamellar phase of a binary amphiphile-solvent mixture: A molecular dynamics study. *J. Chem. Phys.* **119**, p. 13138
60. T. Soddemann, B. Dünweg, and K. Kremer (2001) A generic computer model for amphiphilic systems. *Eur. Phys. J. E* **6**, p. 409
61. M. Müller and M. Schick (1996) New mechanism of membrane fusion. *J. Chem. Phys.* **116**, p. 2342
62. S.-J. Marrink, F. Jähning, and H. Berendsen (1996) Proton transport across transient single-file water pores in a lipid membrane studied by molecular dynamics simulations. *Biophys. J.* **71**, p. 632
63. D. Zahn and J. Brickmann (2002) Molecular Dynamics Study of Water Pores in a Phospholipid Bilayer. *Chem. Phys. Lett.* **352**, p. 441
64. T. V. Tolpekina, W. K. den Otter, and W. J. Briels (2004) Simulations of stable pores in membranes: System size dependence and line tension. *J. Chem. Phys.* **121**, p. 8014
65. Z. J. Wang and D. Frenkel (2005) Pore nucleation in mechanically stretched bilayer membranes. *J. Chem. Phys.* **123**, p. 154701
66. C. Loison, M. Mareschal, and F. Schmid (2004) Pores in bilayer membranes of amphiphilic molecules: Coarse-Grained Molecular Dynamics Simulations Compared with Simple Mesoscopic Models. *J. Chem. Phys.* **121**, p. 1890
67. C. Loison, M. Mareschal, and F. Schmid (2005) Fluctuations and defects in lamellar stacks of amphiphilic bilayers. *Comp. Phys. Comm.* **169**, p. 99
68. J. D. Lister (1975) Stability of lipid bilayers and red blood-cell membranes. *Physics Lett.* **53A**, p. 193
69. E. W. Weisstein (2003) *CRC Concise Encyclopaedia of Mathematics*. Chapman & Hall CRC, http://mathworld.wolfram.com
70. W. Helfrich (1973) Elastic properties of lipid bilayers – theory and possible experiments. *Z. Naturforschung* **C28**, p. 693
71. E. Evans (1974) Bending resistance and chemically induced moments in membrane bilayers. *Biophys. J.* **14**, p. 923
72. U. Seifert (1997) Configurations of fluid membranes and vesicles. *Adv. Phys.* **46**, p. 13
73. V. A. Kazakov, I. K. Kostov, and A. A. Migdal (1985) Critical properties of randomly triangulated planar random surfaces. *Phys. Lett. B* **157**, p. 295
74. A. Billoire and F. David (1986) Scaling properties of randomly triangulated planar random surfaces – a numerical study. *Nucl. Phys. B* **275**, p. 617
75. Y. Kantor, M. Kardar, and D. R. Nelson (1986) Statistical mechanics of tethered surfaces. *Phys. Rev. Lett.* **57**, p. 791
76. J.-S. Ho and A. Baumgärtner (1990) Simulations of fluid self-avoiding membranes. *Europhys. Lett.* **12**, p. 295
77. D. M. Kroll and G. Gompper (1992) The conformations of fluid membranes – Monte-Carlo simulations. *Science* **255**, p. 968
78. G. Gompper and D. M. Kroll (1997) Network models of fluid, hexatic and polymerized membranes. *J. Phys.: Cond. Matt.* **9**, p. 8795
79. P. B. S. Kumar, G. Gompper, and R. Lipowsky (2001) Budding dynamics of multicomponent membranes. *Phys. Rev. Lett.* **86**, p. 3911
80. P. B. S. Kumar and M. Rao (1998) Shape instabilities in the dynamics of a two component fluid membrane. *Phys. Rev. Lett.* **80**, p. 2489

81. G. Gompper and D. M. Kroll (1997) Freezing flexible vesicles. *Phys. Rev. Lett.* **78**, p. 2859
82. G. Gompper and D. M. Kroll (1998) Membranes with fluctuating topology: Monte Carlo simulations. *Phys. Rev. Lett.* **81**, p. 2284
83. J. C. Shillcock and D. H. Boal (1996) Entropy-driven instability and rupture of fluid membranes. *Biophys. J.* **71**, p. 317
84. H. Noguchi and G. Gompper (2005) Fluid vesicles with viscous membranes in shear flow. *Phys. Rev. Lett.* **93**, p. 258102
85. H. Noguchi and G. Gompper (2005) Shape transitions of fluid vesicles and red blood cells in capillary flows. *PNAS* **102**, p. 14159
86. H. Noguchi and G. Gompper (2006) Meshless membrane model based on the moving least-squares method. *Phys. Rev. E* **73**, p. 021903
87. G. T. Linke (2005) Dissertation Universität Potsdam.
URN: urn:nbn:de:kobv:517-opus-5835
URL: http://opus.kobv.de/ubp/volltexte/2005/583/
88. G. T. Linke, R. Lipowsky, and T. Gruhn (2006) Osmotically induced passage of vesicles through narrow pores. *Europhys. Lett.* **74**, p. 916
89. E. Guyon, J.-P. Hulin, L. Petit, and C. D. Mitescu (2001) *Physical Hydrodynamics*. Oxford University Press, Oxford
90. E. B. Bagley, I. M. Cabott, and D. C. West (1958) Discontinuity in the flow curve of polyethylene. *J. Appl. Phys.* **29**, p. 109
91. T. C. B. McLeish and R. C. Ball (1986) A molecular approach to the spurt effect in polymer melt flow. *J. Polym. Sci.* **24**, p. 1735
92. M. E. Cates, T. C. B. McLeish, and G. Marrucci (1993) The rheology of entangled polymers at very high shear rates. *Europhys. Lett.* **21**, p. 451
93. D. C. Roux, J.-F. Berret, G. Porte, E. Peuvrel-Disdier, and P. Lindner (1995) Shear induced orientations and textures of nematic wormlike micelles. *Macromolecules* **28**, p. 1681
94. J.-F. Berret, G. Porte, and J.-P. Decruppe (1996) Inhomogeneous shear rows of wormlike micelles: A master dynamic phase diagram. *Phys. Rev. B* **55**, p. 1668
95. M. M. Britton and P. T. Callaghan (1999) Shear banding instability in wormlike micellar solutions. *Eur. Phys. J. B* **7**, p. 237
96. E. Fischer and P. T. Callaghan (2001) Shear banding and the isotropic-to-nematic transition in wormlike micelles. *Phys. Rev. E* **64**, 011501
97. M. R. Lopez-Gonzalex, W. M. Holmes, P. T. Callaghan, and P. J. Photinos (2004) Shear banding fluctuations and nematic order in wormlike micelles. *Phys. Rev. Lett.* **93**, p. 268302
98. P. D. Olmsted and P. M. Goldbart (1992) Isotropic-nematic transition in shear flow: State selection, coexistence, phase transitions, and critical behavior. *Phys. Rev. A* **46**, p. 4966
99. G. Porte, J.-F. Berret, and J. L. Harden (1997) Inhomogeneous flows of complex fluids: Mechanical instability versus non-equilibrium phase transition. *J. de Physique II* **7**, p. 459
100. V. Schmitt, C. M. Marques, and F. Lequeux (1995) Shear-induced phase separation of complex fluids – the role of flow-concentration coupling. *Phys. Rev. E* **52**, p. 4009
101. P. D. Olmsted and C. Y. D. Lu (1997) Coexistence and phase separation in sheared complex fluids. *Phys. Rev. E* **56**, p. R55
102. M. P. Lettinga and J. K. G. Dhont (2004) Non-equilibrium phase behaviour of rod-like viruses under shear flow. *J. Phys.: Cond. Matt.* **16**, p. S3929

103. P. D. Olmsted and C. Y. D. Lu (1999) Phase separation of rigid-rod suspensions in shear flow. *Phys. Rev. E* **60**, p. 4397
104. P. D. Olmsted (1999) Two-state shear diagrams for complex fluids in shear flow. *Europhys. Lett.* **48**, p. 339
105. S. M. Fielding and P. D. Olmsted (2003) Flow phase diagrams for concentration-coupled shear banding. *Europ Phys. J. E* **11**, p. 65
106. N. K. Ailawadi, B. J. Berne, and D. Forster (1971) Hydrodynamics and collective angular-momentum fluctuations in molecular fluids. *Phys. Rev. A* **3**, p. 1462
107. X. F. Yuan and M. P. Allen (1997) Non-linear responses of the hard-spheroid fluid under shear flow. *Physica A* **240**, p. 145
108. H. See, M. Doi, and R. Larson (1990) The effect of steady flow fields on the isotropicnematic phase transition of rigid rod-like polymers. *J. Chem. Phys.* **92**, p. 792
109. P. D. Olmsted and P. Goldbart (1990) Theory of the nonequilibrium phase transition for nematic liquid crystals under shear flow. *Phys. Rev. A* **41**, p. 4578
110. J. F. Berret, D. C. Roux, and G. Porte (1994) Isotropic-to-nematic transition in wormlike micelles under shear. *J. de Physique II* **4**, p. 1261
111. J. F. Berret, D. C. Roux, G. Porte, and P. Lindner (1994) Shear-induced isotropic-tonematic phase transition in equilibrium polymers. *Europhys. Lett.* **25**, p. 521
112. E. Cappelaere, J.-F. Berret, J. P. Decruppe, R. Cressely, and P. Lindner (1997) Rheology, birefringence, and small-angle neutron scattering in a charged micellar system: Evidence of a shear-induced phase transition. *Phys. Rev. E* **56**, p. 1869
113. J.-F. Berret, D. C. Roux, and P. Lindner (1998) Structure and rheology of concentrated wormlike micelles at the shear-induced isotropic-to-nematic transition. *Eur. Phys. J. B* **5**, p. 67
114. P. T. Mather, A. Romo-Uribe, C. D. Han, and S. S. Kim (1997) Rheo-optical evidence of a flow-induced isotropic-nematic transition in a thermotropic liquid-crystalline polymer. *Macromolecules* **30**, p. 7977
115. R. G. Larson and D. W. Mead (1993) The Ericksen number and Deborah number cascades in sheared polymeric nematics. *Liquid Crystals* **15**, p. 151
116. J. F. Berret, D. C. Roux, G. Porte, and P. Lindner (1995) Tumbling behavior of nematic worm-like micelles under shear-flow. *Europhys. Lett.* **32**, p. 137
117. A. V. Zakharov, A. A. Vakulenko, and J. Thoen (2003) Tumbling instability in a shearing nematic liquid crystal: Analysis of broadband dielectric results and theoretical treatment. *J. Chem. Phys.* **118**, p. 4253
118. S. Hess and M. Kröger (2004) Regular and chaotic orientational and rheological behaviour of liquid crystals. *J. Phys.: Cond. Matt.* **16**, p. S3835
119. S. Sarman and D. J. Evans (1993) Statistical mechanics of viscous flow in nematic fluids. *J. Chem. Phys.* **99**, p. 9021
120. S. Tang, G. T. Evans, C. P. Mason, and M. P. Allen (1995) Shear viscosity for fluids of hard ellipsoids – A kinetic-theory and molecular-dynamics study. *J. Chem. Phys.* **102**, p. 3794
121. M. P. Allen and D. J. Tildesley (1989) *Computer Simulation of Liquids.* Oxford University Press, New York
122. D. J. Evans and T. P. Morriss (1990) *Statistical Mechanics of Nonequilibrium Fluids.* Academic Press, San Diego

123. S. S. Sarman, D. J. Evans, and P. T. Cummings (1992) Recent developments in non-Newtonian molecular dynamics. *Physics Reports* **305**, p. 1
124. F. Varnik and K. Binder (2002) Shear viscosity of a supercooled polymer melt via nonequilibrium molecular dynamics simulations. *J. Chem. Phys.* **117**, p. 6336
125. A. W. Lees and S. F. Edwards (1972) Computer study of transport processes under extreme conditions. *J. Phys. C* **5**, p. 1921
126. D. J. Evans and G. P. Morriss (1984) Nonlinear-response theory for steady planar couette-flow. *Phys. Rev. A* **30**, p. 1528
127. B. J. Edwards, C. Baig, and D. J. Keffer (2005) An examination of the validity of nonequilibrium molecular-dynamics simulation algorithms for arbitrary steady-state flows. *J. Chem. Phys.* **123**, p. 114106
128. F. Zhang, D. J. Searles, D. J. Evans, J. S. D. Hansen, and D. J. Isbister (1999) Kinetic energy conserving integrators for Gaussian thermostatted SLLOD. *J. Chem. Phys.* **111**, p. 18
129. G. A. Pan, J. F. Ely, C. McCabe, and D. J. Isbister (2005) Operator splitting algorithm for isokinetic SLLOD molecular dynamics. *J. Chem. Phys.* **122**, 094114
130. D. Baalss and S. Hess (1986) Nonequilibrium molecular-dynamics studies on the anisotropic viscosity of perfectly aligned nematic liquid crystals. *Phys. Rev. Lett.* **57**, p. 86
131. S. Sarman (1995) Nonequilibrium molecular dynamics of liquid-crystal shear-flow. *J. Chem. Phys.* **103**, p. 10378
132. S. Sarman (1997) Shear flow simulations of biaxial nematic liquid crystals. *J. Chem. Phys.* **107**, p. 3144
133. J. Liam McWhirter and G. N. Patey (2002) Nonequilibrium molecular dynamics simulations of a simple dipolar fluid under shear flow. *J. Chem. Phys.* **117**, p. 2747
134. J. Liam McWhirter and G. N. Patey (2002) Molecular dynamics simulations of a ferroelectric nematic liquid under shear flow. *J. Chem. Phys.* **117**, p. 8551
135. F. Müller-Plathe (1999) Reversing the perturbation in nonequilibrium molecular dynamics: An easy way to calculate the shear viscosity of fluids. *Phys. Rev. E* **59**, p. 4894
136. D. Frenkel and B. Smit (2002) *Understanding Molecular Simulations.* Academic Press, San Diego
137. D. J. Evans and S. Sarman (1993) Equivalence of thermostatted nonlinear responses. *Phys. Rev. E* **48**, p. 65
138. D. Ruelle (2000) A Remark on the Equivalence of Isokinetic and Isoenergetic Thermostats in the Thermodynamic Limit. *J. Stat. Phys.* **100**, p. 757
139. A. Kolb and B. Dünweg (1999) Optimized constant pressure stochastic dynamics. *J. Chem. Phys.* **111**, p. 4453
140. T. Soddemann, B. Dünweg, and K. Kremer (2003) Dissipative particle dynamics: A useful thermostat for equilibrium and nonequilibrium molecular dynamics simulations. *Phys. Rev. E* **68**, 046702
141. R. Zwanzig (1961) Memory effects in irreversible thermodynamics. *Phys. Rev.* **124**, p. 983
142. H. C. Öttinger (1998) General projection operator formalism for the dynamics and thermodynamics of complex fluids. *Phys. Rev. E* **57**, p. 1416

143. A. N. Gorban, I. V. Karlin, H. C. Öttinger, and L. L. Tatarinova (2001) Ehrenfests argument extended to a formalism of nonequilibrium thermodynamics. *Phys. Rev. E* **63**, 066124
144. B. Dünweg and W. Paul (1991) Brownian dynamics simulations without Gaussian random numbers. *Int. J. Mod. Phys. C* **2**, p. 817
145. G. Germano and F. Schmid (2003) Simulation of nematic-isotropic phase coexistence in liquid crystals under shear. In *Publication Series of the John von Neumann Institute for Computing* **20**, p. 311
146. G. Germano and F. Schmid (2005) Nematic-isotropic interfaces under shear: A molecular dynamics simulation. *J. Chem. Phys.* **123**, p. 214703
147. M. P. Allen (2000) Molecular simulation and theory of the isotropic-nematic interface. *J. Chem. Phys.* **112**, p. 5447
148. A. J. McDonald, M. P. Allen, and F. Schmid (2001) Surface tension of the isotropicnematic interface. *Phys. Rev. E* **63**, p. 10701R
149. N. Akino, F. Schmid, and M. P. Allen (2001) Molecular-dynamics study of the nematic isotropic interface. *Phys. Rev. E* **63**, 041706
150. H. Guo, K. Kremer, and T. Soddemann (2002) Nonequilibrium molecular dynamics simulation of shear-induced alignment of amphiphilic model systems. *Phys. Rev. E* **66**, 061503
151. T. Soddemann, G. K. Auernhammer, H. Guo, B. Dünweg, and K. Kremer (2004) Shear induced undulation of smectic-A: Molecular dynamics simulations vs. analytical theory. *Eur. Phys. J. E* **13**, p. 141
152. V. K. Gupta, R. Krishnamoorti, J. A. Kornfield and S. D. Smith (1995) Evolution of microstructure during shear alignment in a polystyrene-polyisoprene lamellar diblock copolymer. *Macromolecules* **28**, p. 4464
153. K. A. Koppi, M. Tirrell, F. S. Bates, K. Almdal, and R. H. Colby (1992) Lamellae orientation in dynamically sheared diblock copolymer melts. *J. Physique II* **2**, p. 1941
154. G. K. Auernhammer, H. R. Brand, and H. Pleiner (2000) The undulation instability in layered systems under shear flow – a simple model. *Rheol. Acta* **39**, p. 215
155. G. K. Auernhammer, H. R. Brand, and H. Pleiner (2002) Shear-induced instabilities in layered liquids. *Phys. Rev. E* **66**, 061707
156. A. D. Rey and M. M. Denn (2002) Dynamical phenomena in liquid-crystalline materials. *Annu. Rev. Fluid Mech.* **34**, p. 233
157. C. Denniston, E. Orlandini, and J. M. Yeomans (2000) Simulations of liquid crystal hydrodynamics in the isotropic and nematic phases. *Europhys. Lett.* **52**, p. 481
158. C. Denniston, D. Marenduzzo, E. Orlandini, and J. M. Yeomans (2004) Lattice Boltzmann algorithm for three-dimensional liquid-crystal hydrodynamics. *Phil. Trans. Royal Soc. London A* **362**, p. 1745
159. A. N. Beris, B. J. Edwards, and M. Grmela (1990) Generalized constitutive equation for polymeric liquid crystals. 1. Model formulation using the Hamiltonian (Poisson bracket) formulation. *J. Non-Newton. Fluid Mechanics* **35**, p. 51
160. D. Marenduzzo, E. Orlandini, and J. M. Yeomans (2003) Rheology of distorted nematic liquid crystals. *Europhys. Lett.* **64**, p. 406
161. D. Marenduzzo, E. Orlandini, and J. M. Yeomans (2004) Interplay between shear flow and elastic deformations in liquid crystals. *J. Chem. Phys.* **121**, p. 582

Mesoscopic Simulations of Biological Membranes

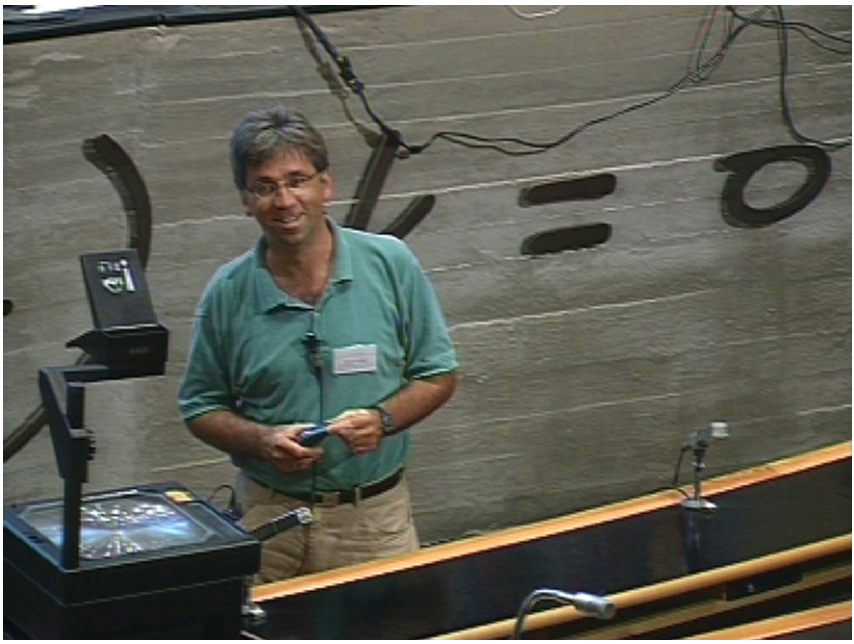
B. Smit[1,2], M. Kranenburg[2], M. M. Sperotto[3], and M. Venturoli[4]

[1] Centre Européen de Calcul Atomique et Moleculaire (Cecam), Ecole Normale Supérieure, 46 Allée d'Italie, 69007 Lyon, France
b.smit@cecam.fr

[2] Van 't Hoff Institute for Molecular Sciences, University of Amsterdam, Nieuwe Achtergracht 166, 1018 WV Amsterdam, The Netherlands
marieke@science.uva.nl

[3] Biocentrum, The Technical University of Denmark, Kgs. Lyngby, Denmark
maria@cbs.dtu.dk

[4] CCS, University College, C. Ingold Labs, 20 Gordon St., London WC1H 0AJ, U.K.
m.venturoli@ucl.ac.uk



Berend Smit

B. Smit et al.: *Mesoscopic Simulations of Biological Membranes*, Lect. Notes Phys. **704**, 259–286 (2006)
DOI 10.1007/3-540-35284-8_11
© Springer-Verlag Berlin Heidelberg 2006

1 Introduction 261

2 Computer Simulations Aspects 263

2.1 Dissipative Particle Dynamics 263
2.2 Surface Tension in Lipid Bilayers 264

3 Phase Behavior of Coarse-Grained Lipid Bilayers 266

3.1 Single-Tail Lipid Bilayers 267
3.2 Double-Tail Lipid Bilayers 272

4 Perturbations of the Membrane Structure 274

4.1 Effect of Alcohol 275
4.2 Bilayers with Transmembrane Proteins 278

5 Concluding Remarks 280

References 281

1 Introduction

A common structure found in all cells and theirs inner organelles is the membrane. Biological membranes act as semi-permeable barriers, allowing a selected passage of small molecules or ions. Biomembranes are constituted of a lipid matrix in which molecules such as proteins or cholesterol are embedded or attached. Lipids are amphiphilic molecules, i.e. molecules constituted of an hydrophilic polar headgroup, which is water soluble, and hydrophobic tails, which are water insoluble. The lipid matrix is formed by the non-covalent self-assembly of two lipid monolayers made of a variety of lipid types.

The combination of hydrophobic and hydrophilic groups in the same molecule is a key factor for the assembly of lipids into supra-molecular aggregates, such as micelles or vesicles, the latter being the templates for the cell membranes. Due to the hydrophobic effect [1,2] membrane lipids assemble in such a way that their hydrophobic part is excluded from a direct contact with the water environment, while the hydrophilic or polar parts are in direct contact with the water. The resulting pseudo two-dimensional system (Fig. 1) is a fluid structure where the molecules may diffuse in the membrane plane, may flip-flop from one monolayer to another, or may even move out of the system.

Biological membranes are not only inert walls, but complex, organized, dynamic, and highly cooperative structures whose physical properties are important regulators of vital biological functions ranging from cytosis and nerve processes, to transport of energy and matter [3].

To relate the biomembrane structure and dynamics to their biological function –the ultimate goal of biomembrane science– is often necessary to consider simpler systems. Lipid bilayers composed of one or two lipid species, and with embedded proteins or natural or artificial peptides, provide a model system for biological membranes. Understanding the physics of such simplified soft-condensed matter systems can yield insight into biological membrane functions.

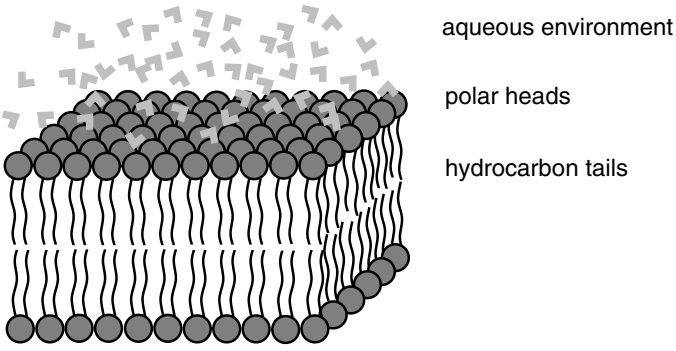


Fig. 1. Schematic representation of a lipid bilayer

Because of the many degrees of freedom involved, the processes that take place even in model biomembranes occur over a wide range of time and length scales [4]. The typical time and length scales of the processes under investigation do pose limitations on the details of the model. Often, this necessity follows the fact that some theoretical methods are limited in their applicability by the long computational times needed to compute statistical quantities.

Molecular Dynamics (MD) simulation methods on atomistic detailed models have been used to study the structural and dynamic properties of membranes [5], the self-assembly of phospholipids into bilayers [6], as well as the interactions of membrane proteins or other molecules with the lipid bilayer [7–12]. MD simulations can provide detailed informations about the phenomena that occur in biomembrane systems, although at the nanoscopic level and on a nanosecond time-scale. Many membrane processes happen though at the mesoscopic length and time scale, i.e. >1–1000 nm, ns, respectively, and involve the collective nature of the system.

An alternative modeling approach consists in neglecting most of the molecular details of the system. The resulting lattice [13, 14], interfacial [15], or phenomenological models [16–18], are computationally very efficient, and can give us insight into the physical properties of reconstituted membranes [19,20]. However, with the help of these models, it is difficult to study some structural and conformational properties of the system, which derive from some molecular details, and which results from its cooperative behavior. To overcome this difficulty, we have developed a model for lipid systems which can be seen as an intermediate between the all-atom models and the models briefly just mentioned. This mesoscopic model considers a system of "particles", or "beads", in which each "particle" represents a complex molecular component of the system whose details are not important to the process under investigation. Models with simplified interactions between the "beads" are called coarse-grain (CG) models. In Fig. 2 the chemical structure of the phospholipid dimyristoylphosphatidylcholine (DMPC) and its CG representation are shown. In the recent years, CG models have been developed to study the phase equilibria of biomembrane-like systems at the mesoscopic level, and both MD

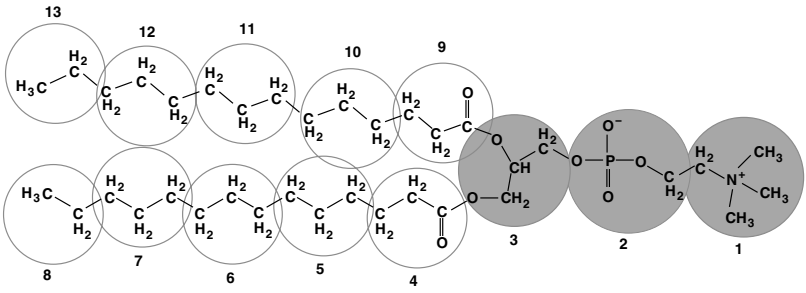


Fig. 2. The atomistic representation of DMPC and its corresponding coarse-grained model

and Monte Carlo (MC) simulation methods were used on such models [21–25]. In this contribution we summarize some of the results of our simulations. A more complete account can be found in [26–34].

2 Computer Simulations Aspects

2.1 Dissipative Particle Dynamics

The simulations presented in this contribution are carried out with Dissipative Particle Dynamics (DPD) [35]. By combining several aspects of Molecular Dynamics and lattice-gas automata, DPD captures hydrodynamic time and length scales much larger than can be reached with the first method and it avoids the lattice artifacts of the latter method. One of the most attractive features of DPD is the versatility in simulating complex fluids. A DPD particle represents a "fluid package" or a cluster of atoms, that moves according to the Newton's equations of motion, interacting with other particles through a dissipative, a random, and a conservative force. By changing the conservative force between different types of particles a fluid can be made "complex". See [36, 37] for nice reviews on DPD and its applications. Recently, DPD has also been used to study the behavior of a lipid bilayer [24, 26, 38, 39].

A DPD particle represents the center of mass of a cluster of atoms. The particles interact via a force consisting of three contributions, all of them pairwise additive. The total force on a particle i consists of a dissipation force $\mathbf{F}^D$, a random force $\mathbf{F}^R$, and a conservative force $\mathbf{F}^C$, and can then be written as the sum of these forces [35, 40]:

$$\mathbf{f}_i = \sum_{i \neq j} (\mathbf{F}_{ij}^D + \mathbf{F}_{ij}^R + \mathbf{F}_{ij}^C) \tag{1}$$

The first two forces in equation 1 are of the form:

$$\begin{aligned} \mathbf{F}_{ij}^D &= -\eta w^D(r_{ij})(\hat{\mathbf{r}}_{ij} \cdot \mathbf{v}_{ij})\hat{\mathbf{r}}_{ij} \\ \mathbf{F}_{ij}^R &= \sigma w^R(r_{ij})\zeta_{ij}\hat{\mathbf{r}}_{ij} \end{aligned} \tag{2}$$

where $\mathbf{r}_{ij} = \mathbf{r}_i - \mathbf{r}_j$ and $\mathbf{v}_{ij} = \mathbf{v}_i - \mathbf{v}_j$, with $\mathbf{r}_i$ and $\mathbf{v}_i$ representing the position and the velocity of particle i, respectively. $\hat{\mathbf{r}}_{ij}$ is the unit vector, η is the friction coefficient, σ the noise amplitude, and ζ_{ij} a random number taken from a uniform distribution, which is independent for each pair of particles. The combined effect of these two forces is a thermostat, which conserves (angular) momentum, and hence gives the correct hydrodynamics at sufficiently long time and length scales.

Español and Warren [41] have shown that the equilibrium distribution of the system is the Gibbs-Boltzmann distribution if the weight functions and coefficients of the drag and the random force satisfy:

$$w^D(r) = [w^R(r)]^2 \tag{3}$$

$$\sigma^2 = 2\eta k_B T \tag{4}$$

The weight function $w^R(r)$ is chosen as

$$w^R(r) = \begin{cases} (1 - r/r_c) & (r < r_c) \\ 0 & (r \geq r_c) \end{cases} \tag{5}$$

where r_c is the cut-off radius, which gives the extent of the interaction range. In this case, all forces assume the same functional dependence on the interparticle distance $\mathbf{r}_{ij}$ as the conservative force $\mathbf{F}_{ij}^C$. Español and Warren have also shown that a DPD Hamiltonian can be defined and the existence of a Hamiltonian implies that only the conservative part of the force (or better, the potential U_C related to it: $\mathbf{F}^C = -\nabla U_C$), determines the equilibrium averages of the system observables. Furthermore, as shown by Willemsen et al. in [42], this implies that DPD can be combined with Monte-Carlo (MC) methods.

Most DPD simulation use soft-repulsive interactions of the form

$$\mathbf{F}_{ij}^C = \begin{cases} a_{ij}(1 - r_{ij}/r_c)\hat{\mathbf{r}}_{ij} & (r_{ij} < r_c) \\ 0 & (r_{ij} \geq r_c) \end{cases} \tag{6}$$

where the coefficient $a_{ij} > 0$ is a parameter expressing the maximum repulsion strength. Alltough any form of conservative interactions can be used in a DPD simulation, these soft-repulsive interactions are often refer to as the DPD model.

2.2 Surface Tension in Lipid Bilayers

Lipid bilayers are self-assembled structures, which are not constrained by the total area, and hence will adopt a conformation that will have the lowest free energy. Since the thermodynamic definition of surface tension is the derivative of the free energy with respect to the area of the interface [43], for a unconstrained bilayer the free energy minimum will be a tensionless state [44]. Experimental results on unilamellar vesicles have also indicated that bilayers are in a stress free state [45]. In molecular simulation, for both self-assembled and pre-assembled membranes, a fixed number of lipid molecules and a fixed area are usually combined with periodic boundary conditions. The periodic boundary conditions correspond to an infinitely large membrane, but the fixed size of the simulation box, and the fixed number of lipids at the interface, impose a constraint on the bilayer area which results in a finite surface tension. Although the constraint on the fixed area can be released by performing simulations of membranes at constant pressure or constant surface tension [46,47], it is an important –and still open– question which value of the surface tension should be used in simulations to reproduce the state and the area per lipid of a real membrane.

In their molecular dynamics simulations Feller and Pastor [48,49] observed that a tensionless state did not reproduce the experimental value of the area per lipid. They explained this result by considering that, since the typical undulations and and out-of-plane fluctuations of a macroscopic membrane do not develop in a small patch of a membrane (theirs was composed of 36 lipids for monolayer), a positive surface tension (stretching) must be imposed in order to compensate for the suppressed undulations, and hence to recover the experimental values of the area per lipid. Recently, Marrink and Mark [50] investigated the system size dependence of the surface tension in membrane patches ranging from 200 to 1800 lipids, simulated for times up to 40ns. Their calculations show that, in a stressed membrane, the surface tension is size dependent, i.e. it drops if the system size is increased (at fixed area per lipid), which is in agreement with the results by Feller and Pastor. On the other hand, their results show that at zero stress simulation conditions the equilibrium does not depend on the system size. Marrink and Mark then concluded that simulations at zero surface tension correctly reproduce the experimental surface areas for a stress free membrane.

Simulations at constant surface tension have been introduced by Chiu et al. in [46]. A constant surface tension ensemble ($NV\gamma$) has been considered in literature and the corresponding equations of motion for Molecular Dynamics simulations have been derived [47], and applied to the simulation of phospholipid bilayers [49, 51–55]. Alternatively, to ensure a tensionless state, Goetz and Lipowsky [21] performed several simulations to determine the area per lipid that gives a state of zero tension. An alternative approach it to combine DPD or MD with a Monte Carlo move that imposes a given value for the surface tension [26]. The advantage of these constant surface tension scheme is that one can directly observe phase transitions in which the area per lipid changes.

Consider a simulation box (see Fig. 3) with edges $L_x = L_y = L_\parallel$ parallel to the interface (xy plane), and $L_z = L_\perp$ perpendicular to the interface (z axis), so that the system volume is $V = L_\perp L_\parallel^2$ and the area of the interface $A = L_\parallel^2$. We define a transformation of the box sizes which changes the area and the height but keeps the volume constant. Such a transformation can be written in the form

$$\begin{aligned} L'_\parallel &= \lambda\, L_\parallel \\ L'_\perp &= \frac{1}{\lambda^2}\, L_\perp \end{aligned} \tag{7}$$

where λ is the parameter of the transformation. By changing λ, the above expression generates a transformation of coordinates which preserves the total volume of the system, hence no work against the external pressure is performed. In a MC move an attempt of changing the parameter λ is then accepted with a probability

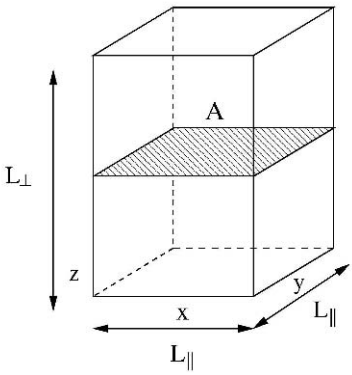


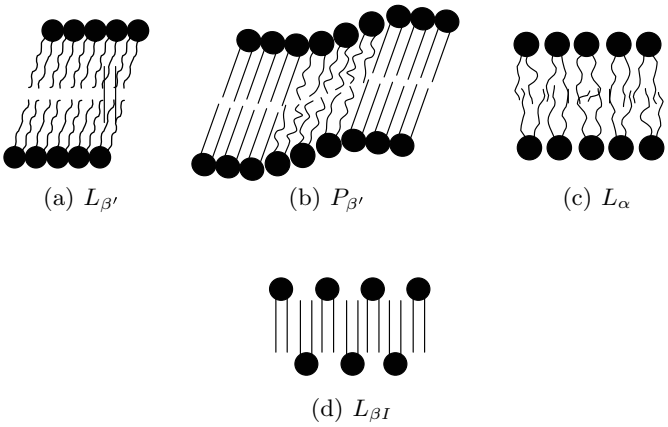
Fig. 3. Schematic representation of a simulation box for a system with a flat interface parallel to the xy-plane. The area of the interface is $A = L_{\parallel}^2$ and the box dimension perpendicular to the interface (z-axis) is $L_{\perp}$

$$\mathrm{acc}(\lambda \to \lambda') = \frac{\exp\left\{-\beta\left[U(\mathbf{s}'^N;\lambda') - \gamma A(\lambda')\right]\right\}}{\exp\left\{-\beta\left[U(\mathbf{s}^N;\lambda) - \gamma A(\lambda)\right]\right\}} \tag{8}$$

where γ is the imposed surface tension. If we choose the particular value $\gamma = 0$, then the explicit term depending on the area in equation 8 drops. The described scheme can be applied to impose any value of the surface tension. It is important to remark that this scheme assumes that the stress tensor is diagonal, which is true for fluid systems.

3 Phase Behavior of Coarse-Grained Lipid Bilayers

The phase behavior of different phosphocholines (PC's) has been determined experimentally (see [56] for a review). All PC's have a low temperature $L_{\beta'}$ phase (see Fig. 4(a)). In this phase the bilayer is a gel: the chains of the phospholipids are ordered and show a tilt relative to the bilayer normal. At higher temperature the L_{α} phase is the stable phase. This phase is the liquid crystalline state of the bilayer in which the chains are disordered and tail overlap due to this thermal disorder is possible. This phase is physiologically the most relevant [57]. Under normal conditions the two monolayers of a bilayer contact each other at the terminal methyl group of their hydrophobic chains, while their hydrophilic headgroups are in contact with water. However, it is known experimentally that at low temperatures an interdigitated state, in which the terminal methyl groups of one monolayer interpenetrate the opposing layer, is also possible. This $L_{\beta I}$ phase does not spontaneously form in bilayers of symmetrical chain phospholipids, like dipalmitoylphosphatidylcholines (DPPC), [58] but can be induced by changes in the environment, like hydrostatic pressure or changes in the pH of the solution [59], or by incorporation, at the membrane interface of small amphiphilic molecules, like



(a) $L_{\beta'}$ (b) $P_{\beta'}$ (c) L_{α}

(d) $L_{\beta I}$

Fig. 4. Schematic drawings of the various bilayer phases. The characteristics of these phases are explained in the text. The filled circles represent the hydrophilic headgroup of a phospholipid and the lines represent the hydrophobic tails

alcohols [60–62], or anesthetics [63]. Interdigitation can also be induced by changes in the lipid structure, for example by introducing an ester-linkage in the headgroup of the phospholipids [64, 65].

An important question in the development of a mesoscopic model is how much chemical detail should be included in the model. Most likely this depend on the properties one would like to study. Here, we focus on the phase behavior and we illustrate how one can use mesoscopic models to obtain insights in how changes in the molecular structure or interactions of the lipid change its phase behavior. This insight can subsequently be used for the development of such a mesoscopic model.

3.1 Single-Tail Lipid Bilayers

Single-tail lipids are studied by several groups as models for phospholipid membranes [26] and it is an interesting question of such a simple amphiphilic molecule can describe the phase behavior of say DMPC. This question was addressed by Kranenburg et al. [27] using a model in which the lipid is represented by one head segment connected to a single tail with variable length as shown in Fig. 5. These authors have used the soft repulsion which are often used in DPD simulations. With this model the authors studied the effect of changes in the head-head interactions, caused, by, for example, adding salt to the system, on the phase diagram.

The computed phase diagram as a function of temperature and head-head interactions for the lipid ht_9 is shown in Fig. 6. If one would use this model to describe DMPC one would take a head-head repulsion of $a_{\mathrm{hh}} = 35$. For this value of the head-head repulsion, an interdigitated gel phase, $L_{\beta I}$ is observed at low temperatures, while at $a_{\mathrm{hh}} = 15$ the non-interdigitated L_{β} phase is

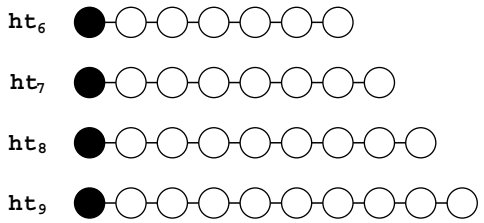


Fig. 5. Schematic drawing of the model lipids used by Kranenburg et al. [27]. The black particles represent the head beads and the white particles the tail beads. Two consecutive beads are connected by harmonic springs with spring constant and a harmonic bond bending potential between three consecutive beads is added with a bending constant $K_\theta = 10$ and an equilibrium angle $\theta_0 = 180^\circ$. The soft-repulsion parameters used are $a_{\rm ww} = a_{\rm tt} = 25$, $a_{\rm wh} = 15$, and $a_{\rm wt} = 80$

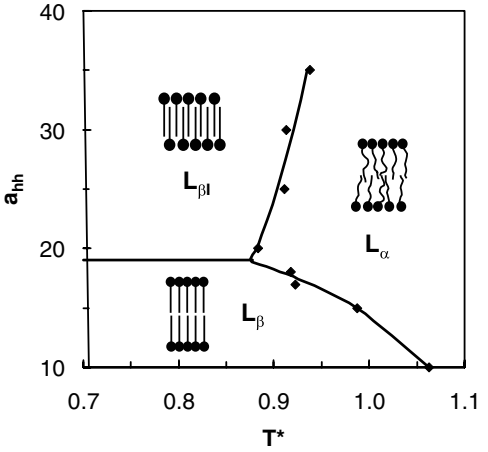
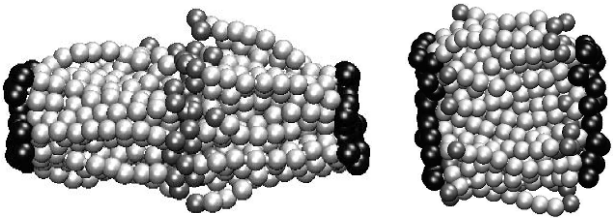


Fig. 6. Computed phase diagram of the lipid ht_9 as function of head-head repulsion parameter $a_{\rm hh}$ and reduced temperature T^*. At high values of the head-head repulsion parameters the interdigitated $L_{\beta I}$ phase is formed, while at low values the non-interdigitated L_β phase is formed. Increasing temperature causes the melting of the bilayer to the L_α phase. These simulations involve a tensionless membrane of 200 lipids and the total number of particles was 3500. The overall density of the system is $\rho = 3$. The results are expressed in the usual reduced units, i.e. using R_c as the unit of length and $k_B T_o = 1$, with T_o room temperature, as unit of the energy. T^* is the temperature expressed in this unit

formed. The high temperature phase is the liquid or L_α phase. A snapshot of the different phases is given in Fig. 7.

Characterization of the Bilayer Phases

Before discussing this transition in detail, it is interesting to see what happens in the low and the high temperature phases. To characterize the ordering of



(a) (b)

Fig. 7. Snapshots of the simulations of a bilayer consisting of the lipid ht_9 at $T^* = 0.85$. (**a**) The non-interdigitated gel phase L_β at $a_{\rm hh} = 15$ and (**b**) the interdigitated gel phase L_β at $a_{\rm hh} = 15$. *Black* represents the hydrophilic headgroup and *gray* represents the hydrophobic tails

the lipids in the bilayer we use the order parameter $S_{\rm tail}$, is a measure of the order in the tails. In Fig. 8 $S_{\rm tail}$ is plotted as a function of temperature. $S_{\rm tail}$ has two regimes; below $T^* = 0.95$ where $S_{\rm tail}$ has values higher than 0.5 indicating that the chains are ordered along the bilayer normal, and above $T^* = 0.95$ where the values of $S_{\rm tail}$ decrease below 0.5, showing an increase in the disorder of the chains.

The density profiles show that in the low temperature region, the two monolayers are interdigitated. The lipids stretch out in the direction normal to the bilayer, inducing interdigitation. This packing results in a larger average distance between the lipids headgroups in each monolayer and in a larger area. In this region an increase of temperature reduces the values of the order parameter, but along the chain the order persists. Thus interdigitation is still present, but is decreasing in depth, resulting in an increase of the bilayer thickness and a decrease of the area per lipid. Above the transition temperature, the chains loose the persisting order and are not interdigitated. Only the terminal tail beads overlap, due to thermal disorder. In this temperature region an increase in temperature increases the effective volume occupied by the molecules, but the extent of tail overlap does not depend significantly of temperature. As a result the area per molecule increases while the bilayer thickness decreases.

Determination of the Phase Boundaries

Three quantities have been used to distinguish among the different phases: the area per lipid $A_{\rm L}$, the extent of tail overlap $D_{\rm overlap}$, and the ordering of the tails $S_{\rm tail}$. Figure 8(a) shows the area per lipid $A_{\rm L}$, as function of temperature and head-head repulsion parameter. For repulsion parameters $a_{\rm hh} \leq 18$, the low temperature phase is the bilayer gel L_β phase, while for repulsion parameters $a_{\rm hh} > 18$, the low temperature phase is the interdigitated gel $L_{\beta I}$.

By increasing temperature all bilayers melt from an ordered into a disordered phase. For bilayers in the L_β phase, the area per molecule and chain

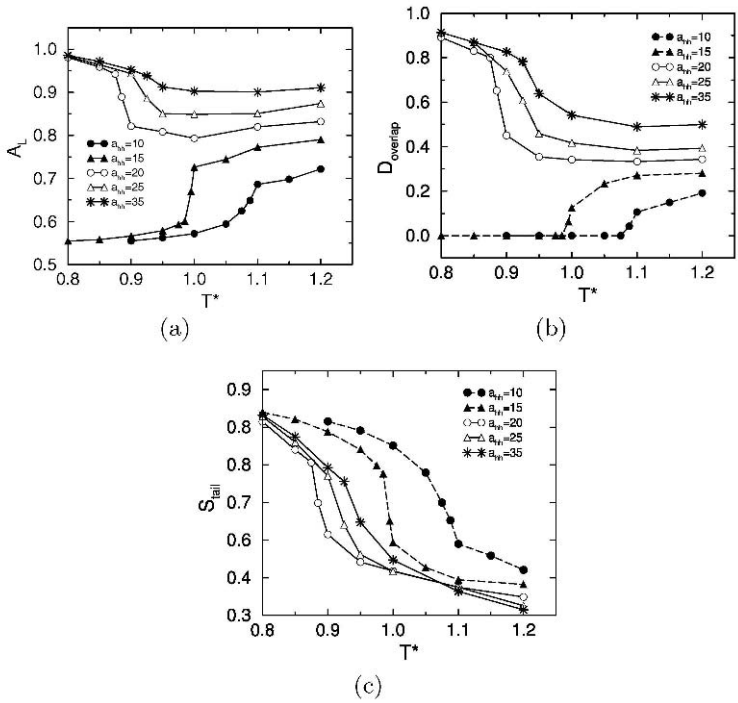


Fig. 8. (**a**) Area per lipid A_{L}, (**b**) extent of chain overlap D_{overlap}, and (**c**) tail order parameter S_{tail} as function of reduced temperature T^* for different repulsion parameters a_{hh}. *Dashed curves* show a transition from the L_β to the L_α phase, *solid curves* show the transition from the $L_{\beta I}$ to the L_α phase

overlap increase upon melting, while for bilayers in the $L_{\beta I}$ phase the area per molecule and chain overlap decrease.

The curves in Fig. 8(c) show that the transition from an ordered phase to a disordered one is very gradual. Much larger systems might be required to observe a sharp transition in these quasi two-dimensional systems. This gradual transition makes it difficult to determine the exact location of the phase boundaries and therefore we used the inflection point as our definition of the phase boundary. The temperature at which the chains get disordered is the same as the temperature of the inflection point in A_{L} and D_{overlap}. We define as the main transition temperature T_{m} the value of temperature at the inflection point of the shown curves. T_{m} is higher for bilayers in the L_β phase than for bilayers in the $L_{\beta I}$ phase. This is in agreement with experimental results [59].

Phase Behavior as a Function of Tail Length

Besides investigating the effect of changing the head-head repulsion parameter, it is also interesting to vary the tail length of the lipid. A similar analysis,

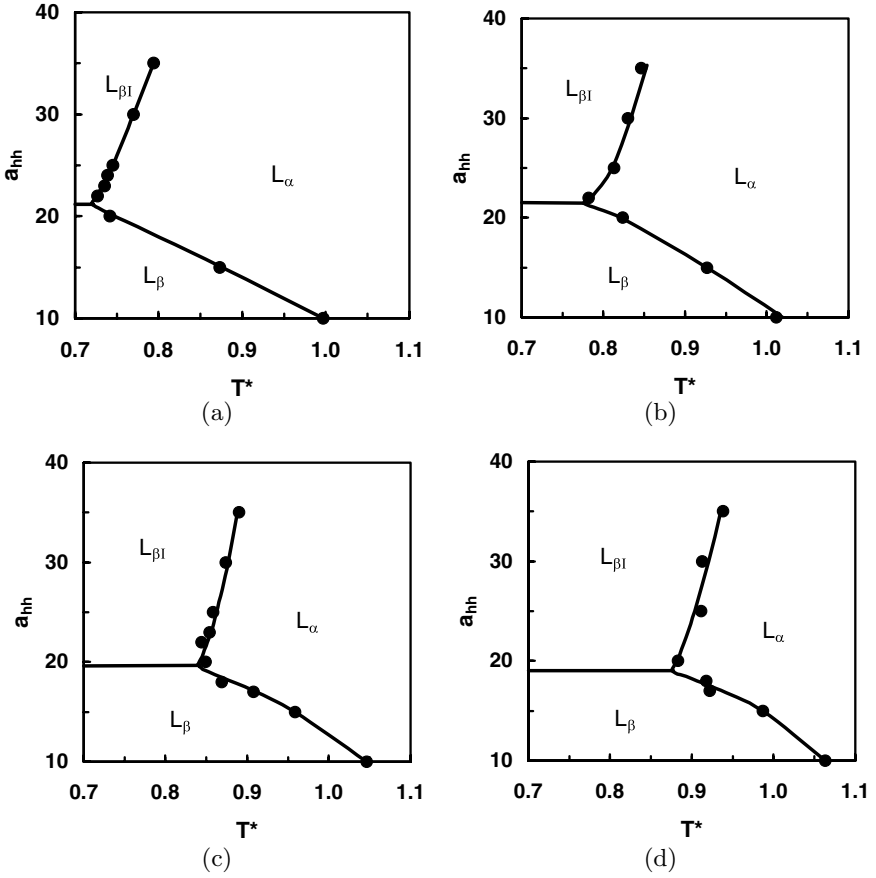


Fig. 9. Phase diagrams as a function of the head-head repulsion parameter a_{hh} and reduced temperature T^* for lipids of different chain lengths: (**a**) ht_6 (**b**) ht_7 (**c**) ht_8, and (**d**) ht_9

as was presented for the lipid ht_9, has been carried out for lipid types ht_6, ht_7, and ht_8 (see Fig. 9).

Depending on the repulsion parameter we obtain two gel phases $L_{\beta I}$ and L_β for all tail lengths. For high head-head repulsion the system can gain energy by adding water particles in between the heads. As a result the distance between the head groups increases and the interdigitated phase is stabilized. For low values of a_{hh} the headgroups expel water and the stable phase is the non-interdigitated phase. In between we find a^*_{hh} for which the transition from $L_{\beta I}$ to L_β occurs. The difference between the two phases is that in the $L_{\beta I}$ phase the tail ends are in direct contact with water, whereas in the L_β phase the tail ends face each other. Therefore, the critical value a^*_{hh} to induce interdigitation is higher than the value of a_{wh}.

As we increase the tail length the gel phases are stabilized and the transition shifts to higher temperatures. The effect of increasing the head-head repulsion on the gel to liquid crystalline transition temperature is much more pronounced for the $L_\beta \rightarrow L_\alpha$ compared to $L_{\beta I} \rightarrow L_\alpha$. This can be understood from the fact that in the interdigitated phase the average distance between the heads is already much larger compared to the non-interdigitated phase, and a further increase in this distance does not have a dramatic effect on the stability of the gel phase.

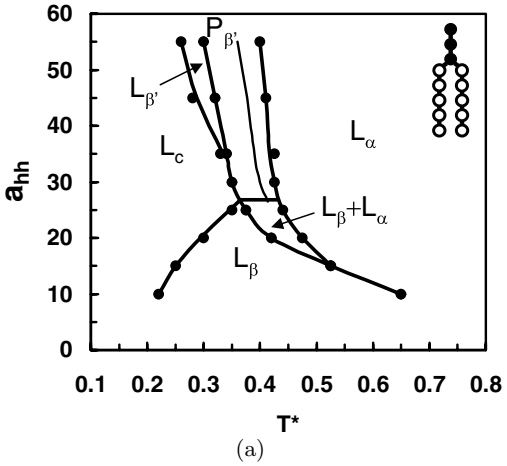
For lipids ht_8 and ht_9 the $L_{\beta I}$ phase occurs at slightly lower repulsion parameters than for lipids ht_6 and ht_7. This is consistent with experimental results [66]. Since the interdigitated phase is more closely packed than the non-interdigitated phase, the van der Waals energy is greater. This energy gain is proportional to the number of carbon atoms in the phospholipid chain and thus interdigitation becomes energetically more favorable for longer chains. Also in our simulations we observe that the interdigitated phase is more compact and hence a^*_{hh} decreases slightly with increasing tail length.

It is interesting to compare these results with the experimental data. Misquitta and Caffrey in [67] systematically investigate the phase diagrams of monoacylglycerols, a single-tail lipid, and show a similar tail length dependence for the $\mathrm{L}_\beta \rightarrow \mathrm{L}_\alpha$ transition. Interestingly, as we will show later, for a similar model of a double-tail lipid we do not observe the spontaneous formation of an interdigitated phase. This corresponds to the experimental observation that for the most common double-tail lipids the interdigitated phase does not form spontaneously, but should be induced by the addition of, for example, alcohol [58].

3.2 Double-Tail Lipid Bilayers

In the previous section we have discussed the phase behavior of single-tail lipid bilayers. In this section we extend the model and investigate the phase behavior of a double-tail lipid with three head-beads and two tails of five beads each ($h_3(t_5)_2$). The computed phase diagram is shown in Fig. 10. No interdigitation was found for the chosen values of the repulsion parameter between the headgroups. Also, an increase of a_{hh} up to 55 does not lead to any interdigitation (data not shown). This result is consistent with the experimentally observed structure of symmetric PC's bilayers, for which no spontaneous interdigitation is found. The snapshots in Fig. 11(c) shows a typical configuration of the system in the various phases.

At very low temperatures the system is in the $L_{\beta'}$ gel phase, which is characterized by having ordered chains, hence a high value of the bilayer thickness and of the tail order parameter. While single-tail lipids are not tilted in the gel phase, for the double-tail lipid we observe that the lipid chains are tilted with respect to the bilayer normal. We find a tilt angle of 25°, which is slightly lower than the value of $\approx 32°$ measured experimentally for DMPC

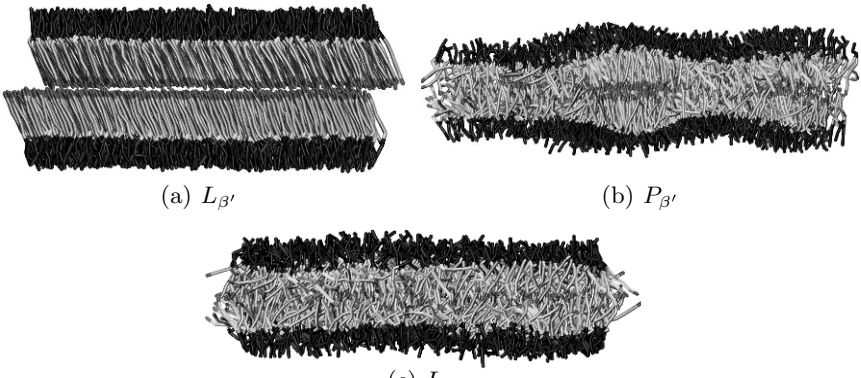


(a)

Fig. 10. Phase diagram of model lipid, $h_3(t_5)_2$, as a function of head-head repulsion and reduced temperature

lipid bilayers [68]. A typical configuration at this temperature can be see in the snapshot in Fig. 11(a).

Between the L_α and the $L_{\beta'}$ phases, when the temperature is increased above T^*=0.35, we observed a third phase. This phase, which disappears again as the temperature reaches the main-transition temperature, is characterized by having striated regions made of lipids in the gel-state intercalated by



(a) $L_{\beta'}$ (b) $P_{\beta'}$

(c) L_α

Fig. 11. Snapshots of typical configurations of the $h_3(t_5)_2$ bilayer simulated at reduced temperatures: (**a**) $T^* < 0.35$, corresponding to the gel phase, or $L_{\beta'}$; (**b**) $0.35 \leq T^* < 0.425$ corresponding to the ripple-like "striated" phase, or $P_{\beta'}$; and (**c**) $T^* > 0.425$ corresponding to the fluid phase, or L_α. The lipid headgroups are represented by *black lines* and the lipid tails by *gray lines*, with the terminal tail beads darker gray. The water is not shown

Table 1. Values from simulation of the bilayer hydrophobic thickness, D_c, and area per lipid, A_L, at different temperatures, compared with experimental values. The error on the simulation data is 0.2 Å for D_c and 0.4 Å^2 for A_L.

T [°C]	Phase	D_c [Å]		A_L [Å^2]	
		sim	exper	sim	exper
10	$L_{\beta'}$	34.3	30.3[†]	48.6	47.2[†]
30	L_α	26.3	25.6[‡]	60.4	60.0[‡]
50	L_α	24.3	24.0[‡]	64.4	65.4[‡]
65	L_α	23.6	23.4[‡]	65.7	68.5[‡]

[†] From [68]. The error for D_c is 0.2 Å, and for A_L 0.5 Å^2.
[‡] From [70]. The error is not reported in the cited reference.

regions made of lipids in the fluid-state. This modulated structure can be seen in the snapshot in Fig. 11(b). This phase resembles the $P_{\beta'}$, or ripple-phase. The ripple-phase occurs in phospholipid bilayers at the so-called pre-transition temperature, and is characterized by a rippling of the bilayer, with a wave length of the order of 150 Å [69].

It is interesting to make a more quantitative comparison of our course-grained model with a real phospholipid. The double-tail lipid $h_3(t_5)_2$ can be mapped onto DMPC, if a coarse-grained representation is used in which one DPD bead has a volume of 90 Å.[3] The unit of length is derived, from the volume of one DPD bead, and is equal to $R_c = 6.4633$ Å. To relate the temperatures in our model to real temperatures, we use the the main- and pre-transition temperatures for a pure DMPC phospholipid bilayer (≈24 °C and ≈14 °C [56] of the main- and pre-transition temperatures, respectively).

The values of the bilayer hydrophobic thickness and the area per lipid obtained from our simulations can now be compared with the corresponding experimental values for fully hydrated DMPC bilayers, as shown in Table 1 The values from simulations are in good quantitative agreement with the experimental data, although some deviations from the experimental values are observed for the area per lipid at high temperature (65°C), and for the bilayer hydrophobic thickness in the gel phase (10°C). The larger thickness, compared with the experimental value, found in our simulations in the gel phase, could be due to the smaller tilt angle shown by the coarse grained lipids compared to the DMPC lipids.

4 Perturbations of the Membrane Structure

In the previous section we have shown that one can obtain a very reasonable description of the phase behavior of a phospholipid membrane. Here we

show that this model can be used to study the effect of adding alcohol or a transmembrane peptide on the structure of a membrane.

4.1 Effect of Alcohol

Small molecules such as alcohols or anesthesia can adsorb inside a membrane and this adsorption may induce changes in the structure of the membrane which can have an influence on the functioning of the membrane [58]. To study the effect of these molecules on the properties studies on various model membranes have been carried out by various groups [66, 71–76]. These studies show that at low temperatures adding alcohol leads the formation of an interdigitated or $L_{\beta I}$ phase in stead of the commonly formed L_c or $L_{\beta'}$ (gel phases).

The structure of the interdigitated phase as been studied in detail and based on the available experimental data Adachi et al. [62] proposed the model shown in Fig. 12, in which the alcohol molecules fill the space and prevent the hydrophobic tails of the lipids to be exposed to water. This model nicely explains the experimentally observed density profiles and provides a simple molecular explanation why the alcohol molecules stabilize the interdigitated phase. This model also suggests that the optimal alcohol to lipid ratio in the membrane is 2:1. Therefore, measuring this ratio in a membrane would be an important conformation of this model. However, experimentally it is very difficult to measure the concentration of alcohol in the membrane directly. Most studies use theoretical models to relate the alcohol concentration in the water phase to the alcohol concentration in the membrane. Since the validity of these models has not been tested, it is not yet clear whether the 2:1 ratio can indeed be confirmed experimentally.

The model of the alcohols consists of one hydrophilic head bead an a tail that varies in length from one to three hydrophobic beads. In the coarse-graining procedure the alcohols methanol through pentanol correspond with the coarse-grained models ht and ht_2 and hexanol and heptanol correspond

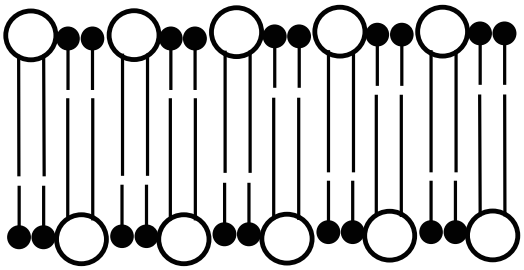


Fig. 12. Proposed model of the interdigitated phase by alcohols [62], in which every tail end of a phospholipid is facing the tail end of an alcohol. *Black* molecules represent the phospholipids and *gray* molecules the alcohols

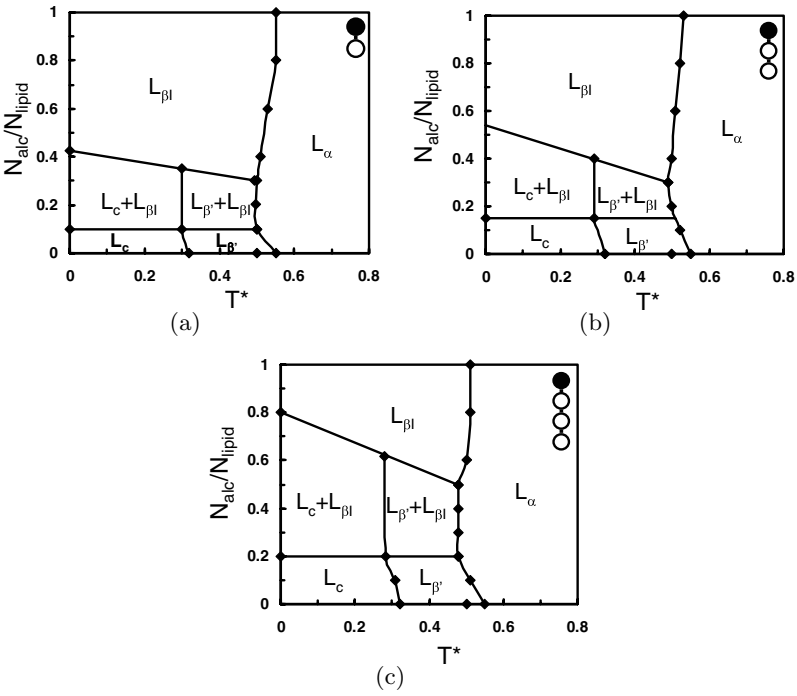


Fig. 13. Phase diagrams of the lipid $h_3(t_7)_2$ with the model alcohols ht (**a**), ht_2 (**b**), and ht_3 (**c**)

with the model ht_3. The bond-bending potential with $k_\theta = 6$ and $\theta = 180°$ was also applied between three consecutive beads of the these alcohols.

Figure 13 shows the effect of alcohols with different tail lengths on the phase behavior of a membrane. At zero alcohol concentration we nicely recover the experimentally observed phases. At low temperatures we observe the L_c phase, in which the tails are order and tilted. If we increase the temperature the tails lose their order in the $L_{\beta'}$ phase, and at even higher temperatures the tails also loose their tilt in the fluid or L_α phase.

At low concentrations of alcohol, the molecules are homogeneously distributed in the $L_{\beta'}$ or L_c phase. At high alcohol concentrations we find the $L_{\beta I}$, in which the lipid tails do not have a tilt with respect to the bilayer normal. The tails of the lipids of one monolayer are fully interpenetrated into the opposing layer and the tail ends are facing the tail end of the alcohol. In between these two extremes we find that there is coexistence between the interdigitated and non-interdigitated phase. The alcohols are inhomogeneously distributed in the bilayer and mainly located in the interdigitated part of the bilayer.

At lower temperatures the concentration of alcohol required to obtain the fully interdigitated phase increases, which is in agreement with the experimental observations of Nambi et al. [72]. Our simulations also report the so-called

"biphasic effect" on the transition from the low temperature gel phase to the high temperature fluid phase [71]; at low concentrations of alcohol, the main transition temperature shifts to a lower temperature, while at high concentrations this transition temperature shifts to a higher temperature compared to a pure lipid bilayer. Figure 13 shows that the increase or decrease in temperature are directly related to the effect of the alcohol on the stability of the low temperature phases [66]. The $L_{\beta'}$, phase is destabilized by alcohol, while the $L_{\beta I}$ can easily incorporate alcohol molecules.

The phase diagrams also show the effects of changes of the length of the alcohol. We observe little influence on the $L_{\beta I} \rightarrow L_{\alpha}$ transition. However, we do observe a difference in the stability of the $L_{\beta I}$ phase as a function of the concentration of alcohol; the shorter the alcohol, the more stable $L_{\beta I}$ phase.

An important issue is the concentration of alcohol in the bilayer. The critical *bulk* concentration at which interdigitation is complete decreases with increasing length of the alcohol [71,75,77]. Various theories have been used to relate the bulk concentration of alcohol to the concentration in the lipid. At present there is little consensus in the literature whether this procedure yields a reliable estimate of the alcohol concentration in the bilayer [78–82]. In our simulations we can compute the number of alcohols in the bilayer directly. Comparison of the phase diagrams (Fig. 13) for three different alcohols shows that if we increase the length of the alcohol more alcohol molecules are needed to destabilize the L_c and L_{β} phases. The $L_{\beta I}$ requires a higher concentration of the longer chain alcohols to be stable, which agrees nicely with experimental results [77].

Adachi et al. [62] proposed a model of the structure of the bilayer in which the terminal methyl group of the alcohol faces a terminal methyl group of a lipid chain (see Fig. 12). The assumption is based on the experimental observation that the membrane thickness increases by about 0.08 nm per one methylene unit in either the alcohol molecules or the phospholipids [62, 83]. This distance of 0.1 nm is the length of one CH_2-unit in the stretched chain of an alkane [84]. Furthermore, Adachi et al. [62] show that two alcohol molecules can occupy a volume surrounded by the PC head groups of one layer. From this it follows that the number of alcohol molecules should be twice as high as the number of lipids in the bilayer independent of the length of the alcohol. This conclusion clearly differs with the results of our simulations which do indicate that the stability of the interdigitated phase depends on the length of the alcohol.

We observe that the fully interdigitated phase occurs at much lower number of alcohols than twice the number of lipids in the lipid bilayer. This can be explained by taking into account an energy balance between the non-interdigitated and interdigitated phase. By the incorporation of alcohols at the membrane interface, voids are created in the hydrophobic core, which are energetically
unfavorable. The more alcohol, the higher the energy of the membrane. In the interdigitated phase the tail ends of the lipid are in contact with the

interfacial water. Adding alcohol reduces these energetically unfavorable interactions and hence increasing the alcohol concentration decreases the energy of the membrane. Clearly at the 2 : 1 ratio, the energy of the interdigitated phase will be lowest, but at much lower alcohol concentrations the energy can already be lower compared to the $L_{\beta'}$ phase. To use this model quantitatively, one would need to take entropy effects into account as well, but this model does rationalize why we observe in our simulations already an interdigitated phase at much lower alcohol concentrations.

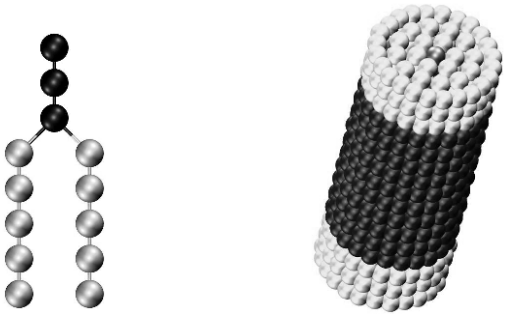
4.2 Bilayers with Transmembrane Proteins

The hydrophobic matching between the lipid bilayer hydrophobic thickness and the hydrophobic length of integral membrane proteins has been proposed as a generic physical principle on which the lipid-protein interaction in biomembranes is based [14, 20, 85–88]. The energy cost of exposing polar moieties, from either hydrocarbon chains or protein residues, is so high that the hydrophobic part of the lipid bilayer should match the hydrophobic domain of membrane proteins. The results from a number of investigations have indeed pointed out the relevance of the hydrophobic matching in relation to the lipid-protein interactions, hence to membrane organization and biological function.

In this work we study the perturbation caused by a transmembrane peptide on the surrounding lipids, its possible dependence on hydrophobic mismatch, protein size, and on temperature. We have investigated whether and to which extent – due to hydrophobic mismatch and via the cooperative nature of the system – a protein may prefer to tilt (with respect to the normal to the bilayer plane), rather than to induce a bilayer deformation without (or even with) tilting.

The systems that we have simulated are made of model lipids having three headgroup beads and two tails of five beads each; this corresponds to the case of acyl chains with fourteen carbon atoms, namely to a model for a dimyristoylphosphatidylcholine (DMPC) phospholipid, as illustrated in in Fig. 14(a).

Within the model formulation, a peptide is considered as a rod-like object, with no appreciable internal flexibility, and characterized by a hydrophobic length d_{P}. The model for the transmembrane peptide is built by connecting $\mathrm{n_{t_P}}$ hydrophobic beads into a chain, to which ends $\mathrm{n_h}$ headgroup-like beads are attached. N_{P} of these amphiphatic chains are linked together into a bundle. In each model protein, all the N_{P} chains are linked to the neighboring ones by springs, to form a relatively rigid body. We have considered three typical model-protein sizes, two of them referring to a "skinny" peptide-like molecule, and the third type to a "fat" protein. These model proteins consist of $N_{\mathrm{P}} = 4$, 7, or 43 chains linked together in a bundle. The bundle of $N_{\mathrm{P}} = 7$ chains is formed by a central chain surrounded by a single layer of six other chains. The $N_{\mathrm{P}} = 43$ bundle is made of three layers arranged concentrically around a central chain, and containing each six, twelve, and twenty four amphiphatic



(a) (b)

Fig. 14. Schematic representation of a model-lipid (**a**), and a model protein ($N_P = 43$ and $\tilde{d}_P = 41$Å) (**b**)

chains, respectively. Figure 14(b) shows a cartoon of a model protein of size $N_P = 43$.

One of the quantities that can be measured experimentally is the tilt of the peptide. The peptide tilt angle with respect to the bilayer normal as function of Δd and peptide size N_P is shown in Fig. 15. The snapshots on the right show typical configurations of the system, for a fixed value of the peptide hydrophobic length, for the three peptide sizes, $N_P = 4$, 7, and 43, and for the largest (positive) value of mismatch, $\Delta d = 26$ Å, at the considered temperature, $T^* = 0.7$. For $\Delta d < 0$ the tilt angle is very small, and is within the statistical tilt-fluctuations to which the peptide is subject in the bilayer; as the peptide hydrophobic length increases (and the mismatch becomes positive), the peptide undergoes a significant tilting. Also, for equal values of hydrophobic mismatch, the "thinner" peptide ($N_P = 4$) is much more tilted than the "fatter" one ($N_P = 43$). These results, combined with the one discussed above, suggest that in the case of peptides with small surface area, the main mechanism to compensate for a large hydrophobic mismatch is the tilt, while in the case of peptides with a large surface area, that cannot accommodate a too large tilt, the mismatch is mainly compensated for by an increase of the bilayer thickness around the peptide, as is clearly illustrated by the snapshot in Fig. 15 ($N_P = 43$).

The occurrence of peptide/protein tilting has also been confirmed experimentally. In fact, the results from very recent experimental investigations by solid state NMR spectroscopy [89] show that α-helical model peptides – of fixed hydrophobic length and with a hydrophobic leucine-alanine core, and tryptophan flanked ends – experience tilt when embedded in phospholipid bilayers of varying hydrophobic thickness (such that $d_P \geq d_L^o$, i.e. $\Delta d > 0$). It was found that the tilt angle increases by systematically increasing hydrophobic mismatch; however, the tilt dependence on hydrophobic mismatch was not as pronounced as one would have expected, given the degree of mismatch. This result brought the authors to conclude that the tilt of these peptides is ener-

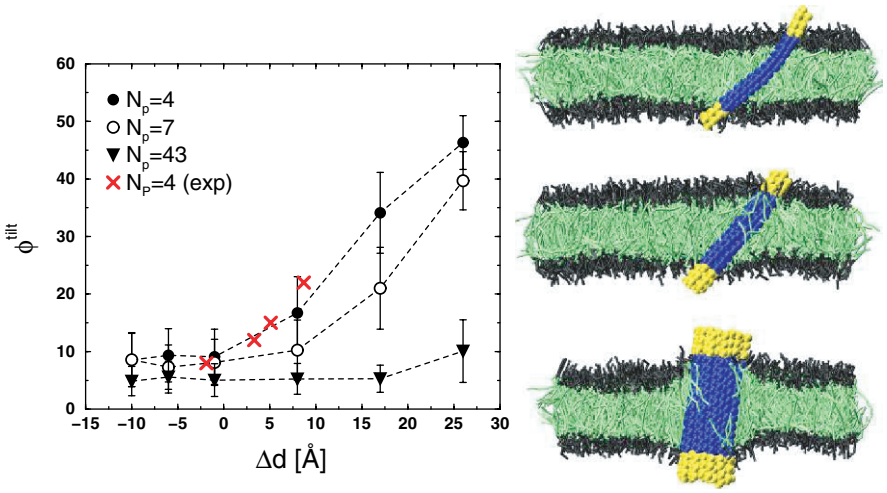


Fig. 15. peptide tilt-angle ϕ^{tilt} as function of mismatch, Δd. The data refer to a reduced temperature of $T^* = 0.7$ and to the three considered peptide sizes $N_P = 4$, 7 and 43. The dashed lines are only meant to be a guideline for the eye and the crosses are experimental data. Typical configurations of the systems resulting from the simulations are shown on the right. Starting from the top, the snapshots refer to peptides sizes $N_P = 4$, 7 and 43. In the three cases the peptide hydrophobic length is $\tilde{d}_P = 50$ Å, hence the hydrophobic mismatch is $\Delta d = 26$ Å

getically unfavorable, and to suggest that the (anchoring) effects by specific residues such as tryptophans are more dominant than mismatch effect. A large tilt is instead experienced by the M13 coat peptide peptide when embedded in phospholipid bilayer of varying hydrophobic thickness [90]. For values of mismatch of the same order of the one experienced by the synthetic peptides just mentioned [89], the degree of tilting experienced by M13 peptide is much higher. Our simulation data indicate a dependence of the peptide-tilt angle on mismatch in agreement with the experimental data just discussed. Incidentally, the results from our simulations suggest that, when a skinny peptide ($N_p = 4$) is subjected to a large positive mismatch ($d_P > d_L^o$), it might bend – besides to experience a tilt – as can be seen by looking at the snapshot shown on the top-right of Fig. 15. Also, as soon as the positive mismatch decreases, the bending disappears, although the peptide still tends to remain tilted.

5 Concluding Remarks

In this contribution we have shown that a simple, mesoscopic, representation of a phospholipid can give a surprisingly realistic description of the phase behavior of biological membranes. The simulations showed that different stable phases are obtained for a wide range of temperatures. We characterized the low

temperature phase as a gel phase, and we reproduced the main order/disorder phase transition from a gel to a liquid crystalline phase.

We also demonstrate that this model can be used to study the effect of alcohol on the structure of the membrane. Our model nicely reproduces the experimental phase diagram and the alcohol length dependence of the thickness of the interdigitated phase. Our simulations show that the interdigitated phase is stable at much lower alcohol concentrations in the membrane than suggested in the literature. This points at an alternative interpretation of the structure of the interdigitated phase.

This model has also been extended to lipid bilayers containing just one lipid species and an embedded protein. For which we have investigated the effect due to mismatch and protein size on the perturbation induced by the protein on the surrounding lipid bilayer.

References

1. C. Tanford (1978) The hydrophobic effect and the organization of living matter. *Science* **200**, pp. 1012–1018
2. C. Tanford (1980) *Hydrophobic Effect: Formation of Micelles and Biological Membranes.* Wiley, New York
3. E. Sackmann (1995) Biological membranes architecture and function. In *Structure and Dynamics of Membranes* (R. Lipowsky and E. Sackmann, eds.), (Amsterdam), Elsevier pp. 1–65
4. S. König and E. Sackmann (1996) Molecular and collective dynamics of lipid bilayers. *Current Opinion in Colloid and Interface Sci.* **1**, pp. 78–82
5. D. P. Tieleman, S. J. Marrink, and H. J. C. Berendsen (1997) A computer perspective of membranes: molecular dynamics studies of lipid bilayer systems. *Biochim. Biophys. Acta* **1331**, pp. 235–270
6. S. J. Marrink and D. P. Tieleman (2001) Molecular dynamics simulation of a lipid diamond cubic phase. *J. Am. Chem. Soc.* **123**, pp. 12383–12391
7. L. Shen, D. Bassolino, and T. Stouch (1997) Transmembrane helix structure, dynamics, and interactions: Multi-nanosecond molecular dynamics simulations. *Biophys. J.* **73**, pp. 3–20
8. K. Belohorcová, J. A. Davis, T. B. Woolf, and B. Roux (1997) Structure and dynamics of an amphiphilic peptide in a bilayer: a molecular dynamics study. *Biophys. J.* **73**, pp. 3039–3055
9. K. Belohorcová, J. Qian, and J. A. Davis (2000) Molecular dynamics and ^{2}h-nmr study of the influence of an amphiphilic peptide on membrane order and dynamics. *Biophys. J.* **79**, pp. 3201–3016
10. M. Ø. Jensen, E. Tajkhorshid, and K. Schulten (2001) The mechanism of glycerol conduction in aquaglyceroporins. *Structure* **9**, pp. 1083–1093
11. M. Ø. Jensen and O. G. Mouritsen (2004) Lipids do influence protein function – the hydrophobic matching hypothesis revised. *Biochim. Biophys. Acta: Biomembranes* vol. in press, pp. 205–226
12. T. Stouch and D. Bassolino (1996) Movement of small molecules in lipid bilayers: Molecular dynamics simulation studies. In *Biological Membranes* (J. Mertz and R. B. eds.), (Boston), pp. 255–277; Birkhauser

13. O. G. Mouritsen, M. M. Sperotto, J. Risbo, Z. Z. Zhang, and M. J. Zuckermann (1996) Computational approach to lipid-protein interactions in membranes. *Adv. Comp. Biol.* **2**, pp. 15–64
14. T. Gil, J. H. Ipsen, O. G. Mouritsen, M. C. Sabra, M. M. Sperotto, and M. J. Zuckermann (1998) Theoretical analysis of protein organization in lipid membranes. *Biochim. Biophys. Acta* **1376**, pp. 245–266
15. T. Gil and J. H. Ipsen (1997) Capillary condensation between disks in two dimensions. *Phys. Rev. E* **55**, pp. 1713–1721
16. S. May (2000) Theories on structural perturbations of lipid bilayers. *Curr. Opin. Colloid Interface Sci.* **5**, pp. 244–249
17. D. R. Fattal and A. Ben-Shaul (1993) A molecular model for lipid-protein interactions in membranes: the role of hydrophobic mismatch. *Biophys. J.* **65**, pp. 1795–1809
18. C. Nielsen, M. Goulian, and O. S. Andersen (1998) Energetics of inclusion-induced bilayer deformations. *Biophys. J.* **74**, pp. 1966–1983
19. A. G. Lee (2003) Lipid-protein interactions in biological membranes: a structural perspective. *Biochim. Biophys. Acta* **1612**, pp. 1–40
20. J. A. Killian (1998) Hydrophobic mismatch between proteins and lipids in membranes. *Biochim. Biophys. Acta* **1376**, pp. 401–416
21. R. Goetz and R. Lipowsky (1998) Computer simulations of bilayer membranes: Selfassembly and interfacial tension. *J. Chem. Phys.* **108**, pp. 7397–7409
22. J. C. Shelley, M. Y. Shelley, R. C. Reeder, S. Bandyopadhyay, P. B. Moore, and M. L. Klein (2001) Simulations of phospholipids using a coarse grain model. *J. Phys. Chem. B* **105**, pp. 9785–9792
23. J. C. Shelley, M. Y. Shelley, R. C. Reeder, S. Bandyopadhyay, and M. L. Klein (2001) A coarse grain model for phospholipid simulations. *J. Phys. Chem. B* **105**, pp. 4464–4470
24. R. D. Groot and K. L. Rabone (2001) Mesoscopic simulation of cell membrane damage, morphology change and rupture by nonionic surfactants. *Biophys. J.* **81**, pp. 725–736
25. S. J. Marrink, A. H. de Vries, and A. E. Mark (2004) Coarse grained model for semiquantitative lipid simulations. *J. Phys. Chem. B* **108**, pp. 750–760
26. M. Venturoli and B. Smit (1999) Simulating the self-assembly of model membranes. *Phys. Chem. Comm.* **2**, part. no. 10
27. M. Kranenburg, M. Venturoli, and B. Smit (2003) Molecular simulations of mesoscopic bilayer phases. *Phys. Rev. E* **67**, part. no. 060901
28. M. Kranenburg, M. Venturoli, and B. Smit (2003) Phase behavior and induced interdigitation in bilayers studied with dissipative particle dynamics. *J. Phys. Chem. B* **107**, pp. 11491–11501
29. M. Kranenburg and B. Smit (2004) Simulating the effect of alcohol on the structure of a membrane. *FEBS Lett.* **568**, no. 1–3, pp. 15–18, FEBS LETT
30. M. Kranenburg, J.-P. Nicolas, and B. Smit (2004) Comparison of mesoscopic phospholipid-water models. *Phys. Chem. Chem. Phys.* **6**, no. 16, pp. 4142–4151
31. M. Kranenburg, M. Vlaar, and B. Smit (2004) Simulating induced interdigitation in membranes. *Biophys. J.* **87**, pp. 1596–1605
32. M. Kranenburg, C. Laforge, and B. Smit (2004) Mesoscopic simulations of phase transitions in lipid bilayers. *Phys. Chem. Chem. Phys.* **6**, pp. 4531–4534
33. M. Venturoli, B. Smit, and M. M. Sperotto (2005) Simulation studies of protein-induced bilayer deformations, and lipid-induced protein tilting, on a mesoscopic model for lipid bilayers with embedded proteins. *Biophys. J.* **88**, pp. 1778–1798

34. M. Kranenburg and B. Smit (2005) Phase behavior of model lipid bilayers. *J. Phys. Chem. B* **109**, pp. 6553–6563
35. P. J. Hoogerbrugge and J. M. V. A. Koelman (1992) Simulating microscopic hydrodynamics phenomena with dissipative particle dynamics. *Europhys. Lett.* **19**, pp. 155–160
36. P. Español (2002) Dissipative particle dynamics revisited. *SIMU, Challenges in Molecular Simulations* **4**, pp. 59–77
37. P. B. Warren (1998) Dissipative particle dynamics. *Curr. Opinion Coll. Interface Sci.* **3**, pp. 620–624
38. J. C. Shillcock and R. Lipowsky (2002) Equilibrium structure and alteral stress distribution of amphiphilic bilayers from dissipative particle dynamics simulations. *J. Chem. Phys.* **117**, pp. 5048–5061
39. G. Ayton and G. A. Voth (2002) Bridging microscopic and mesoscopic simulations of lipid bilayers. *Biophys. J.* **83**, pp. 3357–3370
40. R. D. Groot and P. B. Warren (1997) Dissipative particle dynamics: bridging the gap between atomistic and mesoscopic simulation. *J. Chem. Phys.* **107**, pp. 4423–4435
41. P. Español and P. Warren (1995) Statistical mechanics of dissipative particle dynamics. *Europhys. Lett.* **30**, pp. 191–196
42. S. M. Willemsen, T. J. H. Vlugt, H. C. J. Hoefsloot, and B. Smit (1998) Combining dissipative particle dynamics and Monte Carlo techniques. *J. Comp. Phys.* **147**, pp. 507–517
43. J. S. Rowlinson and B. Widom (1982) *Molecular Theory of Capillarity.* Oxford: Clarendon Press
44. F. Jähnig (1996) What is the surface tension of a lipid membrane? *Biophys. J.* **71**, pp. 1348–1349.
45. R. Kwok and E. Evans (1981) Thermoelasticity of large lecithin bilayer vesicles. *Biophys. J.* **35**, pp. 637–652
46. S.-W. Chiu, M. Clark, S. Subramaniam, H. L. Scott, and E. Jakobsson (1995) Incorporation of surface tension into molecular dynamics simulation of an interface: a fluid phase lipid bilayer membrane. *Biophys. J.* **69**, pp. 1230–1245
47. R. Zhang, W. Sun, S. Traistram-Nagle, R. L. Headrick, R. M. Suter, and J. F. Nagle (1995) Critical fluctuations in membranes. *Phys. Rev. Lett.* **74**, pp. 2832–2835
48. S. E. Feller and R. W. Pastor (1996) On simulating lipid bilayers with an applied surface tension: periodic boundary conditions and undulations. *Biophys. J.* **71**, pp. 1350–1355
49. S. E. Feller and R. W. Pastor (1999) Constant surface tension simulations of lipid bilayers: the sensitivity of surface areas and compressibilities. *J. Chem. Phys.* **111**, pp. 1281–1287
50. S. J. Marrink, E. Lindahl, O. Edholm, and A. E. Mark (2001) Simulation of the spontaneous aggregation of phospholipids into bilayers. *J. Am. Chem. Soc.* **123**, pp. 8638–8639
51. S. E. Feller, Y. Zhang, and R. W. Pastor (1995) Computer simulation of liquid/liquid interfaces. ii. surface tension-area dependence of a bilayer and monolayer. *J. Chem. Phys.* **103**, pp. 10267–10276
52. D. P. Tieleman and H. J. C. Berendsen (1996) Molecular dynamics simulations of a fully hydrated dipalmitoylphosphatidylcholine bilayer with different macroscopic boundary conditions and parameters. *J. Chem. Phys.* **105**, pp. 4871–4880

53. S.-W. Chiu, E. Jakobsson, S. Subramaniam, and H. L. Scott (1999) Combined Monte Carlo and molecular dynamics simulation of fully hydrated dioleyl and palmitoyloleyl phosphatidylcholine lipid bilayers. *Biophys. J.* **77**, pp. 2462–2469
54. R. M. Venable, B. R. Brooks, and R. W. Pastor (2000) Molecular dynamics simulations of gel ($l_{\beta i}$) phase lipid bilayers in constant pressure and constant surface area ensemble. *J. Chem. Phys.* **112**, pp. 4822–4832
55. F. Sun (2002) Constant normal pressure, constant surface tension, and constant temperature molecular dynamics simulation of hydrated 1,2-dilignoceroylphosphatidylcholine monolayer. *Biophys. J.* **82**, pp. 2511–2519
56. R. Koynova and M. Caffrey (1998) Phases and phase transitions of the phosphatidylcholines. *Biochim. Biophys. Acta* **1376**, pp. 91–145
57. J. Katsaras and T. Gutberlet (2001) *Lipid Bilayers, Structure and Interactions* Berlin: Springer
58. J. L. Slater and C. H. Huang (1988) Interdigitated bilayer membranes. *Prog. Lipid. Res.* **27**, pp. 325–359
59. S. Furuike, V. G. Levadny, S. J. Li, and M. Yamazaki (1999) Low pH induces an interdigitated gel to bilayer gel phase transition in dihexadecylphosphatidylcholine membrane. *Biophys. J.* **77**, pp. 2015–2023
60. T. J. McIntosh, R. V. McDaniel, and S. A. Simon (1983) Induction of an interdigitated gel phase in fully hydrated phosphatidylcholine bilayers. *Biochim. Biophys. Acta* **731**, pp. 109–114
61. T. J. McIntosh, H. Lin, S. Li, and C.-H. Huang (2001) The effect of ethanol on the phase transition temperature and the phase structure of monounsaturated phosphatidylcholines. *Biochim. Biophys. Acta* **1510**, pp. 219–230
62. T. Adachi, H. Takahashi, K. Ohki, and I. Hatta (1995) Interdigitated structure of phospholipid-alcohol systems studied by X-ray diffraction. *Biophys. J.* **68**, pp. 1850–1855
63. T. Hata, H. Matsuki, and S. Kaneshina (2000) Effect of local anesthetics on the bilayer membrane of dipalmitoylphosphatidylcholine: interdigitation of lipid bilayer and vesicle-micelle transition. *Biophys. Chem.* **87**, pp. 25–36
64. R. N. A. H. Lewis, I. Winter, M. Kriechbaum, K. Lohner, and R. N. McElhaney (2001) Studies of the structure and organization of cationic lipid bilayer membranes: Calorimetric, spectroscopic, and X-ray diffraction studies of linear saturated P-O-ethyl phosphatidylcholines. *Biophys. J.* **80**, pp. 1329–1342
65. I. Winter, G. Pabst, M. Rappolt, and K. Lohner (2001) Refined structure of 1,2-diacyl-P-O-ethylphosphatidylcholine bilayer membranes. *Chem. Phys. Lip.* **112**, pp. 137–150
66. S. A. Simon and T. J. McIntosh (1984) Interdigitated hydrocarbon chain packing causes the biphasic transition behavior in lipid/alcohol suspensions. *Biochim. Biophys. Acta* **773**, pp. 169–172
67. Y. Misquitta and M. Caffrey (2001) Rational design of lipid molecular structure: a case study involving the c19:1c10 monoacylglycerol. *Biophys. J.* **81**, pp. 1047–1058
68. S. Tristram-Nagle, Y. Liu, J. Legleiter, and J. F. Nagle (2002) Structure of gel phase dmpc determined by X-ray diffraction. *Biophys. J.* **83**, pp. 3324–3335
69. R. S. Cantor (1998) The lateral pressure profiles in membranes: a physical mechanism of general anesthesia. *Toxicology Lett.* **101**, pp. 451–458
70. J. G. Petrov, E. E. Polymeropoulos, and Möhwald (2000) Langmuir monolayers with fluorinated groups in the hydrophilic head. 1. comparison of trifluoroethyl

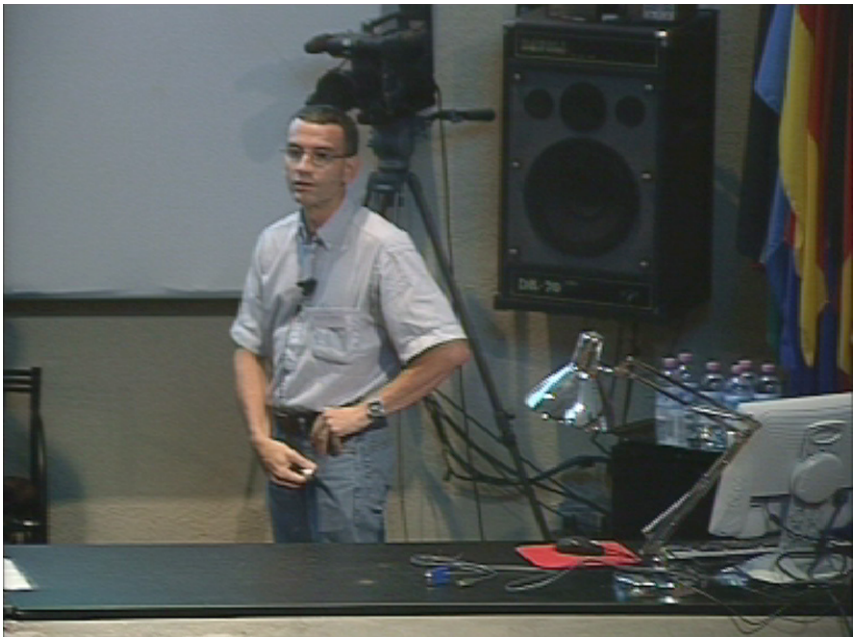
behenate and ethyl behenate monolayers: Molecular models, mechanical properties, stability. *Langmuir* **16**, pp. 6411–7420
71. E. S. Rowe (1983) Lipid chain length and temperature dependence of ethanol-phosphatidylcholine interactions. *Biochemistry* **22**, pp. 3299–3305
72. P. Nambi, E. S. Rowe, and T. J. McIntosh (1988) Studies of the ethanol-induced interdigitated gel phase in phosphatidylcholines using the fluorophore 1,6-diphenyl-1,3,5-hexatriene. *Biochemistry* **27**, pp. 9175–9182
73. K. Ohki, K. Tamura, and I. Hatta (1990) Ethanol induces interdigitated gel phase ($l_{\beta i}$) between lamellar gel phase ($l_{\beta'}$) and ripple phase ($p_{\beta'}$) in phosphatidylcholine membranes: a scanning density meter study. *Biochim. Biophys. Acta* **1028**, pp. 215–222
74. U. Vierl, L. Löbbecke, N. Nagel, and G. Cevc (1994) Solute effects on the colloidal and phase behavior of lipid bilayer membranes: ethanol-dipalmitoylphosphatidylcholine mixtures. *Biophys. J.* **67**, pp. 1067–1079
75. L. Löbbecke and G. Cevc (1995) Effects of short-chain alcohols on the phase behavior and interdigitation of phosphatidylcholine bilayer membranes. *Biochim. Biophys. Acta* **1237**, pp. 59–69
76. J. A. Veiro, P. Nambi, L. L. Herold, and E. S. Rowe (1987) Effect of *n*-alcohols and glycerol on the pretransition of dipalmitoylphosphatidylcholine. *Biochim. Biophys. Acta* **900**, pp. 230–238
77. E. S. Rowe and J. M. Campion (1994) Alcohol induction of interdigitation in distearoylphosphatidylcholine: fluorescence studies of alcohol chain length requirements. *Biophys. J.* **67**, pp. 1888–1895
78. S. A. Simon, W. L. Stone, and P. Busto-Latorre (1977) A thermodynamic study of the partition of n-hexane into phosphatidylcholine and phosphatidylcholine-cholesterol bilayers. *Biochim. Biophys. Acta* **468**, pp. 378–388
79. C. Trandum and P. Westh (2000) Alcohol-liposome interactions: a thermodynamic investigation. *Prog. Anesth. Mech.* vol. **6**
80. C. Trandum, P. Westh, K. Jörgensen, and O. G. Mouritsen (1999) A calorimetric investigation of the interaction of short chain alcohols with unilamellar DMPC liposomes. *J. Phys. Chem. B* **103**, pp. 4751–4756
81. P. Westh and C. Trandum (1999) Thermodynamics of alcohol-lipid bilayer interactions: application of a binding model. *Biochim. Biophys. Acta* **1421**, pp. 261–272
82. P. Westh, C. Trandum, and Y. Koga (2001) Binding of small alcohols to a lipid bilayer membrane: does the partitioning coefficient express the net affinity?. *Biophys. Chem.* **89**, pp. 53–63
83. T. Adalsteinsson and H. Yu (2000) Lipid lateral diffusion in multi-bilayers, and in monolayers at the air/water and heptane/water interfaces. *Langmuir* **16**, pp. 9410–9413
84. L. Pauling (1960) *The nature of the chemical bond.* New York: Cornell University Press
85. O. G. Mouritsen and M. Blom (1984) Matress model of lipid-protein interactions in membranes. *Biophys. J.* **46**, pp. 141–153
86. E. Sackmann (1984) Physical basis of trigger processes and membrane structure. In *Biological Membranes vol. 5* (D. Chapman ed.), pp. 105–143, Academic Press, London
87. O. G. Mouritsen and M. M. Sperotto (1993) Thermodynamics of lipid-protein interactions in lipid membranes. In *Thermodynamics of Membrane Receptors and*

Channels (M. Jackson, ed.), (Boca Raton, Florida), pp. 127–181, CRC Press, Inc.
88. F. Dumas, M. C. Lebrum, and J.-F. Tocanne (1999) Is the protein/lipid hydrophobic matching principle relevant to membrane organization and functions? *FEBS Lett.* **485**, pp. 271–277
89. E. Strandberg, S. Özdirekcan, D. T. S. Rijkers, P. C. A. van der Wel, R. E. Koeppe II, R. M. J. Liskamp, and J. A. Killian (2004) Tilt angles of transmembrane model peptides in oriented and non-oriented lipid bilayers as determined by ^{2}H solid state NMR. *Biophys. J.* **86**, pp. 3709–3721
90. R. B. M. Koehorst, R. B. Spruijt, F. J. Vergeldt, and M. A. Hemminga (2004) Lipid bilayer topology of the transmembrane α-helix of M13 major coat protein and bilayer polarity profile by site-directed fluorescence spectroscopy. *Biophys. J.* **87**, pp. 1445–1455

Microscopic Elasticity of Complex Systems

J.-L. Barrat

Laboratoire de Physique de la Matière Condensée et Nanostructures. Université Claude Bernard Lyon I and CNRS, 6 rue Ampère, 69622 Villeurbanne Cedex France
barrat@lpmcn.univ-lyon1.fr



Jean-Louis Barrat

J.-L. Barrat: *Microscopic Elasticity of Complex Systems*, Lect. Notes Phys. **704**, 287–307 (2006)
DOI 10.1007/3-540-35284-8_12
© Springer-Verlag Berlin Heidelberg 2006

1 Introduction 289

2 Some Definitions 289

3 Finite Temperature Elastic Constants: Born and Fluctuation Terms 291

4 Amorphous Systems at Zero Temperature: Nonaffine Deformation 292

5 Numerical Results 294

6 Polymeric Systems: Stresses and Self Consistent Field Theory 299

A Expression for the Stress Tensor in SCFT 304

References 305

1 Introduction

Elasticity has the reputation of being a rather boring and un-physical subject. In fact, dealing with second (stress-strain) or fourth (elastic constants) rank tensor guarantees that the notations are in general rather heavy, and that the underlying physics is not easily captured.[1] The elastic stress-strain behavior is, however, a very basic property of all solid materials, and one that is rather easy to obtain experimentally. Hence simulation methods for solid systems (hard or soft) should in general consider the obtention of these elastic properties. The determination of the response of a system to an imposed external stress is an important topic, which was first addressed in the simulation community through the pioneering work of Parrinello and Rahman [2,3]. While the main issue addressed by the Parrinello-Ray-Rahman method was that of allotropic transformations in crystalline solids, the current interest in nanostructured materials opens a number of new questions. What are the appropriate measures of stress and strains at small scale? Can one scale down the constitutive laws of macroscopic elasticity, and if yes to what scale? Are the elastic constants of nanometric solids identical to those of the bulk? What about the elastic/plastic transition at small scales? Many of these questions are still unanswered, and the object of active studies. In the following, I will describe recent studies that investigate some of these questions. After briefly recalling the appropriate definitions, I will particularly concentrate on the case of amorphous systems, in which some surprises arise due to the non-affine character of deformations at small scales. In the last section, I will describe how the Parrinello–Rahman scheme can be extended to systems in which the particle representation has been replaced by a field representation, as is common in polymer or amphiphilic systems that self organize at a mesoscopic scale.

2 Some Definitions

A convenient way to study elasticity with a microscopic viewpoint is to think of systems contained in periodic cells with variable shape. This approach, which was initiated by Parrinello and Rahman, is useful from a practical, but also conceptual viewpoint. In a system with periodic boundary conditions, the simulation cell can be defined by three (in general nonorthogonal) independent vectors $\mathbf{h}_1, \mathbf{h}_2, \mathbf{h}_3$ forming the sides of the parallelepiped cell. The Cartesian coordinates of these vectors can be used to construct a 3×3 matrix $\underline{h}$ defined by $\underline{h} = (\mathbf{h}_1, \mathbf{h}_2, \mathbf{h}_3)$. The Cartesian coordinates of any point $\mathbf{R}$ in the cell can be expressed as

$$\mathbf{R} = \underline{h}\mathbf{X} \tag{1}$$

[1] To appreciate the very nice physical content of elementary elasticity theory, the reader is referred to the "Feynman lectures on Physics" or in a more formal style to the first chapters of the Landau and Lifschitz textbook [1]

where $\mathbf{X}$ is a rescaled vector whose components lie in $[0,1]$. Integrals on $\mathbf{R}$ can be converted into integrals over $\mathbf{X}$ by using a scaling factor $\det \underline{h}$, which represents the volume of the cell, V. In the case of a particle or monomer number density, for example, one can write

$$\rho(\mathbf{R}) = \rho(\mathbf{X})(\det \underline{h})^{-1} \tag{2}$$

The metric tensor $\underline{G}$ is constructed from $\underline{h}$ as

$$\underline{G} = \underline{h}^T \underline{h} \tag{3}$$

where $\underline{h}^T$ is the transpose of $\underline{h}$. $\underline{G}$ is used in transforming dot products from the original Cartesian to rescaled coordinates, according to

$$\mathbf{R} \cdot \mathbf{R}' = \mathbf{X} \cdot \underline{G} \cdot \mathbf{X}' = X_\alpha G_{\alpha\beta} X'_\beta \tag{4}$$

where here and in the following summation over repeated indexes is implicit.

Elasticity theory describes the deformation of any configuration from a reference configuration in terms of a strain tensor. This tensor is constructed by relating the vector connecting two points in the deformed configuration to the corresponding displacement of the same points in the reference configuration. If the reference configuration of the simulation box is denoted by $\underline{h}_0$, the displacement is $\mathbf{u} = \mathbf{R} - \mathbf{R}_0 = (\underline{h}\underline{h}_0^{-1} - 1)\mathbf{R}_0$, and the strain is given by [2,3]

$$\underline{\epsilon} = \frac{1}{2}\left[(\underline{h}^T{}_0)^{-1}\underline{h}^T\underline{h}(\underline{h}_0)^{-1} - \underline{1}\right] = \frac{1}{2}\left[(\underline{h}^T{}_0)^{-1}\underline{G}(\underline{h}_0)^{-1} - \underline{1}\right] \tag{5}$$

where $\underline{1}$ denotes the unit tensor. It is important to note that this expression, usually known as the *Lagrangian strain tensor* is not limited to small deformations [1]. Usually, the reference configuration $\underline{h}_0$ will be defined as a state of the system under zero applied external stress. If one starts with a cubic cell, $\underline{h}_0$ is the identity matrix and the relation between $\underline{\epsilon}$ and $\underline{G}$ simplifies. The thermodynamic variable conjugate to this strain tensor, in the sense that the elementary work done on the system can be written in the form

$$\delta W = V_0 \mathrm{Tr}(\underline{t}\ \delta\underline{\epsilon})\ , \tag{6}$$

is the thermodynamic *tension* tensor $\underline{t}$ [4], also known as Piola-Kirchhoff second stress tensor. $V_0 \equiv \det \underline{h}_0$ denotes the volume of the system in the reference configuration. This thermodynamic tension tensor can be related to the more usual *Cauchy stress tensor* $\underline{\sigma}$ through

$$\underline{\sigma} = \frac{V_0}{V}\underline{h}\,(\underline{h}_0)^{-1}\underline{t}\,(\underline{h}^T{}_0)^{-1}\underline{h}^T \tag{7}$$

The tension is the derivative of the free energy with respect to the strain, which is calculated from the reference configuration. The Cauchy stress, on the other hand, is the derivative of the free energy with respect to an incremental strain

taken with respect to the actual configuration. This Cauchy stress tensor is the one that enters momentum conservation and whose expression is given by the usual Irving-Kirkwood formula for pairwise additive potentials (see below). The difference between these two quantities can be understood, qualitatively, from the fact that the strain is not an additive quantity, as can be seen from the existence of the nonlinear term in (5). While the Cauchy stress has a mechanical meaning in terms of forces within the sample, the thermodynamic tension is a purely thermodynamic quantity, and does not in general have a simple mechanical interpretation.

Fortunately, in the limit of small deformations which I will concentrate on, the differences between these various expressions of stress tensors can be forgotten. This is not the case, however, for large deformations, where these differences result in a whole variety of stress/strain relations and associated elastic constants. This is especially important when dealing with solids under high pressure, where for example one has to be careful as to which of these elastic constants is used to compute e.g. sound velocities. I will refer the interested reader to the reference publication of Klein and Baron [5] for an in depth discussion of these subtleties.

3 Finite Temperature Elastic Constants: Born and Fluctuation Terms

The elastic constants for a material made of particles interacting through a pair potential $\phi(r)$ (I'll keep this simplifying assumption in the following) can be determined from simulations using an approach presented by Hoover and coworkers [6]. These authors start from the explicit expression of the free energy in terms of a configuration integral

$$\exp(-\beta F) = V^N \int d\mathbf{X}_1 d\mathbf{X}_2 .. d\mathbf{X}_N \exp(-\beta H(\{\underline{h}\mathbf{X}_i\})) \tag{8}$$

where $H(\{\mathbf{R}_i\}) = \sum_{ij} \phi(\mathbf{R}_i - \mathbf{R}_j)$ is the total interaction energy.

The derivative with respect to strain is taken using

$$dF = \mathrm{Tr}(\frac{\partial F}{\partial \underline{G}} d\underline{G}) = 2\mathrm{Tr}\left(\underline{h}_0 \frac{\partial F}{\partial \underline{G}} \underline{h}^T{}_0 d\underline{\epsilon}\right) \tag{9}$$

which gives for the thermodynamic tension matrix

$$\begin{aligned} V_0 t_{\alpha\beta} &= N k_B T \underline{h}_{0,\alpha\gamma} \underline{G}^{-1}_{\gamma\delta} \underline{h}^T{}_{0,\delta\beta} + \underline{h}_{0,\alpha\gamma} \left\langle \sum_{ij} X_{ij,\gamma} X_{ij,\delta} \frac{\phi'(R_{ij})}{R_{ij}} \right\rangle \underline{h}_{0,\delta\beta} \\ &= N k_B T \underline{h}_{0,\alpha\gamma} \underline{G}^{-1}_{\gamma\delta} \underline{h}^T{}_{0,\delta\beta} + \langle \hat{\underline{T}}_{\alpha\beta} \rangle \end{aligned} \tag{10}$$

where $X_{ij} = X_i - X_j$, $\sum_{ij}$ is the summation over all distinct pairs of particles, and the pair potential ϕ is assumed to depend only on the particle separation

R_{ij}. The brackets $\langle\rangle$ denote a thermal average. This obviously reduces to the usual Kirkwood formula for small deformations ($\underline{h}_0 = \underline{h}$). Note that the first term arises here from the volume factor V^N in the configuration integral, but could also be obtained by introducing the momenta and the kinetic energy contribution in the partition function. The last term defines the potential energy contribution to the microscopic stress tensor, denoted by $\underline{\hat{T}}$. Carrying out one more derivation with respect to strain, one obtains the elastic constants in the limit of zero strain (more general expressions for arbitrary strain can be found in refs [7,8]):

$$C_{\alpha\beta\gamma\delta} = \frac{\partial t_{\alpha\beta}}{\partial \epsilon_{\gamma\delta}} = 2Nk_BT(\delta_{\alpha\gamma}\delta_{\beta\delta} + \delta_{\alpha\delta}\delta_{\beta\gamma}) - \frac{V_0}{k_BT}\left[\langle\hat{T}_{\alpha\beta}\hat{T}_{\gamma\delta}\rangle - \langle\hat{T}_{\alpha\beta}\rangle\langle\hat{T}_{\gamma\delta}\rangle\right] + C^{Born}_{\alpha\beta\gamma\delta} \tag{11}$$

Where the last (Born) term is written in terms of potential energy functions

$$C^{Born}_{\alpha\beta\gamma\delta} = \frac{1}{V_0}\left\langle \sum_{ij} R_{ij,\alpha}R_{ij,\beta}R_{ij,\gamma}R_{ij,\delta}\left(\frac{\phi"(R_{ij})}{R_{ij}^2} - \frac{\phi'(R_{ij})}{R_{ij}^3}\right)\right\rangle \tag{12}$$

with $\mathbf{R}_{ij} = \mathbf{R}_i - \mathbf{R}_j$. The term in square brackets in (11) is called the fluctuation term, and is generally expected to be a correction to the main Born term at finite temperature. We will see below that this term, even in the low temperature limit, remains an essential contribution to the elastic properties of disordered systems. Formulae that generalize the equations above to local stresses and elastic constants can be found in references [9,10].

Note that a different approach must be used for hard core potentials. In that case, one possibility is to study strain fluctuations either in a cell of variable cell [11], or to directly study the stress-strain relation in a deformed cell [12].

4 Amorphous Systems at Zero Temperature: Nonaffine Deformation

The simulation of amorphous systems at low temperatures is interesting from the point of view of the physics of glasses, but also because these models can serve as very elementary examples of "athermal" systems such as granular piles or foams. The elastic/plastic response of such complex systems is not well understood, and is currently the subject of many experimental studies [13–16].

A naive approach to the calculation of elastic properties in such systems would consist in taking the second derivative of the potential energy $H = \sum_{i<j}\phi(R_{ij})$ with respect to strain. Such an approach is easily shown to yield elastic constants that correspond to the Born expression, without the thermal average brackets. Although such an approach seems natural –

the "fluctuation" term could be ignored at zero temperature –, it proves in fact completely incorrect for disordered systems, or even for crystals with a complex unit cell.

The essential point is that the derivatives have to be taken not at constant X, but rather keeping the force on each atom equal to zero in the deformed configuration. In other words, one has to allow for relaxation of the deformed configuration before computing the energy and stresses. It was shown by Lutsko [7] (see also the recent work by Lematre and Maloney, [8]) that this relaxation gives a contribution to elasticity which is identical to the zero temperature limit of the fluctuation term. The corresponding proof can be briefly summarized as follows.

The elastic constant is written as

$$C_{\alpha\beta\gamma\delta} = \frac{\partial t_{\alpha\beta}}{\partial \epsilon_{\gamma\delta}}|_{\mathbf{F}_i=0} \tag{13}$$

${}_F F_i = 0$ indicates the constraint that forces on the particles must remain zero during the deformation. The variables in the problem are the reduced coordinates, $\mathbf{X}_i$, and the strain $\underline{\epsilon}$. The force $\mathbf{F}_i$ is a function of these variables, so that the constrained derivative above can be written as (for simplicity, we drop in this formula and the following the Greek indexes for Cartesian coordinates):

$$\underline{C} = \frac{\partial t}{\partial \epsilon}|_{\mathbf{X}_i} - \frac{\partial t}{\partial \mathbf{X}_i}\left(\frac{\partial \mathbf{F}_j}{\partial \mathbf{X}_i}\right)^{-1}\frac{\partial \mathbf{F}_j}{\partial \epsilon} \tag{14}$$

where terms such as $\left(\frac{\partial \mathbf{F}_j}{\partial \mathbf{X}_i}\right)$ have to be understood in a matrix sense. In fact, $D_{ij} = \left(\frac{\partial \mathbf{F}_j}{\partial \mathbf{X}_i}\right)$ is nothing but the *dynamical matrix* of second derivatives of the potential energy with respect to atomic positions. Finally, using the definition of the force $\mathbf{F}_j$ and of the tension $\underline{t}$ as derivatives of the potential energy, 14 can be rewritten in the more symmetric form

$$\underline{C} = \underline{C}^{Born} - \frac{\partial t}{\partial X_i}\left(\frac{\partial F_j}{\partial X_i}\right)^{-1}\frac{\partial t}{\partial X_j} \tag{15}$$

The direct evaluation of the second term in this equation is not straightforward, hence the actual procedure to obtain the zero temperature elastic constants generally consists in carrying out explicitly an affine deformation of all coordinates, then letting the atomic positions relax (using e.g. conjugate gradient minimization) [27] to the nearest energy minimum. Equation (15) however can be used to show that the resulting elastic constants are identical to those obtained at a finite, low temperature using (11). The proof goes simply by expanding both the stress and energy in terms of the atomic displacements in the unstrained reference configuration

$$H = H_0 + \frac{1}{2}D_{ij}\delta X_i \delta X_j \quad ; \quad t = t_0 + \frac{\partial t}{\partial X_i}\delta X_i \tag{16}$$

Performing the resulting gaussian integrals, it is seen that the "fluctuation" term in (11) and the "relaxation" term in (15) are identical in the limit of zero temperature.

5 Numerical Results

We now address the importance, qualitative and quantitative, of this "relaxation-fluctuation" contribution. Quantitatively, the importance of this contribution can be estimated from Fig. 1. In this figure, the Lamé coefficients of a three dimensional, amorphous, Lennard–Jones system (see [17]) at zero temperature are computed using the Born approximation and the exact formula, (11). It is seen that the relaxation term can account for as much as 50% of the absolute value of elastic constants. Although this fraction may obviously be system dependent, the situation is very different compared to simple crystals (with one atom per unit cell, e.g. FCC in the Lennard–Jones system) in which the elastic constants are exactly given by the Born term (see e.g. [18] for a comparison between amorphous systems and simple crystal structures). The relaxation contribution tends to lower the shear modulus μ, and to increase the coefficient λ. Remarkably, the bulk modulus $K = \lambda + 2\mu/d \approx 57$ ($d = 3$ is the dimensionality of space) would be correctly predicted by the Born calculation.

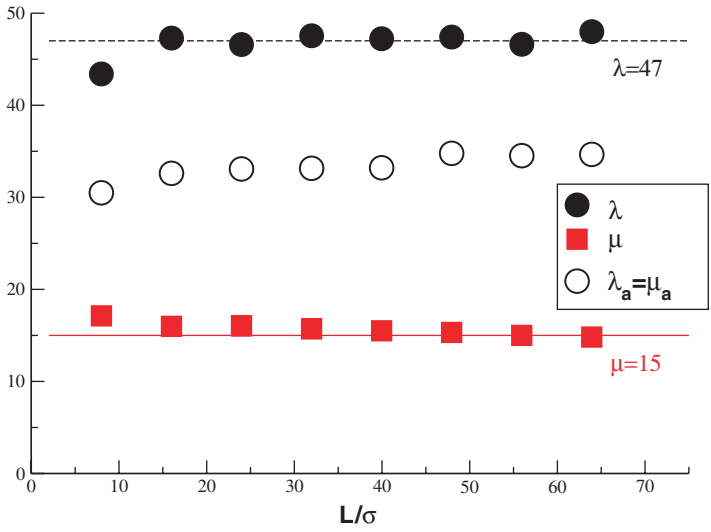


Fig. 1. Lamé coefficients λ (*spheres*) and μ (*squares*) *vs.* system size L, for a simple polydisperse Lennard–Jones "glass". *Full* symbols correspond to a direct measurement using Hooke's law with relaxation, *open* symbols correspond to the Born approximation). The effect of system size is weak. For large boxes we get $\mu \approx 15$ and $\lambda \approx 47$

As discussed above, the Born formulae would be exact at zero temperature if the global deformation was equivalent to an affine deformation of atomic coordinates at all scales, i.e. a mere rescaling of $\underline{h}$ at fixed values of $\{X_i\}$. The failure of the Born calculation can therefore be traced back to the existence of a *non-affine deformation field*, which stores part of the elastic deformation energy. This field is defined by substraction from the actual displacement of the atoms (after relaxation) the displacement that would be obtained in the affine hypothesis. The existence of this non-affine deformation field was pointed out in several recent publications [9,19–21]. In [21,22], it was in particular shown that this non-affine contribution is correlated over large distances, and is organized in vortex like structures (i.e. is mostly rotational in nature). These properties are illustrated in Figs. 2 and 3, for a simple Lennard–Jones two dimensional system.

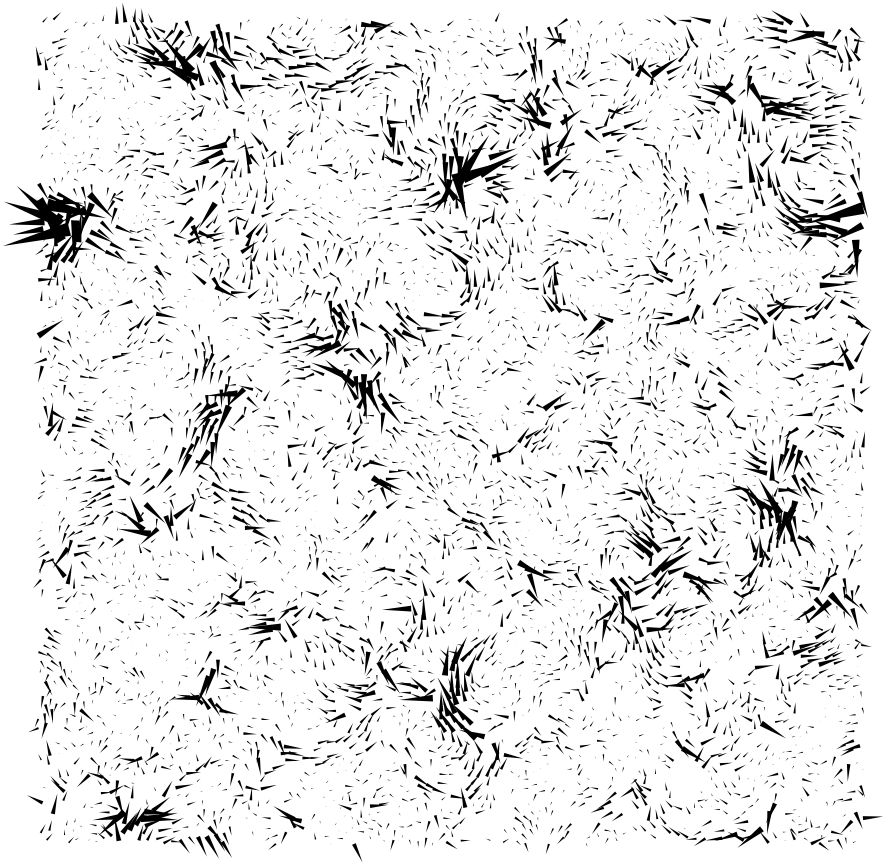


Fig. 2. Snapshot of the nonaffine displacement field in a 2d Lennard–Jones amorphous system undergoing uniaxial extension. Note the large scale, vortex like structures. The sample contains about 20000 particles

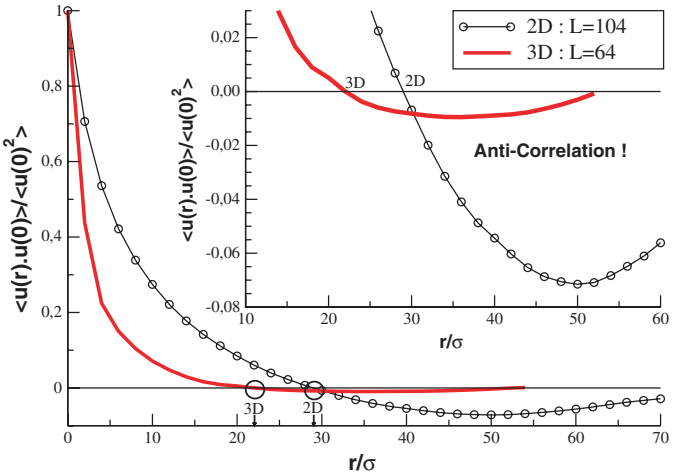


Fig. 3. Correlation function $C(r) = \langle \mathbf{u}^{NA}(\boldsymbol{r}) \cdot \mathbf{u}^{NA}(0) \rangle$ of the nonaffine displacement field $\mathbf{u}^{NA}(\boldsymbol{r})$ in an amorphous sample undergoing a simple uniaxial extension. Both the 2d and 3d case show correlations that extend over scales of typically 20-30 particle sizes. A negative tail can be associated with "vortex like" structures that reflect the essentially rotational character of the non affine field

A slightly different, more local and general, definition of the nonaffine displacement field (or "displacement fluctuation") was proposed in [9]. In this reference, the nonaffine field is defined by removing from the actual displacement a local displacement field built that is obtained using a coarse-graining procedure. This allows in principle to deal with situations in which the displacement field has a complex structure. In the case of simple shear considered here, our definition should be sufficient. Reference [9] also demonstrates that, even when the displacements are not locally affine, there exists a local linear relation between the stress and strain fields, at sufficiently large scales (resolution). These fields are evaluated at the same position with a chosen resolution. The derivation does assume that the displacement fluctuations are uncorrelated over sufficiently large scales, i.e. it will be valid at scales larger than the one discussed here.

A quick study of a simple one dimensional model is useful to understand the importance of the nonaffine displacement field. Let us consider a chain of N atoms, connected by springs k_i, submitted to a force F. The extension of spring i (linking site i to $i+1$) is $\delta_i = F/k_i$. One can therefore write the displacement of atom p

$$u_p = F \times \sum_{i=1}^{p} k_i^{-1} \tag{17}$$

The affine displacement is just $u_p^{aff} = (p/N) \times F \times \sum_{i=1}^{N} k_i^{-1} = pF\langle k^{-1} \rangle$, where the $\langle\rangle$ refer to an average over the distribution of elastic constants and

the large N limit has been taken to compute the affine displacement. As a result we have for the nonaffine displacement of atom p

$$u_p^{NA} = u_p - u_p^{aff} = F \times \sum_{i=1}^{P} (k_i^{-1} - \langle k^{-1} \rangle) \tag{18}$$

which shows that in this simple 1d situation the mean squared value $\langle \langle u_p^{NA} \rangle^2 \rangle \propto p \langle (\delta k^{-1})^2 \rangle$ is increasing linearly with p, and proportional to the variance of $1/k_i$ (see also the discussion by DiDonna and Lubensky, cond-mat/0506456).

The existence and nature of the length scale over which the non affine field is correlated is still a matter of debate.[2] Clearly, the correlation length ξ is a lower limit for the applicability of continuous elasticity theory. This limit manifests itself in several different ways. If one considers the vibrations of a system of size L, these vibrations will be properly described by the classical elasticity theory only if the corresponding wavelength is larger than ξ. If one considers the response to a point force, this response will be correctly described by the continuum theory only beyond the length scale ξ. More precisely, it was found that the *average* response is described by continuum theory essentially down to atomic size, but that the *fluctuations* (from sample to sample) around this average are dominant below ξ [22].

Finally, it is very likely that the existence of this length scale is related to a prominent feature of many disordered systems, the so called "boson peak". This feature actually corresponds to an excess (as compared to the standard Debye prediction, $g(\omega) \propto \omega^{d-1}$ in d dimensions) in the vibrational density of states $g(\omega)$ of many amorphous systems. This excess shows up as a peak in a plot of $g(\omega)/\omega^{d-1}$ vs ω, that usually lies in the THz range. In terms of length scales, we found that this peak typically corresponds to wavelengths of the order of magnitude of ξ (see Fig. 4). A simple description [23] is therefore to assume that waves around this wavelength are scattered by inhomogeneities, and see their frequencies shifted to higher values. Pressure studies show that the boson peak is shifted to higher frequencies under pressure, consistent with a shift to smaller values for ξ obtained in simulations. Another very interesting evidence for the existence of mesoscale inhomogeneities was recently provided by Masciovecchio and coworkers [24], by studying Brillouin spectra in the ultraviolet range. The width of the Brillouin peak shows a marked change for

[2] In a recent preprint (cond-mat/0506456, "Nonaffine correlations in Random Elastic Media"), DiDonna and Lubensky argued that the nonaffine field has logarithmic (in 2d) or $1/r$ (in 3d) correlations, and hence no characteristic length scale. That such singular behaviour is possible is already illustrated in the simple 1d example above. Such a behaviour, however, is not evident in our numerical results. Their calculation, based on the fact that elastic propagator have $1/k^2$ behaviour in Fourier space, is perturbative in the disorder strength, and it could be that we are investigating a "strong disorder" limit. In any case, further investigations are needed to assess the actual existence of such long range correlations

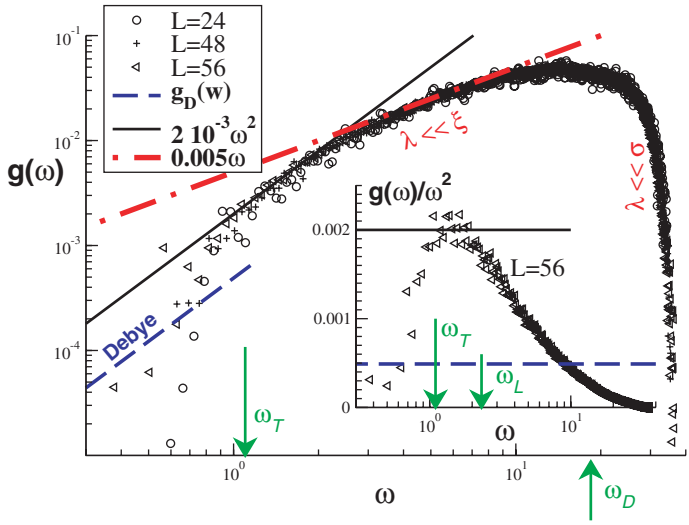


Fig. 4. Vibrational density of states in a 3d Lennard–Jones amorphous system. The "boson peak" is apparent as the deviations from the Debye prediction in the ratio $g(\omega)/\omega^2$. This peak is observed at frequencies of order $\omega_{L,T} = 2\pi c_{L,T}/\xi$ where $c_{L,T}$ is the speed of longitudinal or transverse sound

wavelength between 50 and 80 nm, indicative of scattering by elastic inhomogeneities.

A very interesting question is whether this characteristic correlation length for elastic inhomogeneities, which can reach rather large values compared to atomic sizes, is somehow associated with a "critical" phenomenon. An idea that was recently suggested by Nagel and co-workers [25] is that this correlation length should diverge at the so-called "jamming" transition in purely athermal systems. The jamming density is defined, in a system with purely repulsive interactions at zero temperature, by the density at which the system will start to exhibit mechanical rigidity. Below the jamming density ϕ_c, an infinitesimal temperature results in diffusion, while systems above ϕ_c remain in a frozen state on macroscopic time scales. Based on general arguments concerning isostaticity of the packing at ϕ_c, Nagel and coworkers [25] suggested the existence of a correlation length associated with soft modes, that diverges at the transition. Although the arguments are in principle valid only for contact type interactions, it would be quite interesting to follow the evolution of ξ for a system *with attractive interactions, but under tension*, expecting perhaps a divergence close to the rupture threshold.

Finally, let us mention that a different way of studying local elastic properties was proposed by de Pablo and coworkers through the study of local elastic moduli, which can be defined by using the definition 11 to a small, finite box [10]. Depending on the scale at which they are measured, these moduli can take negative values. Such regions would be unstable, if they were

not immersed in a matrix of " normal" regions. The size over which local elastic constants are found to be negative is small (typically 3 particle sizes), and could probably be considered as a first coarse-graining scale for using classical methods for disordered systems [26].

6 Polymeric Systems: Stresses and Self Consistent Field Theory

Self consistent field theory (SCFT) is a powerful approach to the determination of phase equilibria in polymer systems with complex architectures. The theory directly deals with density fields rather than particles, and minimizes a mean field like free energy. The method is particularly suitable for polymers, in which interactions on length scales comparable to the chain size are effectively very soft. Here, I will show how a stress control (rather than volume control) can be introduced in such a theory. The goal of such an approach is, ultimately, to be able to explore morphological phase transformations in soft solids without any a priori assumption on the symmetry of the simulation cell, letting the system spontaneously equilibrate under zero stress.

The SCFT method is well known and has been described in many publications (see e.g. [28, 29]). Here I will only describe briefly the main steps and the way a constant stress method can be introduced in the simulation. This option allows one to obtain relaxed configurations at zero imposed stress easily, or to study the effect of anisotropic tension on phase behavior.

As a representative example, I consider a model for an incompressible AB diblock copolymer melt [29]. The melt consists of n identical diblock copolymer chains composed of monomer species A and B and is contained in a volume V. Each of the chains has a total of N statistical segments; a fraction f of these segments are type A and constitute the A block of each macromolecule. For simplicity, the volume occupied by each segment, v_0, and the statistical segment length, b, are assumed to be the same for the A and B type segments. The Hamiltonian for this system can be written

$$H = \sum_{i=1}^{n} \frac{k_B T}{4R_{g0}^2} \int_0^1 ds \left(\frac{d\mathbf{R}_i(s)}{ds} \right)^2 + v_0 \chi_{AB} k_B T \int d\mathbf{r}\, \hat{\rho}_A(\mathbf{r}) \hat{\rho}_B(\mathbf{r}) \qquad (19)$$

where $\mathbf{R}_i(s)$ with $s \in [0, 1]$ is a space curve describing the conformation of the ith copolymer and $R_{g0}^2 = b^2 N/6$ is the radius of gyration of an ideal chain of N statistical segments. Interactions between dissimilar segments A and B are described by the Flory parameter χ_{AB}. The densities $\hat{\rho}_{A,B}(\mathbf{r})$ are microscopic segment density fields defined by

$$\hat{\rho}_A(\mathbf{r}) = N \sum_{i=1}^{n} \int_0^f ds\, \delta(\mathbf{r} - \mathbf{R}_i(s)) \qquad (20)$$

and

$$\hat{\rho}_B(\mathbf{r}) = N \sum_{i=1}^{n} \int_f^1 ds\ \delta(\mathbf{r} - \mathbf{R}_i(s)) \tag{21}$$

A local incompressibility constraint $\hat{\rho}_A(\mathbf{r}) + \hat{\rho}_B(\mathbf{r}) = \rho_0$ is imposed in this standard copolymer melt model for all points $\mathbf{r}$ in the simulation domain. The total segment density ρ_0 can evidently be expressed as $\rho_0 = nN/V = 1/v_0$. Using the rescaled coordinates $\mathbf{X}(s)$ (taken in $[0,1]^3$), the generalized partition function that has to be sampled *for a fixed value of the thermodynamic tension* $\underline{t}$ reads

$$\begin{aligned} Z = & \int d(\underline{h})(\det \underline{h})^{nN} \delta(\det \underline{h} - V_0)\ \exp(-\beta V_0 \underline{t}\ :\underline{\epsilon}) \\ & \times \prod_{i=1}^{n} \int \mathcal{D}\mathbf{X}_i(s) \exp(-\beta H) \prod_{\mathbf{x}} \delta(\hat{\rho}_A(\mathbf{x}) + \hat{\rho}_B(\mathbf{x}) - nN) \end{aligned} \tag{22}$$

($\underline{t}\ :\underline{\epsilon}$ is the contraction $t_{\alpha\beta}\epsilon\alpha\beta$). The final factor in the above expression imposes the constraint of local incompressibility. Moreover, incompressibility implies globally that the cell volume remains fixed at its initial value, i.e. $\det \underline{h} = V = V_0 = \det \underline{h}_0$. This is enforced by the delta function in the first line above. Hence all shape transformations should be volume preserving. The practical implementation of this constraint will be discussed below.

Hubbard–Stratonovich transformations are used to convert the particle-based partition function (22) into a field theory [28]. These can be carried out straightforwardly on the polymer partition function for a given cell shape $\underline{h}$, $Z(\underline{h})$, with the result

$$\begin{aligned} Z(\underline{h}) &\equiv \prod_{i=1}^{n} \int \mathcal{D}\mathbf{X}_i(s) \exp(-\beta H) \prod_{\mathbf{x}} \delta(\hat{\rho}_A(\mathbf{x}) + \hat{\rho}_B(\mathbf{x}) - nN) \\ &= \int \mathcal{D}w\ \exp(n \ln Q[w, \underline{h}] - E[w]) \end{aligned} \tag{23}$$

where $Q[w, \underline{h}]$ is the partition function of a single copolymer chain experiencing a chemical potential field $w(\mathbf{x}, s)$, $\int \mathcal{D}w$ denotes a functional integral over the field w, and $E[w]$ is a local quadratic functional of w that reflects the A-B monomer interactions and the local incompressibility constraint: [28]

$$E[w] = \frac{n}{2} \int d\mathbf{x}\ \left[\frac{1}{2\chi N} (w_B - w_A)^2 - (w_A + w_B) \right] \tag{24}$$

Here we have noted that for an AB diblock copolymer melt, the potential $w(\mathbf{x}, s)$ amounts to a two-component potential, i.e. $w(\mathbf{x}, s) = w_A(\mathbf{x})$ for $s \in [0, f]$ and $w(\mathbf{x}, s) = w_B(\mathbf{x})$ for $s \in [f, 1]$.

The partition function $Q[w, \underline{h}]$ can be obtained from a single-chain propagator $q(\mathbf{x}, s)$ that is the solution of a modified diffusion equation

$$\frac{\partial q}{\partial s} = R_{g0}^2 (\underline{G}^{-1})_{\alpha\beta} \frac{\partial^2 q}{\partial x_\alpha \partial x_\beta} - w(\mathbf{x}, s) q(\mathbf{x}, s) \tag{25}$$

subject to $q(\mathbf{x}, 0) = 1$. The single chain partition function is given by $Q[w, \underline{h}] = \int d\mathbf{x}\; q(\mathbf{x}, 1)$.

Finally, the partition function for an incompressible diblock copolymer melt confined to a cell of variable shape can be expressed as a field theory in the variables $\underline{h}$ and w:

$$Z = \int d(\underline{h}) (\det \underline{h})^{nN} \int \mathcal{D}w\; \delta(\det \underline{h} - V_0)\; \exp(-F[w, \underline{h}]) \tag{26}$$

where $F[w, \underline{h}]$ is an effective Hamiltonian given by

$$F[w, \underline{h}] = \beta V_0 \underline{t}\; : \underline{\epsilon} + E[w] - n \ln Q[w, \underline{h}] \tag{27}$$

In the mean-field approximation (SCFT), for a given shape $\underline{h}$ of the simulation box, we approximate the functional integral over w in (26) by the saddle point method. For this purpose, the functional $Q[w, \underline{h}]$ can be evaluated for any w and $\underline{h}$ by solving the modified diffusion equation (using e.g. a pseudo-spectral approach). The saddle point (mean-field) value of w, w^*, is obtained by applying a relaxation algorithm [28,30] to solve

$$\left. \frac{\delta F[w, \underline{h}]}{\delta w(\mathbf{x}, s)} \right|_{w=w^*} = 0 \tag{28}$$

In the mean-field approximation, $F[w^*, \underline{h}]$ corresponds to the free energy of the copolymer melt (in units of $k_B T$).

In a simulation at constant tension, the relaxation equation for the fields must be supplemented by a corresponding evolution equation for the cell. This equation is chosen to be a simple relaxation

$$\frac{d\underline{h}}{dt} = -\lambda_0\, \underline{h} \underline{D} \underline{h}^{-1} \frac{\partial F[w, \underline{h}]}{\partial \underline{h}} \tag{29}$$

where the tensor $\underline{D}$ is a projection operator whose action on an arbitrary tensor $\underline{M}$ is a traceless tensor, i.e. $\underline{D}\,\underline{M} \equiv \underline{M} - (1/3)\mathrm{Tr}(\underline{M})\underline{1}$. Equation (29) corresponds to a cell shape relaxation that (for $\lambda_0 > 0$) is down the gradient $\partial F / \partial \underline{h}$, approaching a local minimum of the mean-field free energy $F[w^*, \underline{h}]$. The "mobility" tensor $\underline{h} \underline{D} \underline{h}^{-1}$ is chosen so that the cell shape dynamics described by (29) conserves the cell volume.

Application of (29) requires an expression for the thermodynamic force $\partial F / \partial \underline{h}$. Explicit differentiation, noting the constraint of constant $\det \underline{h}$, leads to

$$\frac{\partial F[w, \underline{h}]}{\partial \underline{h}} = \beta V_0 \left(\frac{\partial}{\partial \underline{h}} \mathrm{Tr}(\underline{\epsilon t}\;) + \underline{h \Sigma} \right) \tag{30}$$

where $\underline{\Sigma}$ is a symmetric tensor defined by

$$\Sigma_{\alpha\beta}[w,\underline{h}] = -\frac{2k_BTn}{V}\frac{\partial \ln Q[w,\underline{h}]}{\partial G_{\alpha\beta}}$$
$$= \frac{k_BTn}{2VR_{g0}^2}\left\langle \int_0^1 ds \frac{dX_\alpha(s)}{ds}\frac{dX_\beta(s)}{ds}\right\rangle \qquad (31)$$

The angular brackets in the second expression denote an average over all conformations $\mathbf{X}(s)$ of a single copolymer chain that is subject to a prescribed chemical potential field w and fixed cell shape $\underline{h}$.

The first term on the right hand side of (30) can be conveniently rewritten as

$$\frac{\partial}{\partial h_{\alpha\beta}}(\mathrm{Tr}(\underline{\epsilon t}\,)) = \frac{\partial}{\partial h_{\alpha\beta}}(\mathrm{Tr}\frac{1}{2}\underline{h}^{T^{-1}_0}\underline{h}^T\underline{h}\,\underline{h}_0^{-1}\underline{t}\,) = (\underline{h}\,\underline{h}_0^{-1}\underline{t}\,\underline{h}^{T^{-1}_0})_{\alpha\beta}\,. \qquad (32)$$

Hence, (29) can be compactly expressed as

$$\frac{d\underline{h}}{dt} = -\lambda \underline{hD}\left[(\underline{h}_0^{-1}\underline{t}\,\underline{h}^{T^{-1}_0}) + \underline{\Sigma}\,\right] \qquad (33)$$

where $\lambda > 0$ is a new relaxation parameter defined by $\lambda = \beta V_0 \lambda_0$.

Equation (33) will evolve the cell shape to a configuration of minimum free energy (in the mean-field approximation). This configuration can either be metastable (local minimum) or stable (global minimum). Addition of a noise source to the equation provides a means for overcoming free energy barriers between metastable and stable states, i.e. a simple simulated annealing procedure.

An equilibrium solution of the cell shape (33) is evidently obtained when

$$(\underline{h}_0^{-1}\underline{t}\,\underline{h}^{T^{-1}_0}) + \underline{\Sigma}\; = 0 \qquad (34)$$

Combining (7), (31) and (34), it is seen that this equilibrium condition corresponds to a balance between the *externally* applied Cauchy stress, $\underline{\sigma}$, and the *internal* elastic stress, $\underline{\sigma}^{\,int}$, sustained by the polymer chains

$$\underline{\sigma}\; + \underline{\sigma}^{\,int} = 0 \qquad (35)$$

where

$$\sigma^{int}_{\alpha\beta}[w,\underline{h}] \equiv (\underline{h\Sigma}\;\underline{h}^T)_{\alpha\beta} = \frac{k_BT}{2VR_{g0}^2}\sum_{i=1}^{n}\left\langle \int_0^1 ds \frac{dR_{i\alpha}}{ds}\frac{dR_{i\beta}}{ds}\right\rangle\,. \qquad (36)$$

This expression for the internal polymer stress is well-known in the polymer literature [31].

Equation (33) drives a change in the shape of the simulation cell (at constant cell volume) to approach the equilibrium condition (34) at which the internal elastic stress of the copolymers balances the imposed external stress.

The last step is to find an expression for the internal stress tensor $\underline{\sigma}^{\,int}$ (36) or $\underline{\Sigma}$ (31) in terms of the single chain propagator, which is the central object

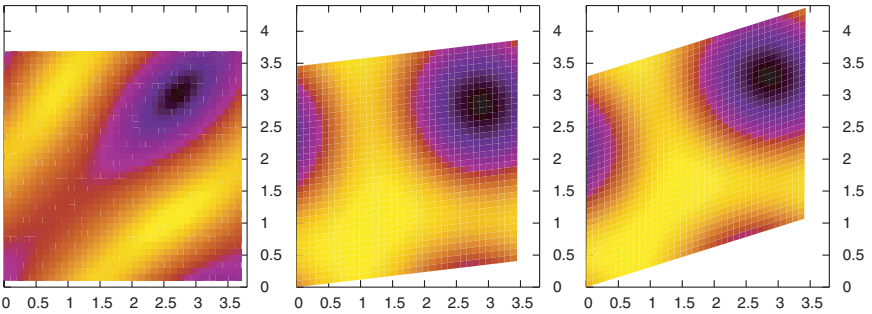


Fig. 5. Transformation of a square cell under zero external stress, when the melt is quenched into the stability region of the cylindrical phase ($\chi N = 15.9$, $f = 0.64$). Colors represent the concentration field

computed in a field-theoretic simulation [28]. The appropriate expression turns out to be

$$\frac{\sigma^{int}_{\alpha\beta}}{(n/V)k_BT} = -\frac{2R^2_{g0}}{Q} h^{-1}_{\gamma\alpha} \int d\mathbf{X} \int_0^1 ds \, \frac{\partial q(\mathbf{X}, s)}{\partial X_\gamma} \frac{\partial q(\mathbf{X}, 1-s)}{\partial X_\delta} h^{-1}_{\delta\beta} \qquad (37)$$

we refer to the appendix and to [32] for the derivation of this expression.

From the above derivations, it is clear that the quantity that is externally imposed in the method is not the Cauchy stress, but rather the thermodynamic tension. The Cauchy stress, which is the experimentally accessible quantity, is a result of the simulation, as is the cell shape. This feature is general in any application of the Parrinello–Rahman method in which a partition function of the form refzpol is sampled using Monte-Carlo, Langevin or molecular dynamics. The Cauchy stress has therefore to be obtained independently, using (44) in the present case, or in the case of molecular system through the Irving-Kirkwood formula.

To illustrate the method, Fig. 5 shows the evolution of the simulation cell under zero tension in a simple case. In the system under study, the parameters have been chosen so that the equilibrium phase is ordered on a triangular lattice, under zero external stress. Starting with a square simulation cell, evolution to the correct rhombohedral shape is obtained after a few relaxation steps. Other examples of application may be found in [32].

Acknowledgments

This work was supported by the CNRS. The work presented in these notes results from collaborations with F. Leonforte, A. Tanguy, J. Wittmer, L.J. Lewis, S.W Sides and G.H. Fredrickson. A very careful reading of the manuscript by Peter Henseler is also acknowledged. Finally, I would like to thank Mike Klein for the many stimulating discussions we had, over the last 20 years, on various topics, including -but not restricted to!- elasticity and glasses.

A Expression for the Stress Tensor in SCFT

To obtain (37), we start with the definition of $\underline{\Sigma}$

$$\beta V \Sigma_{\alpha\beta} = -2n \frac{1}{Q} \frac{\partial Q[w,G]}{\partial G_{\alpha\beta}} \tag{38}$$

The derivative of the single chain partition function can be calculated by discretizing the paths with a small contour step Δ.

$$\begin{aligned} -\frac{2}{Q}\frac{\partial Q[w,G]}{\partial G_{\alpha\beta}} &= \frac{1}{2R_{g0}^2 Q}\int_0^1 ds \int d\mathbf{X} \int d\mathbf{X}' \\ &\int \mathcal{D}\mathbf{X}(s)\delta(\mathbf{X}-\mathbf{X}(s))\delta(\mathbf{X}'-\mathbf{X}(s+\Delta))\left(\frac{X_\alpha - X'_\alpha}{\Delta}\right)\left(\frac{X_\beta - X'_\beta}{\Delta}\right) \\ &\exp\left(-\frac{1}{4R_{g0}^2}\int_0^1 ds \frac{dX_\alpha(s)}{ds} G_{\alpha\beta} \frac{dX_\beta(s)}{ds} - \int_0^1 ds\, w(\mathbf{X}(s),s)\right) \end{aligned} \tag{39}$$

Except between the points $\mathbf{X}, s$ and $\mathbf{X}', s+\Delta$ one can replace the path integrals with propagators q, so that

$$\begin{aligned} -\frac{2}{Q}\frac{\partial Q[w,G]}{\partial G_{\alpha\beta}} &= \frac{1}{2R_{g0}^2 Q}\int_0^1 ds \int d\mathbf{X} \int d\mathbf{X}' q(\mathbf{X},s) q(\mathbf{X}', 1-s-\Delta) \\ &\times \left(\frac{X_\alpha - X'_\alpha}{\Delta}\right)\left(\frac{X_\beta - X'_\beta}{\Delta}\right) \\ &\times \exp\left(-\frac{1}{4\Delta R_{g0}^2} G_{\alpha\beta}(X_\alpha - X'_\alpha)(X_\beta - X'_\beta)\right) \end{aligned} \tag{40}$$

One can then set $\mathbf{X}' = \mathbf{X} + \mathbf{u}$, and expand for small $\mathbf{u}$ and Δ according to

$$q(\mathbf{X}+\mathbf{u}, 1-s-\Delta) = q(\mathbf{X}, 1-s) - \Delta\frac{\partial q}{\partial s} + u_\gamma \frac{\partial q}{\partial X_\gamma} + \frac{1}{2} u_\gamma u_\delta \frac{\partial^2 q}{\partial X_\gamma \partial X_\delta} \tag{41}$$

The derivative w.r.t. s can be eliminated by applying the modified diffusion (25). One also requires second and fourth moments of the Gaussian distribution of displacements $\mathbf{u}$,

$$\overline{u_\alpha u_\beta} = G^{-1}_{\alpha\beta}(2R_{g0}^2\Delta)$$

$$\overline{u_\alpha u_\beta u_\gamma u_\delta} = (2R_{g0}^2\Delta)^2 (G^{-1}_{\alpha\beta}G^{-1}_{\gamma\delta} + G^{-1}_{\alpha\gamma}G^{-1}_{\beta\delta} + G^{-1}_{\alpha\delta}G^{-1}_{\beta\gamma})$$

By means of these results, we have

$$\begin{aligned} -\frac{2}{Q}\frac{\partial Q[w,G]}{\partial G_{\alpha\beta}} &= G^{-1}_{\alpha\beta}\frac{\int d\mathbf{X}\rho(\mathbf{X})}{\Delta} + G^{-1}_{\alpha\beta}\int d\mathbf{X} w(\mathbf{X})\rho(\mathbf{X}) \\ -\frac{R_{g0}^2}{Q} G^{-1}_{\alpha\beta} G^{-1}_{\gamma\delta} &\int d\mathbf{X} \int_0^1 ds q(\mathbf{X},s) \frac{\partial^2 q}{\partial x_\gamma \delta x_\beta} \end{aligned}$$

$$+\frac{R_{g0}^2}{Q}(G_{\alpha\beta}^{-1}G_{\gamma\delta}^{-1}+G_{\alpha\gamma}^{-1}G_{\beta\delta}^{-1}+G_{\alpha\delta}^{-1}G_{\beta\gamma}^{-1})\int d\mathbf{X}\int_0^1 dsq(\mathbf{X},s)\frac{\partial^2 q}{\partial x_\gamma \partial x_\delta} \tag{42}$$

where $\rho(\mathbf{X})=Q^{-1}\int_0^1 ds\, q(\mathbf{X},s)q(\mathbf{X},1-s)$ is the single-chain total monomer density operator. There is a partial cancellation in the last two terms so that

$$\begin{aligned}-\frac{2}{Q}\frac{\partial Q[w,G]}{\partial G_{\alpha\beta}} = \; & G_{\alpha\beta}^{-1}\tfrac{\int d\mathbf{X}\rho(\mathbf{X})}{\Delta}+iG_{\alpha\beta}^{-1}\int d\mathbf{X}w(\mathbf{X})\rho(\mathbf{X}) \\ & +\tfrac{2R_{g0}^2}{Q}G_{\alpha\gamma}^{-1}G_{\beta\delta}^{-1}\int d\mathbf{X}\int_0^1 dsq(\mathbf{X},s)\tfrac{\partial^2 q}{\partial X_\gamma \partial X_\delta}\end{aligned} \tag{43}$$

The internal polymer stress is obtained after matrix multiplication by $\underline{h}$ on the left and $\underline{h}^T$ on the right. This implies that the first two terms become a simple isotropic stress contribution, and are therefore not relevant to an incompressible system. The final formula for the internal stress tensor is therefore, apart from this diagonal contribution,

$$\sigma_{\alpha\beta}^{int}=\left(\frac{nk_BT}{V}\right)\frac{2R_{g0}^2}{Q}h_{\gamma\alpha}^{-1}\int d\mathbf{X}\int_0^1 ds\, q(\mathbf{X},s)\frac{\partial^2 q(\mathbf{X},1-s)}{\partial X_\gamma \partial X_\delta}h_{\delta\beta}^{-1} \tag{44}$$

The tensor Σ appearing in (33) is given by an expression similar to (44), with $\underline{G}$ replacing $\underline{h}$. The factor k_BTn/V accounts for the total number of chains, and produces a stress with the correct dimensions. In practice, the stress will be made dimensionless by dividing by this factor, so that the dimensionless stress is given by

$$\frac{\sigma_{\alpha\beta}^{int}}{(n/V)k_BT}=\frac{2R_{g0}^2}{Q}h_{\gamma\alpha}^{-1}\int d\mathbf{X}\int_0^1 ds\, q(\mathbf{X},s)\frac{\partial^2 q(\mathbf{X},1-s)}{\partial X_\gamma \partial X_\delta}h_{\delta\beta}^{-1} \tag{45}$$

Equation (37) is obtained after integrating by parts.

A local (rather than volume averaged) version of this connection between the stress tensor and the polymer propagator was derived previously in [33]. Numerically, $\sigma_{\alpha\beta}$ is evaluated from (45) using a pseudo-spectral scheme. The derivatives with respect to spatial coordinates are obtained by multiplying the propagator by the appropriate components of the wavevector in Fourier space, and transforming back into real space.

References

1. L. D. Landau and E. M. Lifshitz (1995) *Theory of Elasticity.* Butterworth-Heinemann, London
2. M. Parrinello and A. Rahman (1981) Polymorphic transitions in single-crystals – a new molecular-dynamics method. *J. Appl. Phys.* **52**, pp. 7182–7190

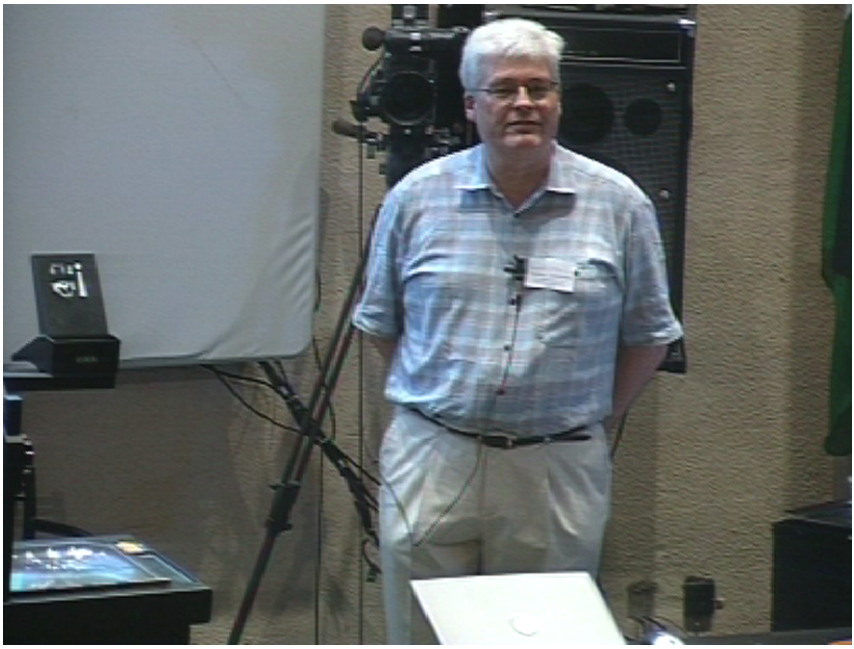
3. J. R. Ray and A. Rahman (1984) Statistical ensembles and molecular-dynamics studies of anisotropic solids. *J. Chem. Phys.* **80**, pp. 4423–4428
4. D. Wallace (1973) *Thermodynamics of crystals.* Wiley, New-York
5. T. H. K. Barron and M. L. Klein (1965) Second-order elastic constants of a solid under stress. *Proc. Phys. Soc.* **65**, pp. 523–532
6. D. R. Squire, A. C. Holt, and W. G. Hoover (1968), Isothermal elastic constants for argon: theory and Monte Carlo calculations. *Physica,* **42**, pp. 388–398
7. J. F. Lutsko (1989) Generalized expressions for the calculation of elastic-constants by computer-simulation. *J. Appl. Phys.* **65**, pp. 2991–2997
8. A. Lemaître and C. Maloney (2006) Sum rules for the quasi-static and visco-elastic response of disordered solids at zero temperature. *J. Stat. Phys.* **123**, pp. 415–453
9. I. Goldhirsch and C. Goldenberg (2002) On the microscopic foundations of elasticity. *Eur. Phys. J. E* **9**, pp. 245–251
10. K. Yoshimoto, T. S. Jain, K. van Workum, P. F. Nealey, and J. J. de Pablo (2004) Mechanical heterogeneities in model polymer glasses at small length scales. *Phys. Rev. Lett.* **93**, p. 175501
11. S. Sengupta, P. Nielaba, M. Rao, and K. Binder (2000) Elastic constants from microscopic strain fluctuations. *Phys. Rev. E* **61**, pp. 1072–1080
12. S. Pronk and D. Frenkel (2003) Large difference in the elastic properties of fcc and hcp hard-sphere crystals. *Phys. Rev. Lett.* **90**, p. 255501
13. J. Geng, G. Reydellet, E. Clement, and R. P. Behringer (2003) Green's function measurements of force transmission in 2D granular materials. *Physica D* **182**, pp. 274–303
14. E. Pratt and M. Dennin (2003) Nonlinear stress and fluctuation dynamics of sheared disordered wet foam. *Phys. Rev. E* **67**, p. 051402
15. A. Kabla and G. Debrégeas (2003) Local stress relaxation and shear banding in a dry foam under shear. *Phys. Rev. Lett.* **90**, p. 258303-1
16. M. Aubouy, Y. Jiang, J. A. Glazier, and F. Graner (2003) A texture tensor to quantify deformations. *Granular Matter* **5**, pp. 67–70
17. F. Leonforte, R. Boissiére, A. Tanguy, J. Wittmer, and J.-L. Barrat (2005) Continuum limit of amorphous elastic bodies. III. Three-dimensional systems. *Phys. Rev. B* **72**, p. 224206
18. J.-L. Barrat, J.-N. Roux, J.-P. Hansen, and M. L. Klein (1988) Elastic response of a simple amorphous binary alloy near the glass-transition. *Europhysics Letters* **7**, pp. 707–713
19. M. L. Falk and J. S. Langer (1998) Dynamics of viscoplastic deformation in amorphous solids. *Phys. Rev. E* **57**, pp. 7192–7205
20. S. A. Langer and A. J. Liu (1998) Effect of random packing on stress relaxation in foam. *J. Phys. Chem. B* **101**, pp. 8667–8671
21. J. P. Wittmer, A. Tanguy, J.-L. Barrat, and L. J. Lewis (2002) Vibrations of amorphous, nanometric structures: When does continuum theory apply? *Europhys. Lett.* **57**, pp. 423–430; A. Tanguy, J. P. Wittmer, F. Leonforte, and J.-L. Barrat (2002) Continuum limit of amorphous elastic bodies: A finite-size study of low-frequency harmonic vibrations. *Phys. Rev. B* **66**, p. 174205
22. F. Leonforte, A. Tanguy, J. P. Wittmer, and J.-L. Barrat (2004) Continuum limit of amorphous elastic bodies II: Linear response to a point source force. *Phys. Rev. B* **70**, p. 014203
23. E. Duval and A. Mermet (1998) Inelastic x-ray scattering from nonpropagating vibrational modes in glasses. *Phys. Rev. B* **58**, pp. 8159–8162

24. G. Baldi, S. Caponi, L. Comez, S. Di Fonzo, D. Fioretto, A. Fontana, A. Gessini, C. Masciovecchio, M. Montagna, G. Ruocco, S. C. Santucci, and G. Viliani (2005) Brillouin ultraviolet light scattering on vitreous silica. *Journal of non-crystalline solids* **351**, pp. 1919–1923
25. L. E. Silbert, A. J. Liu, and S. R. Nagel (2005) Vibrations and diverging length scales near the unjamming transition. *Phys. Rev. Letters* **95**, p. 098301; M. Wyart, L. E. Silbert, S. R. Nagel, and T. A. Witten (2005) Effects of compression on the vibrational modes of marginally jammed solids. *Phys. Rev. E* **72**, p. 051306
26. E. Maurer and W. Schirmacher (2004) Local oscillators vs. elastic disorder: A comparison of two models for the boson peak. *J. Low Temp. Phys.* **137**, pp. 453–470
27. W. H. Press, B. P. Flannery, and S. A. Teukolsky (1986) *Numerical Recipes.* Cambridge University Press, Cambridge
28. G. H. Fredrickson, V. Ganesan, and F. Drolet (2002) Field-theoretic computer simulation methods for polymers and complex fluids. *Macromolecules* **35**, pp. 16–39
29. M. W. Matsen and M. Schick (1994) Stable and unstable phases of a diblock copolymer melt. *Phys. Rev. Lett.* **72**, pp. 2660–2663
30. S. W. Sides and G. H. Fredrickson (2003) Parallel algorithm for numerical self-consistent field theory simulations of block copolymer structure. *Polymer* **44**, pp. 5859–5866
31. M. Doi and S. F. Edwards (1988) *The Theory of Polymer Dynamics.* Oxford University Press, Oxford
32. J-L. Barrat, G. H. Fredrickson, and S. W. Sides (2005) Introducing variable cell shape methods in field theory simulations of polymers. *J. Phys. Chem.* **B109**, pp. 6694–6700
33. G. H. Fredrickson, Dynamics and rheology of inhomogeneous polymeric fluids: A complex Langevin approach. *J. Chem. Phys.* **117**, pp. 6810–6820

Mesoscopic Simulations for Problems with Hydrodynamics, with Emphasis on Polymer Dynamics

B. Dünweg

Max-Planck-Institut für Polymerforschung, Ackermannweg 10, 55128 Mainz, Germany
duenweg@mpip-mainz.mpg.de



Burkhard Dünweg

B. Dünweg: *Mesoscopic Simulations for Problems with Hydrodynamics, with Emphasis on Polymer Dynamics*, Lect. Notes Phys. **704**, 309–340 (2006)
DOI 10.1007/3-540-35284-8_13 © Springer-Verlag Berlin Heidelberg 2006

1 Polymer Dynamics 311

1.1 Overview 311
1.2 Static Scaling in Polymer Solutions 313
1.3 Rouse Model 315
1.4 Zimm Model 320
1.5 Hydrodynamic Screening and Dynamic Crossover 322

2 Simulation Methods 323

2.1 Overview 323
2.2 Dissipative Particle Dynamics (DPD) 326
2.3 Multi-Particle Collision Dynamics (MPCD) 330
2.4 Lattice Boltzmann (LB) 331

3 Some Final Remarks 335

References 336

The contribution discusses two central models for polymer dynamics, the Rouse model and the Zimm model. The latter takes into account hydrodynamic interactions, and is hence appropriate for dilute solutions. In dense melts, the hydrodynamic interactions are screened, and the Rouse model is applicable (within limitations) as long as the chains are short enough to preclude reptation. The physics of this screening is discussed. The chapter then focuses on the methodological issue how to take hydrodynamic interactions into account in computer simulations. So far, the most successful methods are hybrid approaches where standard Molecular Dynamics for the polymer system is coupled to a mesoscopic model for momentum transport in the solvent. The most popular mesoscopic models are lattice Boltzmann and Dissipative Particle Dynamics. These methods are briefly discussed and contrasted. We describe a recent application of such an approach to the problem of hydrodynamic screening.

1 Polymer Dynamics

1.1 Overview

In this chapter, we will deal with the dynamic behavior of polymers [1] and discuss the simplest systems only: We will focus on *linear*, *flexible*, and *uncharged* macromolecules, disregading polydispersity (i.e. the broad distribution of chain lengths which usually occur in real systems), and study them in the *bulk*, in (or near) *thermal equilibrium*. We look at the systems from the point of view of coarse-grained models: We are not interested in the dependence of the properties on the details of the local chemistry, but rather ask for universal scaling laws. The aim of computer simulations within this subfield of polymer physics is to carefully check the pertinent predictions, and to try to elucidate the underlying physics. For certain properties, simulations can be numerically much more accurate than experiments, and they therefore complement the latter.

Although our point of view and our restriction on the physical conditions may look quite narrow, there is nevertheless a rich host of phenomena which need explanation. Since we look at *solutions*, we can ask for the dependence of static and dynamic properties on *monomer concentration* c, *chain length* N (i.e. the number of monomers per chain), and the solvent quality. Usually polymer and solvent are the more miscible the higher the temperature is, i.e. the solvent quality can be parameterized in terms of temperature T – although there exist some cases where miscibility occurs at low temperatures, such that the system unmixes upon heating. Disregarding such "pathological", entropy-driven situations, Fig. 1 shows the generic phase diagram of such a solution.

One sees that even the statics is quite non-trivial, giving rise to a host of scaling laws and various crossovers, and that the dynamics is even more

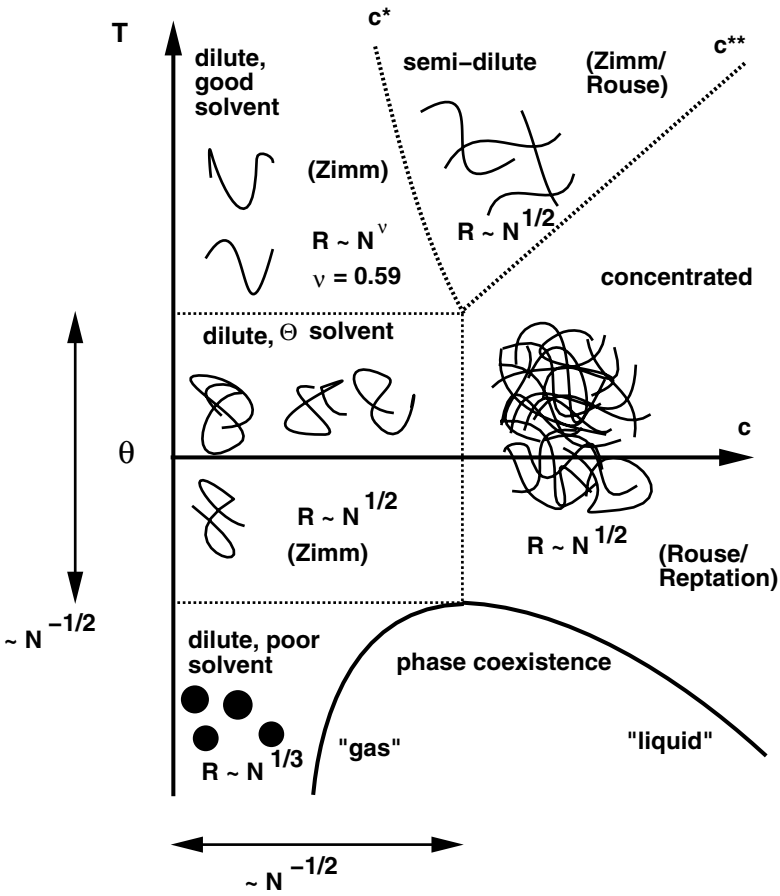


Fig. 1. Phase diagram of a polymer solution, in the plane monomer concentration c vs. temperature T (parameterizing solvent quality). The static properties are characterized by the scaling laws which describe the dependence of the chain size R (like the gyration radius or the end-to-end-distance) on the degree of polymerization N. In the dilute limit $c \to 0$, the so-called theta transition occurs, where at $T = \Theta$ single isolated chains collapse from a swollen random coil to a compact globule. For finite chain length N, this transition is "smeared out" over a temperature region $\Delta T \propto N^{-1/2}$, in which the chain conformations are Gaussian. Below Θ, phase coexistence between a "gas" of globules and a "liquid" of strongly interpenetrating Gaussian chains occurs. The corresponding critical point occurs at a very low concentration, $c_c \propto N^{-1/2}$, and in the vicinity of Θ, $\Theta - T_c \propto N^{-1/2}$. The crossover region which connects the regime of swollen isolated coils with that of the concentrated (Gaussian) solution at high temperatures is called the semidilute regime. The dynamics is characterized by the Zimm model in the dilute limit where hydrodynamic interactions are important, and by the Rouse model for dense systems where they are screened. For very dense systems and/or sufficiently long chains, where curvilinear motion dominates, the Rouse model must be replaced by the reptation model (or the crossover behavior between these two cases). Rouse and Zimm model are described in the text in detail. The Zimm-Rouse crossover which occurs in the semidilute regime is a central topic of the present chapter

complex. Our present state of understanding is based on three fundamental models: The Rouse model [2], the Zimm model [3], and the reptation model [1]. In the present chapter we will discuss the Rouse and the Zimm model (see below). Reptation, which is characterized by curvilinear motion, and occurs in dense long-chain systems, does not play a role for the simulations to be described later, and shall hence not be discussed here.

So far, simulations have successfully treated the following cases:

- A single chain in good solvent with Zimm dynamics [4–10].
- A dense melt and its crossover from Rouse to reptation dynamics [11–14].
- Semidilute solutions characterized by a crossover from Zimm to Rouse dynamics as the concentration is increased [15].

Quite promising attempts have also been made to study the dynamics of the theta transition (i.e. a single chain collapses upon decrease of the solvent quality) [16–18]. Nevertheless, a systematic exploration of the plane concentration vs. solvent quality has not yet been done. After reading this article, the reader may perhaps understand why. For this reason, the present chapter will also disregard solvent quality effects and focus on good solvents (high temperatures in Fig. 1) only. The methods to be discussed in this chapter aim at an optimal exploitation of the different physical nature of the solvent and the solute, and this holds for any solvent quality.

To summarize: Our concern is to use computer simulations to put dynamic scaling laws in dilute, semidilute, and concentrated systems under scrutiny. We will take the underlying static scaling laws for granted (i.e. checked previously by other simulations and/or experiments). Nevertheless, in order to understand the dynamic models, it is necessary to first briefly review the statics.

1.2 Static Scaling in Polymer Solutions

The static conformations of flexible polymer chains are described via the statistics of a random coil [19]. We model the chain as a sequence of N "monomers" with positions $\boldsymbol{r}_i$, $i = 1, \ldots, N$. From the chemist's point of view, a monomer is one chemical repeat unit, usually comprising several atoms (for instance, the repeat unit of polyethylene consists of one carbon and two hydrogen atoms). The position $\boldsymbol{r}_i$ can thus be viewed as the center of mass of the repeat unit i (or some similar quantity characterizing where the unit is). We denote the typical bond length with b,

$$\left\langle b^2 \right\rangle = \left\langle \left(\boldsymbol{r}_{i+1} - \boldsymbol{r}_i\right)^2 \right\rangle \ . \tag{1}$$

It is however also possible to combine several repeat units into a "supermonomer". In such a case, the same chain would be described by a larger value of b, and, correspondingly, by a smaller value of N. The notion of a monomer is therefore somewhat arbitrary. This leads us to the important principle of *scale*

invariance: The measurable large-scale properties of the chain may not depend explicitly on the way in which the chain has been decomposed into monomers. By iterating this coarse-graining procedure ("renormalization group"), and applying dimension arguments, it is then easy to show that scale invariance implies power-law behavior. For a single chain which has no other important length scale than R (size of the chain as a whole), and b_{min} (shortest atomistic length scale below which a yet finer decomposition is impossible), the power law reads

$$R \sim bN^{\nu} \; . \tag{2}$$

Here we have assumed that one specific definition of a monomer (implying one specific value of N) has been chosen. The exponent ν depends on the physical conditions: For isolated chains in good solvent, the chains are swollen as a result of the excluded-volume interaction, with $\nu \approx 0.588$ in three dimensions. The statistics is that of a self-avoiding walk (SAW). In dense systems, the excluded volume interaction is screened [19, 20], hence ν has the Gaussian or random walk (RW) value $\nu = 1/2$. In what follows, the letter ν will either denote both values (in cases where the distinction between RW and SAW does not matter), or the SAW value (in cases where it does).

We have not specified precisely how to measure R; the scaling law applies to *all* ways of defining it. Convenient measures are the end-to-end-distance,

$$\left\langle R_E^2 \right\rangle = \left\langle \left(\boldsymbol{r}_N - \boldsymbol{r}_1 \right)^2 \right\rangle \; , \tag{3}$$

the gyration radius

$$\left\langle R_G^2 \right\rangle = \frac{1}{N} \sum_i \left\langle \left(\boldsymbol{r}_i - \boldsymbol{R}_{CM} \right)^2 \right\rangle \tag{4}$$

($\boldsymbol{R}_{CM}$ denoting the chain's center of mass, $\boldsymbol{R}_{CM} = N^{-1} \sum_i \boldsymbol{r}_i$), and the hydrodynamic radius

$$\left\langle \frac{1}{R_H} \right\rangle = \frac{1}{N^2} \sum_{i \neq j} \left\langle \frac{1}{r_{ij}} \right\rangle \; , \tag{5}$$

where $r_{ij} = |\boldsymbol{r}_i - \boldsymbol{r}_j|$.

Another important way to characterize the conformations is via the single-chain static structure factor $S(k)$, defined as

$$S(k) = \frac{1}{N} \left\langle \left| \sum_i \exp\left(i\boldsymbol{k} \cdot \boldsymbol{r}_i \right) \right|^2 \right\rangle = \frac{1}{N} \left\langle \sum_{ij} \exp\left(i\boldsymbol{k} \cdot \left(\boldsymbol{r}_i - \boldsymbol{r}_j \right) \right) \right\rangle \; . \tag{6}$$

On length scales $b \ll k^{-1} \ll R$, $S(k)$ does not depend on N (note that the addition of exponentials is just a RW in the complex plane). On the other hand, $S(k)$ must have the scaling form $S(k) = Nf(kR)$. This implies a power law decay $S(k) \propto k^{-1/\nu}$.

The conformations can always be described in terms of an effective potential V (potential of mean force) such that the equilibrium probability density is

$$P\left(\{\boldsymbol{r}_i\}\right) \propto \exp\left(-\frac{V\left(\{\boldsymbol{r}_i\}\right)}{k_B T}\right) , \tag{7}$$

k_B and T denoting the Boltzmann constant and the absolute temperature, respectively.

For a Gaussian chain, V is just a harmonic potential:

$$V = \frac{3k_B T}{2b^2} \sum_{i=1}^{N-1} \left(\boldsymbol{r}_{i+1} - \boldsymbol{r}_i\right)^2 . \tag{8}$$

For a SAW chain, additional repulsive potentials between the monomers must be added in order to model the excluded-volume interaction. Models of this type are called "bead-spring" models.

Let us now discuss the crossover from SAW to RW behavior when the concentration is increased. We will always assume good solvent conditions. The coil size remains independent of concentration as long as the chains do not overlap. The concentration $c^\star$ where overlap starts to happen [19] is estimated via the requirement that an arrangement of unperturbed SAWs is just space-filling:

$$c^\star \sim \frac{N}{R^3} \sim \frac{N}{b^3 N^{3\nu}} \sim b^{-3} N^{-(3\nu-1)} . \tag{9}$$

Solutions with concentration $c \gg c^\star$ (well above overlap) but $c \ll b^{-3}$ (i.e. monomer concentration still very small) are called *semi-dilute*. Such solutions have another important length scale ξ, intermediate between b_{min} and R. ξ is called the "blob size" and can be defined as follows: If the chains were cut into sub-chains, each with size ξ, then the solution would be just at the overlap concentration corresponding to this lower molecular weight. On length scales below ξ, the statistics corresponds to SAW behavior, while for length scales beyond ξ RW behavior applies. Denoting the number of monomers within the blob with n, we have $\xi \sim b n^\nu$, and $c \sim n/\xi^3$, hence

$$\xi \sim b \left(c b^3\right)^{-\frac{\nu}{3\nu-1}} \propto c^{-0.77} . \tag{10}$$

The chain is then viewed as a RW sequence of blobs, with $R \sim \xi (N/n)^{1/2}$. The single-chain structure factor decays as $S(k) \propto k^{-1/\nu}$ for $b \ll k^{-1} \ll \xi$, and as k^{-2} for $\xi \ll k^{-1} \ll R$.

1.3 Rouse Model

The Rouse model [1, 2] is the simplest model of polymer dynamics, while the Zimm model and the reptation model are slightly more complicated modifications. Essentially, the Rouse model considers a single chain described by effective interactions as outlined in the previous paragraph, and its overdamped

Brownian motion resulting from its coupling to a simple viscous background and to thermal noise. The Zimm model [3] replaces the viscous background by a hydrodynamic continuum which can transport momentum, and thus takes hydrodynamic interactions into account. Conversely, the reptation model [1] is a generalization of the Rouse model where topological constraints are taken into account in terms of an effective tube to which the chain is confined. In what follows, we will first outline the mathematical description of the Rouse model, and then attempt to critically assess its assumptions. In particular, we will try to briefly discuss the neglect of hydrodynamics (Zimm model physics), and the neglect of entanglements (reptation model physics). This discussion will be followed by an outline of the consequences of the Rouse equation of motion.

The dynamic variables of the Rouse model are just the positions of the monomers of a single test chain, $\boldsymbol{r}_i$, $i = 1, \ldots, N$. Taking into account that the model is supposed to describe its motion in an environment of other chains (a melt), we see that there is a colossal reduction in the number of degrees of freedom – solving the full many-body problem would require the positions and momenta of all monomers. Now, the Rouse model assumes the following overdamped Langevin equation of motion:

$$\frac{d}{dt}\boldsymbol{r}_i = \frac{1}{\zeta}\boldsymbol{F}_i + \boldsymbol{\rho}_i \ , \tag{11}$$

where ζ is the monomer friction coefficient, $\boldsymbol{\rho}_i$ the random displacement (per unit time) acting on monomer i, and $\boldsymbol{F}_i$ the force acting on monomer i, derived from the effective potential V (see previous subsection):

$$\boldsymbol{F}_i = -\frac{\partial V}{\partial \boldsymbol{r}_i} \ . \tag{12}$$

The random displacements are Gaussian white noise satisfying the standard fluctuation-dissipation theorem,

$$\langle \rho_i^\alpha \rangle = 0, \qquad \left\langle \rho_i^\alpha(t)\rho_j^\beta(t') \right\rangle = 2\frac{k_B T}{\zeta}\delta_{ij}\delta_{\alpha\beta}\delta(t-t') \ , \tag{13}$$

such that the equilibrium distribution produced by the Langevin process is the correct one. (Greek letters denote Cartesian indices.) The stochastic displacements in different directions, and of different monomers, are assumed as statistically independent.

At this point, it is clear that many important aspects have been disregarded. Firstly, the neglect of the momenta (i.e. that fact that only positions occur as dynamical variables) is safe only at the first, but not at second glance. The motivation of the neglect is the idea that in a dense simple fluid the motion of particles is essentially an oscillation in a local cage, and escape occurs only after many collisions, such that the memory of the initial momentum is completely lost on the time scale on which a monomer moves its own size.

However, since the discovery of the long-time tails [21] we know that this is not quite true: Since momentum is a conserved quantity, it can only be transported away, but not simply destroyed. Let us therefore discuss the physics of momentum transport in some more detail. On long time and length scales, it can be described by the Stokes equation (hydrodynamic equation of motion for an incompressible fluid, where the nonlinear term is neglected):

$$\rho \frac{\partial}{\partial t} \boldsymbol{u} = \eta \nabla^2 \boldsymbol{u}, \quad \nabla \cdot \boldsymbol{u} = 0 \,, \tag{14}$$

where ρ is the fluid density, η its viscosity, and $\boldsymbol{u}$ the velocity flow field. The momentum transport hence takes place in a *diffusive* fashion, where the so-called "kinematic viscosity" $\eta_{kin} = \eta/\rho$ plays the role of a diffusion constant. Within the time t, an initial momentum therefore spreads into a sphere whose radius is of order $(\eta_{kin} t)^{1/2}$, or whose volume is of order $(\eta_{kin} t)^{3/2}$. This is the reason for the $t^{-3/2}$ decay of the velocity autocorrelation function. It should also be noted that a hydrodynamic description of a fluid is *always* valid on sufficiently large length and time scales for *any* fluid (including polymer melts).

These considerations, however, generate a puzzle for the validity of the Rouse model: Shouldn't one expect, from the physics of the long-time tails, that the monomer will need time "forever" until it really forgets its original velocity? Doesn't that imply that the description in terms of a position-only Langevin equation is expected to fail, and shouldn't one rather introduce an appropriate memory function to describe such a lack of forgetfulness?

However, let us look at the systems for which the Rouse model works (at least to a good approximation). These are short-chain melts. In such a system, the collisions between monomers do not occur in a nice and orderly fashion as in a simple fluid. Rather, the typical collision process is a chain-chain collision, such that an incoming kick will mainly result in a chain elongation against the connectivity forces, rather than being transported straight along. These processes destroy the memory to a large extent, and they are also the reason for hydrodynamic screening (see below). The random arrangement of chains results in a randomization of the scattering, and correlations are removed. Therefore, the monomer *does* ("at third glance") forget its initial velocity rather quickly – except for (i) an extremely small contribution to the velocity autocorrelation function which comes from the long-time, large-length hydrodynamics of the melt (on scales beyond the gyration radius and the chain relaxation time), and can safely be neglected (it has so far never been resolved in simulational or experimental studies of melts), and (ii) a long-time negative contribution which describes the slowing down due to the spreading of correlations along the chain backbone [22]. However, phenomenon (ii) is actually faithfully described *within* the Rouse model. What this means is that the Rouse model monomer diffusion coefficient $k_B T/\zeta$ must be viewed as a *short-time* diffusion coefficient – not to be mistaken for the long-time diffusion coefficient, which is identical to that of the chain as a whole (i.e. much smaller)

and which can be obtained as the time integral of the velocity autocorrelation function. Altogether, this argument restores the validity of the simple friction coefficient ansatz. Of course, the value of ζ depends on the definition of what one calls a monomer, just as the value of b does.

As a side remark, let us note that the typical "simple fluid" collision processes *will* play a role in a dilute solution, and the Rouse model is not expected to work. However, the hydrodynamic effects will not only induce correlations of the monomer with itself at later times, but also with *other* monomers on the chain. This phenomenon is called "hydrodynamic interaction". Taking these correlations into account, one obtains the Zimm model, which is discussed below in more detail. Considering again the single-monomer velocity autocorrelation function, one will have a short-time decay (giving rise to a short-time diffusion coefficient, or a monomer friction coefficient), followed by a negative contribution resulting from both the Rouse-like slowing down and long-time-tail memory.

As a second caveat of the Rouse equation of motion, it must be stressed that the model is a *single-chain* theory, i.e. all correlation effects with other chains are ignored. This latter neglect may be fine in dilute solution, but it is far from obvious for a dense melt where the chains are very close to each other. Indeed, the dramatic failure of the Rouse model for melts of sufficiently long chains, where rather the reptation model applies, is an obvious hint of this fact. The reptation model, where the entanglements with the other chains are replaced by a tube of a certain diameter [1], is of course yet another single-chain theory. From this point of view, it is not too surprising that the Rouse model does not work *precisely* for polymer melts. Rather, computer simulations and experiments have revealed a number of deviations [23, 24], the most interesting of which is a subdiffusive motion of the chain's center of mass, which is not predicted by the Rouse model. There are attempts by analytical theory [25, 26], but this issue is still under investigation. To some extent, the deviations might be trivially due to the fact that applicability of the Rouse model requires that the chains are long enough to satisfy a Gaussian description, and at the same time short enough that reptation does not yet play a role. Usually this window of chain lengths is very small, and in many cases it practically does not exist.

Let us now discuss the consequences of the Rouse model. It is interesting to note that for a Gaussian (RW) chain the model can be solved *exactly*. The reason is that in this case V is a harmonic potential, i.e. the equation of motion is linear. The problem is then mathematically very similar to phonons in a one-dimensional solid. Analogously to phonons one introduces the so-called Rouse modes

$$\boldsymbol{X}_p = \sqrt{2}\, N^{-1/2} \sum_{i=1}^{N} \boldsymbol{r}_i \cos\left[\frac{p\pi}{N}\,(i-1/2)\right], \quad p = 1, \ldots N-1\,, \tag{15}$$

whose equations of motion decouple. Each mode is characterized by a mean square amplitude

$$\left\langle \boldsymbol{X}_p^2 \right\rangle = \frac{b^2}{4\sin^2\left(\frac{p\pi}{2N}\right)} , \tag{16}$$

and a relaxation function

$$\left\langle \boldsymbol{X}_p(t) \cdot \boldsymbol{X}_p(0) \right\rangle = \left\langle \boldsymbol{X}_p^2 \right\rangle \exp\left(-\frac{t}{\tau_p}\right) , \tag{17}$$

where the mode relaxation time τ_p is given by

$$\tau_p^{-1} = \frac{12\, k_B T}{\zeta b^2} \sin^2\left(\frac{p\pi}{2N}\right) . \tag{18}$$

Note that for long chains and small mode number this is approximated by the scaling law $\tau_p \propto (N/p)^2$. The longest relaxation time is $\tau_1 \equiv \tau_R$, the so-called Rouse time, scaling as $\tau_R \propto N^2 \propto R^4$. At this point it is useful to introduce a dynamic exponent z, which relates the length scale R with the corresponding time scale τ_R via

$$\tau_R \propto R^z . \tag{19}$$

We therefore see $z = 4$ for the RW Rouse model.

The mean square displacement of a single monomer, $\left\langle \Delta \boldsymbol{r}^2 \right\rangle$, can be written exactly as a complicated expression, which however behaves asymptotically as $\left\langle \Delta \boldsymbol{r}^2 \right\rangle \propto t^{1/2}$ for times $\tau_m \ll t \ll \tau_R$, where τ_m is the microscopic time scale given by the time a monomer needs to move its own size, $\tau_m = b^2(\zeta/k_B T)$. The zeroth Rouse mode describes the motion of the center of mass, which is pure diffusion on all time scales, with diffusion constant

$$D = \frac{k_B T}{N\zeta} . \tag{20}$$

For times large compared to τ_R, *all* monomers move just diffusively with diffusion constant D.

It is physically more instructive to derive the results of the Rouse model directly by scaling reasoning, and to do this for RW and SAW statistics simultaneously. Looking at first at the equation of motion for the center of mass, one notes that the drift (force) terms all cancel, due to Newton's third law. Furthermore, the friction coefficients of all the monomers simply add up; hence (20) is derived immediately. The scaling law is

$$D \propto N^{-1} \propto R^{-1/\nu} . \tag{21}$$

Furthermore, one estimates the longest relaxation time τ_R via the consideration that the object will just move its own size within τ_R:

$$D\tau_R \propto R^2 \quad \tau_R \propto R^{2+1/\nu} , \tag{22}$$

from which we read off $z = 2 + 1/\nu$ for the general case. Furthermore, scaling tells us that the mean square displacement of the single monomer should follow a power law for times $\tau_m \ll t \ll \tau_R$, simply because the system has no further important time scale. Requiring $\left\langle \Delta r^2 \right\rangle \sim R^2$ for $t \sim \tau_R$ fixes the exponent as

$$\left\langle \Delta r^2 \right\rangle \propto t^{2/z} \; . \tag{23}$$

The physical picture is that originally a monomer can move freely (with diffusion constant $D_0 = k_B T/\zeta$), while at later times it has to drag more and more neighboring monomers along. Therefore the effective diffusion constant systematically decreases with time, until (at $t = \tau_R$) the whole chain is dragged along.

Finally, the single-chain dynamic structure factor, defined as

$$S(k,t) = \frac{1}{N} \left\langle \sum_{ij} \exp\left(i\boldsymbol{k} \cdot (\boldsymbol{r}_i(t) - \boldsymbol{r}_j(0))\right) \right\rangle \; , \tag{24}$$

satisfies the scaling relation

$$S(k,t) = k^{-1/\nu} f\left(k^2 t^{2/z}\right) \tag{25}$$

for $b \ll k^{-1} \ll R$ and $\tau_m \ll t \ll \tau_R$.

1.4 Zimm Model

As already indicated above, in dilute solutions it is necessary to take hydrodynamic momentum transport into account. The main effect is the so-called "hydrodynamic interaction": A monomer i is randomly kicked by its solvent surrounding, and is moved by a certain random displacement (per unit time) $\boldsymbol{\rho}_i$. Another monomer j suffers a displacement $\boldsymbol{\rho}_j$. Now, the motion of the solvent particles near $\boldsymbol{r}_i$ is *highly correlated* with that at position $\boldsymbol{r}_j$, due to fast diffusive momentum transport through the solvent. Strictly spoken, this correlation only occurs at some later time (the time which the "signal" needs to travel from $\boldsymbol{r}_i$ to $\boldsymbol{r}_j$), but this is usually quite short compared to the time which the monomers i and j need to travel considerably. As already discussed, the momentum transport occurs with the "diffusion constant" η_{kin}, while the particles move (initially) with diffusion constant $D_0 = k_B T/\zeta$. The dimensionless ratio is called the Schmidt number $Sc = \eta_{kin}/D_0$; its value controls how accurate the neglect of retardation effects is. Typical numbers for Sc in dense fluids are of the order $Sc \approx 10^2$.

Thus, in contrast to the Rouse case, where the stochastic displacements exhibit no correlations between different particles and between different spatial directions, we now have a non-trivial correlation function $\langle \boldsymbol{\rho}_i(t) \otimes \boldsymbol{\rho}_j(t') \rangle \propto \delta(t - t')$, where the tensorial nature is due to the incompressibility constraint of the solvent flow, and the delta function in time expresses the neglect of retardation effects. (The symbol $\otimes$ denotes the tensor product.)

Furthermore, a force $\boldsymbol{F}$ acting on a monomer at $\boldsymbol{r}_i$ will generate a surrounding flow around it. Again neglecting retardation, one can use the *stationary* Stokes equation to calculate the resulting flow field. The solution is [1] the so-called Oseen tensor

$$\boldsymbol{u}(\boldsymbol{r}) = \overleftrightarrow{T}\,(\boldsymbol{r} - \boldsymbol{r}_i) \cdot \boldsymbol{F} \tag{26}$$

$$\overleftrightarrow{T}\,(\boldsymbol{r}) = \frac{1}{8\pi\eta r}(\overleftrightarrow{1} + \hat{r} \otimes \hat{r}) \,, \tag{27}$$

where $\hat{r}$ is the unit vector in the direction of $\boldsymbol{r}$.

We now write down the most general Langevin equation which is still memory-free, does not use more variables than the monomer coordinates, and satisfies the fluctuation-dissipation theorem (to assure the correct equilibrium distribution function):

$$\frac{d}{dt}\boldsymbol{r}_i = \sum_j \overleftrightarrow{\mu}_{ij} \cdot \boldsymbol{F}_j + \boldsymbol{\rho}_i \tag{28}$$

where $\overleftrightarrow{\mu}_{ij}$ is the mobility tensor, and the stochastic displacements satisfy the relation

$$\langle \boldsymbol{\rho}_i(t) \otimes \boldsymbol{\rho}_j(t') \rangle = 2k_B T \overleftrightarrow{\mu}_{ij}\, \delta(t - t') \,. \tag{29}$$

We only consider mobility tensors which are divergence-free, and hence we need not worry about "spurious drift" terms [1]. Now, since the monomers are essentially just "embedded" in the surrounding flow, we can identify (at least approximately) $\overleftrightarrow{\mu}_{ij} = \overleftrightarrow{T}\,(\boldsymbol{r}_i - \boldsymbol{r}_j)$ (note that both objects just describe the velocity response to a force). This holds of course only for $i \neq j$; for the diagonal elements we assume the Rouse form $\overleftrightarrow{\mu}_{ii} = \overleftrightarrow{1}\,/\zeta$.

The scaling analysis of the Zimm model now proceeds along the same lines as for the Rouse model. First, we study the center-of-mass diffusion constant. In the short-time limit, it is easy to show that the center of mass moves with the Kirkwood diffusion constant

$$D^{(K)} = \frac{D_0}{N} + \frac{k_B T}{6\pi\eta} \left\langle \frac{1}{R_H} \right\rangle \,, \tag{30}$$

and this differs only marginally from the long-time value [27]. Since the second term strongly dominates in the long-chain limit, we find the scaling law

$$D \propto \frac{1}{R} \,, \tag{31}$$

indicating that the chain as a whole essentially moves like a Stokes sphere. The longest relaxation time (Zimm time τ_Z) is again found by requiring $D\tau_Z \sim R^2$, resulting in $\tau_Z \propto R^3$, or $z = 3$ independently from chain statistics. The dynamics is thus faster than in the Rouse case. All other scaling relations remain the same; one just has to use the appropriate values for ν and z.

1.5 Hydrodynamic Screening and Dynamic Crossover

Upon increasing the concentration, we have a static crossover from SAW to RW behavior; this is controlled by the blob size ξ. There is also a dynamic crossover from Zimm to Rouse behavior the physics of which had not fully been understood until very recently when a computer simulation [15] clarified the last remaining puzzles.

The underlying question is: How does the system get rid of its hydrodynamic correlations? Early attempts [28, 29] tried to attack this by studying the multiple scattering of the flow field. The attractive feature of such considerations is the fact that, under the assumption of a frozen polymer matrix, the analog of the Oseen tensor can be easily calculated: Assuming an array of fixed random obstacles with concentration c and friction coefficient ζ (per obstacle), the solvent flow field $\boldsymbol{u}$ experiences a friction force per unit volume of $-\zeta c\boldsymbol{u}$, such that the Stokes equation is modified to

$$\rho\frac{\partial}{\partial t}\boldsymbol{u} = \eta\nabla^2\boldsymbol{u} - \zeta c\boldsymbol{u} \; . \tag{32}$$

Its Green's function now exhibits a Debye-Hückel-like decay of the form $\propto (1/r)\exp(-r/\xi_H)$, where the hydrodynamic screening length ξ_H is found via $\eta\xi_H^{-2} = \zeta c$. The long-range Oseen decay is replaced by a short-range interaction, and thus Rouse behavior is expected on length scales beyond ξ_H, while Zimm behavior should apply on short length scales.

De Gennes [30] has critized this approach for the following reasons: (i) The chains are not at all fixed obstacles, but, on the contrary, enslaved to the surrounding flow and just dragged along; (ii) the predicted scaling $\xi_H \propto c^{-1/2}$ would imply that ξ and ξ_H are not proportional to each other, which makes a scaling analysis difficult if not impossible.

De Gennes' solution to the puzzle [30] is based on the following argument: A description in terms of fixed obstacles *is* justified, however only on length scales beyond the blob size ξ. For these larger length scales, it is the *entanglements* (not in the sense of reptation theory, but rather in the sense of *mutual interaction*) which cause a hindrance in polymer mobility compared to fast Zimm motion. Since a Zimm chain behaves essentially like a Stokes sphere, one arrives at the picture of blobs "hooked up" in a temporary gel. Therefore the appropriate obstacles are not the monomers, but rather the blobs, with Stokes friction coefficient $\zeta_{blob} \sim \eta\xi$, and concentration $c_{blob} \sim \xi^{-3}$. Inserting these relations into $\eta\xi_H^{-2} = \zeta_{blob}c_{blob}$, one finds $\xi_H \sim \xi$, i.e. the length scales are (apart from prefactors) *identical*.

This has been confirmed by most experiments [31,32]; however, the picture of clean Rouse motion beyond the length scale ξ was questioned by the observation of "incomplete screening" in the data of neutron spin echo experiments on labeled chains [33], where a clear Zimm-like contribution was found. The results of our recent simulation [15] revealed the solution of this puzzle: It is not sufficient to look at the problem just in terms of length scales, but one has

to consider the time scales as well, and distinguish between the cases $t \ll \tau_\xi$ and $t \gg \tau_\xi$, where τ_ξ is the blob (Zimm) relaxation time, $\tau_\xi \sim \eta\xi^3/(k_BT)$. Since one has to wait (on average) for a time of order τ_ξ until an entanglement (or, synonymously: an interaction, a chain-chain collision) occurs, there is no screening whatsoever for short times. The motion is rather free Zimm relaxation on *all* length scales, and the chains are just dragged along with the flow. After τ_ξ, the interactions are felt, and the blob screening mechanism sets in, resulting in Rouse-like motion. This, however, has only an effect on the length scales beyond ξ, since at that time all correlations within the blob have already decayed.

Let us now discuss the numerical results of [15]. In order to simulate a real semidilute solution, it is necessary to resolve both the RW regime at large length scales and the SAW regime within the blob. In order to observe random coil behavior in computer chains, a certain minimum number of monomers is necessary. According to our experience, one needs at least $N \approx 30$ monomers to clearly see the scaling behavior of either a RW or a SAW. For our semidilute solution, this means that we need roughly 30 monomers per blob, and roughly 30 blobs per chain. Therefore the minimum chain length is roughly $N = 1000$. For such a chain we then expect a mean size of $R = 30^\nu \times 30^{1/2} \approx 40$, in units of the bond length. In order to safely exclude self-overlaps, one would like to make the linear size of the simulation box (with periodic boundary conditions) substantially bigger. Our largest system therefore had linear box size $L = 88$. The concentration is then obtained via $\xi \approx 30^\nu \approx c^{-0.77}$ or $c \approx 0.066$. The total number of monomers in the $L = 88$ box then results as 45000. In the actual simulation, we studied 50 chains of length $N = 1000$ *which is essentially the smallest system to study a semidilute solution.* For such a system we calculated the single-chain dynamic structure factor $S(k,t)$. The static ($t = 0$) structure factor revealed the expected RW and SAW regimes. To analyze dynamic scaling, one plots the data as a function of the scaling argument $k^2t^{2/z}$. Indeed we found Zimm behavior ($z = 3$) at short times and small length scales, while the data show RW Rouse behavior ($z = 4$) for late times, large length scales. A particularly careful analysis was necessary to distinguish the short-time and long-time regimes (Zimm vs. Rouse) for the data at large length scales ($k\xi < 1$). In order to enhance the short-time region, Fig. 2 shows $-\ln[S(k,t)/S(k,0)]$ instead of simply $S(k,t)/S(k,0)$ as a function of the scaling argument, for *both* Zimm and Rouse scaling. It is clearly seen that Zimm scaling applies for the short times $t \ll \tau_\xi$, while Rouse scaling holds for the later times $t \gg \tau_\xi$.

2 Simulation Methods

2.1 Overview

We now ask the question: What is the right way to simulate a system with hydrodynamic interactions? For simple problems, many methods will work,

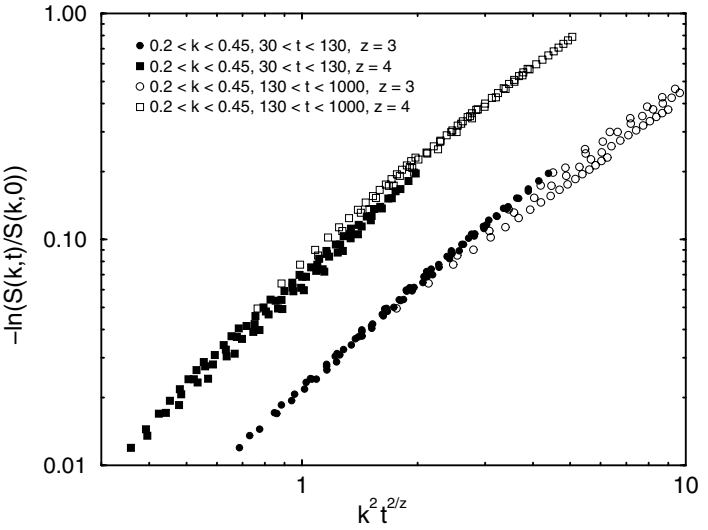


Fig. 2. Scaling plot of single-chain dynamic structure factor data, for both Rouse and Zimm scaling, taken from [15]. For more details, see the text and [15]

but for a challenging application like the system of [15] it is necessary to choose and design the method carefully.

The most straightforward approach would be Brownian dynamics, where just (28) is simulated directly, either for a single-chain system, or a many-chain system. However, this will not work for a system of 50000 monomers. Each time step one would have to calculate a 150000×150000 matrix, and to calculate its square root, in order to find the stochastic displacements. This is beyond the capacity of today's computers. The unfavorable scaling of the computational complexity with the number of Brownian particles makes the method only feasible for small systems. In [27] we studied the Zimm equation of motion for a single chain, and calculated the diffusion constant accurately. The longest chain which was accessible was $N = 200$.

It is therefore quite clear that one needs a method which scales *linearly* with the number of Brownian particles. An $O(N)$ algorithm for evaluating hydrodynamic interactions has indeed been developed, in close analogy to the fast multipole method for electrostatic interactions [34]. However, this is practically only applicable to deterministic problems (like sedimentation) where no thermal noise needs to be considered. For stochastic simulations, the problem of calculating the matrix square root remains.

Therefore the most promising route is to simulate the momentum transport through the solvent *explicitly* via some computational scheme. The most straightforward way to do this is of course Molecular Dynamics (MD), where the Brownian particles are immersed into a bath of solvent particles, and Newton's equations of motion are solved (without modification like a thermostat

etc.). However, this leads to an unnecessarily large computational effort. One needs to follow the motion of each solvent particle down to the time scale of the local oscillation of the particles in their cages, while they have essentially no function except for transporting momentum. One rather would like to simulate the solvent on a somewhat larger time scale, in order to save computer time. Essentially, this is coarse-graining with respect to time scales (not so much with respect to length scales, since one would not like to lose resolution in representing the hydrodynamic interactions).

Indeed, there are several ways to do this. One approach, which is conceptually particularly close to MD, is Dissipative Particle Dynamics (DPD) [35–46], which has become a quite popular method for "mesoscopic" simulations of the dynamics of soft-matter systems. One makes the particles quite soft, in order to afford a large time step, and also adds a momentum-conserving Langevin thermostat. It should be stressed that these two components are conceptually completely independent, and can also be implemented independently. It is therefore relatively straightforward to change an existing MD code into a DPD simulation, by just adding the thermostat. More details on DPD will follow below.

Another simulation aspect which needs appreciation is the issue of equilibration. This is of course completely uninteresting for nonequilibrium studies, which become more and more important, but for studying the dynamics in strict thermal equilibrium this is of paramount importance. Soft matter objects with internal degrees of freedom (like polymer chains, but also membranes) tend to have complex configuration spaces and large relaxation times. On the other hand, one often is not interested in following the dynamic correlation functions all the way up to the longest relaxation time. Such a situation is exactly present in the study of [15], where only the dynamics up to τ_ξ (and somewhat beyond) is needed, but not up to τ_R. One would therefore like to be able to equilibrate the system with a fast Monte Carlo algorithm, in order to shortcut the slow physical dynamics, and to use the generated configurations for starting runs with realistic (slow) dynamics, over which one averages. For dilute systems, the ideal way to do this is to *completely disregard the solvent in the equilibration procedure*. However, this requires the solvent to be *structureless*. A structured solvent (i.e. particles with some non-trivial interaction potential) always modifies the potential of mean force of the solute, and this is not known in advance. Therefore only the coupled solute-solvent system can be equilibrated "cleanly", i.e. without introducing systematic errors. Conversely, for a structureless solvent the potential of mean force of the solute is identical to the "bare" potential (i.e. the interaction without solvent). For a particle method, this means that the solvent should be an ideal gas. In this case, however, MD is not applicable since all particle trajectories are trivial, and there are no collisions. Conversely, DPD is able to simulate an ideal gas with realistic dynamics, because collisions are effectively implemented via the thermostat. Another particle method, which is also based on ideal gas particles, is multi-particle collision dynamics (MPCD) [47], where collisions are

implemented via local stochastic updating rules which conserve energy and momentum. A brief outline of this approach is included below.

Yet another approach to simulate momentum transport through the solvent is to solve the (Navier-) Stokes equation (in a deterministic or stochastic version) on a grid, and to couple an MD system for the solute to such a simulation. This yields a structureless solvent automatically. Solving the hydrodynamics can be done either by a finite-difference scheme, or by the lattice Boltzmann method (LBM) [48]. The latter has become also quite popular for soft matter systems, in particular colloidal suspensions [49–57], and will be discussed below. Compared to DPD, it has the disadvantage that the underlying theory is slightly involved, and that the coupling to the solute system (which is still simulated by some MD-like algorithm) is not a straightforward consequence of the method, but rather must be constructed by hand. The advantage, however, is that it is based on a tight data structure, with the consequence that it is computationally quite efficient, rather straightforward to implement, and ideally suited for parallel computers (the only communication is just the sending of data in the streaming step, while the collision step needs only local data). Another important advantage of a grid-based method is that thermal fluctuations may be both turned on (necessary for Brownian motion, for instance), or off (for some nonequilibrium studies like sedimentation thermal fluctuations are not needed, in particular if the solute system does not have fluctuating internal degrees of freedom). This flexibility is a quite useful aspect, and not present in particle methods, which *always* exhibit thermal fluctuations. A noise-free simulation is of course much cheaper than a noisy one, since no cumbersome averaging is necessary (except, perhaps, over initial conditions).

2.2 Dissipative Particle Dynamics (DPD)

Dissipative Particle Dynamics is essentially MD, where a momentum-conserving Langevin thermostat is added. The method is best understood by contrasting it to the older method of Stochastic Dynamics (SD) [58], which also adds a Langevin thermostat to MD, but does *not* conserve the momentum. The formal development is most transparent if we start from Hamilton's formulation of Newton's equations of motion:

$$\frac{d}{dt} q_i = \frac{\partial \mathcal{H}}{\partial p_i} \tag{33}$$

$$\frac{d}{dt} p_i = -\frac{\partial \mathcal{H}}{\partial q_i} \,, \tag{34}$$

where the q_i denote the generalized coordinates, and the p_i the generalized canonically conjugate momenta, while $\mathcal{H}$ is the Hamiltonian of the system. Adding friction and noise, we obtain the SD equations of motion:

$$\frac{d}{dt}q_i = \frac{\partial \mathcal{H}}{\partial p_i} \tag{35}$$

$$\frac{d}{dt}p_i = -\frac{\partial \mathcal{H}}{\partial q_i} - \zeta_i \frac{\partial \mathcal{H}}{\partial p_i} + \sigma_i f_i \; ; \tag{36}$$

here ζ_i is the friction coefficient for the ith degree of freedom (note that $\partial\mathcal{H}/\partial p_i$, for usual Cartesian coordinates, is nothing but the velocity), σ_i denotes the noise strength, while $\langle f_i \rangle = 0$ and $\langle f_i(t) f_j(t') \rangle = 2\delta_{ij}\delta(t-t')$. We can even allow that the friction constants ζ_i and the noise strengths σ_i depend on the coordinates q_i (but not on the momenta p_i).

We now switch to an equivalent description of the stochastic process, where we study the time evolution of the probability density in phase space, $P(\{q_i\}, \{p_i\}, t)$. This is quite analogous to switching from Hamilton's equations of motion to the Liouville equation in classical mechanics. The equation of motion for P is called the Fokker–Planck equation. Its shape can be derived directly from the Langevin equation, using a standard procedure described in textbooks on stochastic processes (see, e.g. [59, 60] or [61]). For the present case, one obtains

$$\frac{\partial}{\partial t} P = \mathcal{L} P \; , \tag{37}$$

where $\mathcal{L}$ is the Fokker–Planck operator, which is naturally decomposed into two parts,

$$\mathcal{L} = \mathcal{L}_H + \mathcal{L}_{SD} \; , \tag{38}$$

where the first part refers to the Hamiltonian part of the dynamics (it is nothing but the Liouville operator),

$$\begin{aligned} \mathcal{L}_H &= -\sum_i \frac{\partial}{\partial q_i}\frac{\partial \mathcal{H}}{\partial p_i} + \sum_i \frac{\partial}{\partial p_i}\frac{\partial \mathcal{H}}{\partial q_i} \\ &= -\sum_i \frac{\partial \mathcal{H}}{\partial p_i}\frac{\partial}{\partial q_i} + \sum_i \frac{\partial \mathcal{H}}{\partial q_i}\frac{\partial}{\partial p_i} \; , \end{aligned} \tag{39}$$

while the second part is due to friction and noise,

$$\mathcal{L}_{SD} = \sum_i \frac{\partial}{\partial p_i}\left[\zeta_i \frac{\partial \mathcal{H}}{\partial p_i} + \sigma_i^2 \frac{\partial}{\partial p_i}\right] \; . \tag{40}$$

In order to describe a system in thermal equilibrium, the Boltzmann distribution must be the stationary solution of the Fokker–Planck equation:

$$\mathcal{L} \exp\left(-\beta \mathcal{H}\right) = 0 \; , \tag{41}$$

where $\beta = 1/(k_B T)$. For the Hamiltonian part, this relation is identically fulfilled. Therefore, the condition results in

$$\sum_i \frac{\partial}{\partial p_i}\left[\zeta_i \frac{\partial \mathcal{H}}{\partial p_i} - \beta \sigma_i^2 \frac{\partial \mathcal{H}}{\partial p_i}\right] \exp\left(-\beta \mathcal{H}\right) = 0 \; . \tag{42}$$

Hence the relation

$$\sigma_i^2 = k_B T \zeta_i \tag{43}$$

must hold. This is the fluctuation-dissipation theorem (FDT), i.e. the temperature is a result of the balance between friction and noise strength. However, the momentum is *not conserved*, as one can check immediately from the equations of motion. Rather, the center of mass of the system diffuses. The algorithm also violates Galilean invariance, since it dampens the *absolute* velocities, thus labeling the "laboratory frame" as special, which is of course unphysical. These are the reasons why SD is useless for hydrodynamic simulations. It can be shown [61,62] that this unphysical behavior can be expressed in terms of a hydrodynamic screening length $\xi = [\eta/(n\zeta)]^{1/2}$. Here, we have assumed a constant friction, while n is the particle number density. The arguments to derive this are essentially the same as those presented in Sect. 1.5 for a frozen matrix of frictional obstacles.

Dissipative Particle Dynamics (DPD) has been developed to cure this problem, and to simulate hydrodynamic phenomena in fluids on a mesoscopic scale. DPD, as it is usually described in the literature, consists of two parts: (i) Introduction of very soft interparticle potentials in order to facilitate a large time step, and (ii) introduction of a Galilei invariant thermostat, which is similar to SD, but dampens *relative* velocities, and applies the stochastic kicks to *pairs* of particles such that Newton's third law (i.e. momentum conservation) is satisfied. As the procedure is also completely local, it is therefore suitable for the description of (isothermal) hydrodynamics. Unfortunately, it is often not made sufficiently clear that these two parts are *completely unrelated*, i.e. that one can use the DPD thermostat with "conventional" hard potentials, and that one can go from a working MD code to DPD, just as one would go to SD. A technical problem of typical DPD simulations is the fact that, due to the soft potentials, they are run with extremely large time steps. This results in unacceptably large discretization errors. Currently this problem is under thorough investigation [41–46]. We will from now on exclusively focus on the thermostat aspect of DPD. As Espanol and Warren [37] have shown, the structure of the FDT for DPD is very similar to the SD case. A particularly useful application of the DPD thermostat, which is just presently being appreciated, is its use in *nonequilibrium* studies like the simulation of steady-state Couette flow. Nonequilibrium steady states are characterized by a constant nonzero rate of entropy production, usually showing up as viscous heat. This produced entropy must be removed from the system, and therefore such simulations are usually coupled to a thermostat (an alternative approach, which rather removes the entropy by a Maxwell demon, has recently been developed by Müller–Plathe [63, 64]). Before the advent of DPD, it was a non-trivial problem to introduce the thermostat in such a way that it would not prefer a certain profile (so-called "profile-unbiased thermostats", see [65]). The DPD thermostat solves this problem in a very natural and straightforward way [66].

In practice, DPD simulations are done as follows: We first define two functions, $\zeta(r)$, the relative friction coefficient for particle pairs with interparticle distance r, and $\sigma(r)$, the noise strength for a stochastic kick applied to the same particle pair. We will show below that the FDT implies the relation

$$\sigma^2(r) = k_B T \zeta(r) \,, \tag{44}$$

in close analogy to SD. The function has a finite range, such that only near neighbors are taken into account.

Defining $\boldsymbol{r}_{ij} = \boldsymbol{r}_i - \boldsymbol{r}_j = r_{ij}\hat{r}_{ij}$, we then obtain the friction force on particle i by projecting the relative velocities on the interparticle axes:

$$\boldsymbol{F}_i^{(fr)} = -\sum_j \zeta(r_{ij}) \left[(\boldsymbol{v}_i - \boldsymbol{v}_j) \cdot \hat{r}_{ij}\right] \hat{r}_{ij}; \tag{45}$$

it is easy to see that the relation $\sum_i \boldsymbol{F}_i^{(fr)} = 0$ holds. Similarly, we get the stochastic forces along the interparticle axes:

$$\boldsymbol{F}_i^{(st)} = \sum_j \sigma(r_{ij})\, \eta_{ij}(t)\, \hat{r}_{ij} \,, \tag{46}$$

where the noise η_{ij} satisfies the relations $\eta_{ij} = \eta_{ji}$, $\langle \eta_{ij} \rangle = 0$, and

$$\langle \eta_{ij}(t) \eta_{kl}(t') \rangle = 2(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk})\delta(t - t'), \tag{47}$$

such that different pairs are statistically independent. As before, one easily shows $\sum_i \boldsymbol{F}_i^{(st)} = 0$. The equations of motion,

$$\frac{d}{dt}\boldsymbol{r}_i = \frac{1}{m_i}\boldsymbol{p}_i, \tag{48}$$

$$\frac{d}{dt}\boldsymbol{p}_i = \boldsymbol{F}_i + \boldsymbol{F}_i^{(fr)} + \boldsymbol{F}_i^{(st)} \,, \tag{49}$$

where m_i is the mass of the ith particle, and $\boldsymbol{p}_i$ its momentum, therefore indeed conserve the total momentum, as the conservative forces $\boldsymbol{F}_i$ satisfy Newton's third law. The Fokker–Planck operator can then be written as

$$\mathcal{L} = \mathcal{L}_H + \mathcal{L}_{DPD} \,, \tag{50}$$

where $\mathcal{L}_H$ again describes the Hamiltonian part with $\mathcal{L}_H \exp(-\beta\mathcal{H}) = 0$ (cf. 39), and $\mathcal{L}_{DPD}$ is given by

$$\begin{aligned}\mathcal{L}_{DPD} &= \sum_{ij} \zeta(r_{ij}) \hat{r}_{ij} \cdot \frac{\partial}{\partial \boldsymbol{p}_i} \left[\hat{r}_{ij} \cdot \left(\frac{\partial \mathcal{H}}{\partial \boldsymbol{p}_i} - \frac{\partial \mathcal{H}}{\partial \boldsymbol{p}_j} \right) \right] \\ &\quad - \sum_{i \neq j} \sigma^2(r_{ij}) \left(\hat{r}_{ij} \cdot \frac{\partial}{\partial \boldsymbol{p}_i} \right) \left(\hat{r}_{ij} \cdot \frac{\partial}{\partial \boldsymbol{p}_j} \right) \\ &\quad + \sum_i \sum_{j(\neq i)} \sigma^2(r_{ij}) \left(\hat{r}_{ij} \cdot \frac{\partial}{\partial \boldsymbol{p}_i} \right)^2 \\ &= \sum_i \sum_{j(\neq i)} \hat{r}_{ij} \cdot \frac{\partial}{\partial \boldsymbol{p}_i} \left[\zeta(r_{ij}) \hat{r}_{ij} \cdot \left(\frac{\partial \mathcal{H}}{\partial \boldsymbol{p}_i} - \frac{\partial \mathcal{H}}{\partial \boldsymbol{p}_j} \right) \right. \\ &\quad \left. + \sigma^2(r_{ij}) \hat{r}_{ij} \cdot \left(\frac{\partial}{\partial \boldsymbol{p}_i} - \frac{\partial}{\partial \boldsymbol{p}_j} \right) \right] . \end{aligned} \tag{51}$$

In the stochastic term, we have first taken into account the off-diagonal terms (cross-correlations, which are actually anti-correlations between the neighbors). The prefactors for the diagonal terms are given by the sum of all the mean square noise strengths from all the neighbors. Applying this operator to $\exp(-\beta\mathcal{H})$, we find that the FDT is satisfied if $\sigma^2(r) = k_B T \zeta(r)$.

2.3 Multi-Particle Collision Dynamics (MPCD)

As already mentioned, DPD simulations can be run for the special case of vanishing interaction potential, i.e. in the ideal gas limit. If we discretize this procedure in terms of, say, the Verlet algorithm, we arrive at a method where free particle propagation (i.e. update in real space without update in momentum space) alternates with "collisions" (update in momentum space without update in real space). Due to their dissipative nature, these DPD collisions conserve momentum but not the energy.

Malevanets and Kapral [47] have introduced a method which is also based on collisions of ideal gas particles. However, the collisions are now implemented by a simple Monte Carlo procedure such that *both* momentum and energy are conserved. Starting from a set of particle coordinates $\boldsymbol{r}_i$ and a set of particle velocities $\boldsymbol{v}_i$, $i = 1, \ldots, N$, one first performs a streaming step (free propagation by a time step h)

$$\boldsymbol{r}_i(t+h) = \boldsymbol{r}_i(t) + h\boldsymbol{v}_i(t) \; . \tag{52}$$

This is followed by a collision step, which is facilitated by sub-dividing the simulation box into sub-boxes. For each sub-box, one determines the set of particles residing in it. For one particular sub-box, let these particles be enumerated by $i = 1, \ldots, n$. These "collide" with each other by the following procedure:

- Determine the local center-of-mass velocity:

$$\boldsymbol{v}_{CM} = \frac{1}{n}\sum_{i=1}^{n} \boldsymbol{v}_i. \tag{53}$$

- For each particle in the sub-box, perform a Galileo transformation into the local center-of-mass system:

$$\tilde{\boldsymbol{v}}_i = \boldsymbol{v}_i - \boldsymbol{v}_{CM}. \tag{54}$$

- Within the local center-of-mass system, rotate all velocities within the sub-box by a random rotation matrix $\overleftrightarrow{R}$:

$$\tilde{\boldsymbol{v}}_i{}' = \overleftrightarrow{R}\, \tilde{\boldsymbol{v}}_i. \tag{55}$$

- Transform back into the "laboratory" system:

$$\boldsymbol{v}_i' = \tilde{\boldsymbol{v}}_i{}' + \boldsymbol{v}_{CM}. \tag{56}$$

One sees immediately that this procedure satisfies locality as well as the conservation of mass, momentum, and energy. Ihle and Kroll [67] have pointed out that it is necessary to randomly shift the sub-boxes in order to avoid spurious effects and to restore full Galileo invariance. Furthermore, the fact that the dynamics is so simple makes it possible to derive analytic expressions for transport coefficients [68–70].

2.4 Lattice Boltzmann (LB)

The lattice Boltzmann method (LBM) works quite differently. Essentially, the method is the simulation of a fully discretized version of the (linearized) Boltzmann equation known from the kinetic theory of gases. One starts from a regular lattice (usually a simple-cubic lattice) with lattice spacing a; $\boldsymbol{r}$ denotes its sites. Furthermore, we introduce a finite (small) set of (dimensionless) vectors $\boldsymbol{c}_i$, such that $a\boldsymbol{c}_i$ is a vector connecting two sites on the lattice. The set should be consistent with the point symmetry of the lattice. For example, on the simple-cubic lattice one would have six vectors $\boldsymbol{c}_i$ connecting to the nearest neighbors, and another twelve vectors to the next-nearest neighbors. Time is discretized in terms of a time step h, and the model allows only for a finite set of velocities. These are the vectors $(a/h)\boldsymbol{c}_i$. An object residing on a certain lattice site $\boldsymbol{r}$, and having the velocity $(a/h)\boldsymbol{c}_i$, would thus be moved to site $\boldsymbol{r} + a\boldsymbol{c}_i$ within one time step. A commonly used model is the 18-velocity model, where the vectors $\boldsymbol{c}_i$ correspond to the nearest and next-nearest neighbors. Sometimes an additional velocity $\boldsymbol{c}_i = 0$ is included (19-velocity model); this is however not necessary for simulating incompressible flow. The algorithm now works with real-valued variables $n_i(\boldsymbol{r}, t)$, denoting the "number of particles" which reside on site $\boldsymbol{r}$ at time t and have the velocity $(a/h)\boldsymbol{c}_i$. Denoting the particle mass with m, we find for the mass density at site $\boldsymbol{r}$ at time t

$$\rho(\boldsymbol{r},t) = \frac{m}{a^3}\sum_i n_i(\boldsymbol{r},t) \tag{57}$$

and for the momentum density

$$\boldsymbol{j}(\boldsymbol{r},t) = \frac{m}{a^2 h}\sum_i n_i(\boldsymbol{r},t)\boldsymbol{c}_i \ . \tag{58}$$

We can also introduce the streaming velocity $\boldsymbol{u}$ at site $\boldsymbol{r}$ via $\boldsymbol{u} = \boldsymbol{j}/\rho$. It should be noted that in many descriptions of the method the parameters m, a and h are set to unity, thus defining the unit system of the method. However, when coupling the LBM to an MD system, the latter has its own unit system. We prefer to use a unit system built upon MD, and for this purpose we need to keep the parameters. Furthermore, it should be noted that we do not consider the energy density. In this chapter, we only consider LBMs with mass and momentum conservation, while energy conservation (heat conduction etc.) is *not* taken into account. LBMs with proper inclusion of the energy have been developed [71], but are more complicated.

Now, the algorithm proceeds via the following steps:

1. Starting from the variables n_i, one calculates the hydrodynamic variables ρ and $\boldsymbol{j}$ on each lattice site.
2. From ρ and $\boldsymbol{j}$, one calculates a local pseudo-equilibrium distribution n_i^{eq}. It should be stressed that this is done for each site separately. Since the variables ρ and $\boldsymbol{j}$ differ from site to site, one has a different distribution n_i^{eq} on each site. The kinetic-theory analogue would be a Maxwell-Boltzmann velocity distribution centered around the hydrodynamic streaming velocity at position $\boldsymbol{r}$. Since n_i^{eq} and n_i correspond to the *same* hydrodynamic variables, we have $\sum_i n_i = \sum_i n_i^{eq}$ and $\sum_i n_i \boldsymbol{c}_i = \sum_i n_i^{eq}\boldsymbol{c}_i$ at each site.
3. Relaxation ("collisions"): The velocity distribution on the site is rearranged in order to bring it closer to the local equilibrium of that site. This is done via a linear process:

$$n_i \to n_i + \sum_j L_{ij}(n_j - n_j^{eq}) \ . \tag{59}$$

   In many cases, the matrix L_{ij} is just a multiple of the unit matrix. These are the so-called "lattice BGK" (Bhatnagar-Gross-Krook) methods. However, working with a nontrivial matrix causes no practical difficulties and allows one to get rid of non-hydrodynamic modes quickly [50]. In order to ensure mass and momentum conservation, the matrix should satisfy the conditions $\sum_i L_{ij} = 0$ and $\sum_i \boldsymbol{c}_i L_{ij} = 0$.
4. Streaming: The populations are displaced to new sites according to their velocities:

$$n_i(\boldsymbol{r},t) \to n_i(\boldsymbol{r} + a\boldsymbol{c}_i, t+h) \ . \tag{60}$$

   This is the only step which is not completely local.

Further specification of the algorithm requires to give prescriptions for the calculation of n_i^{eq}, and of the relaxation procedure. A common procedure is to use the polynomial ansatz [50]

$$n_i^{eq} = \frac{a^3}{m}\rho\left(A_i + B_i \boldsymbol{c}_i \cdot \boldsymbol{u}\frac{h}{a} + C_i u^2 \frac{h^2}{a^2} + D_i(\boldsymbol{c}_i \cdot \boldsymbol{u})^2 \frac{h^2}{a^2}\right) , \tag{61}$$

where symmetry requires that A_i should only depend on the neighbor shell, but not on the direction within it, and the same holds also for B_i, C_i, D_i. The 18-velocity model thus has eight coefficients. These are determined via the following requirements:

- n_i^{eq} should produce the correct hydrodynamic variables ρ and $\boldsymbol{j}$, as mentioned above.
- The stress tensor constructed from n_i^{eq},

$$\overleftrightarrow{\Pi}^{eq} = \frac{m}{a^3}\frac{a^2}{h^2}\sum_i n_i^{eq} \boldsymbol{c}_i \otimes \boldsymbol{c}_i , \tag{62}$$

  should have the hydrodynamic form

$$\overleftrightarrow{\Pi}^{eq} = \rho c_s^2 \overleftrightarrow{1} + \rho \boldsymbol{u} \otimes \boldsymbol{u} ; \tag{63}$$

  here we have assumed the equation of state of an ideal gas with sound velocity c_s (other equations of state can be implemented [72]).
- The viscosity tensor (which, on a cubic lattice, will in general be a fourth-rank tensor with cubic anisotropy) should exhibit the full rotational symmetry, such that there are only shear and bulk viscosity. This is the main reason why nearest and next-nearest neighbor shells are used: The coefficients can be adjusted in such a way that the anisotropic contributions from the two shells just cancel.
- For $\boldsymbol{u} = 0$ both shells should contain the same number of particles. This is useful for numerical stability [50].

These conditions suffice to determine the coefficients and the parameter c_s uniquely.

The relaxation operator is determined via the following considerations: Apart from mass and momentum conservation, which already give four conditions, one observes that the linear relaxation of n_i towards n_i^{eq} corresponds to a linear relaxation of the stress tensor

$$\overleftrightarrow{\Pi} = \frac{m}{a^3}\frac{a^2}{h^2}\sum_i n_i \boldsymbol{c}_i \otimes \boldsymbol{c}_i \tag{64}$$

towards $\overleftrightarrow{\Pi}^{eq}$. We now require that this process exhibits two relaxation rates, one for the trace of $\overleftrightarrow{\Pi}$, corresponding to the bulk viscosity, and one for the

trace-free part, corresponding to the shear viscosity. These parameters can thus be freely adjusted. Finally, we require [50] that the higher-order moments (non-hydrodynamic modes) are immediately removed after the relaxation process (this corresponds to eigenvalues -1). Under these circumstances, it turns out that the calculation of the new population (after the relaxation step) does not even require the implementation of the L_{ij} matrix. One rather has to simply update the pressure tensor, using the prescribed rates, and to use that result to calculate the new populations (again, the coefficients $A_i, \ldots, D_i$ are used) [50]. Via a Chapman–Enskog expansion one can show that this procedure yields hydrodynamic behavior in the macroscopic limit [50] *if the flow is incompressible, and the flow velocity is small compared to the sound velocity.* One particular advantage of the formulation based on the stress tensor is that the inclusion of thermal noise is quite straightforward: According to linear fluctuating hydrodynamics [73], the noise term occurs in the stress tensor, and therefore it can be directly added in the simulation code. For further details, see the original literature [50]. Recently, Ladd's procedure to include thermal noise has been extended, in order to produce better results for smaller length and time scales [74].

When coupling this to a system of Brownian particles, one can use two methods: The original approach by Ladd [50,51] for colloidal suspensions was to use extended particles with a surface, and to implement a bounce-back rule to simulate the modification of the flow, plus the momentum transfer onto the particle. Combined with a lubrication correction for suspensions at high densities, this approach has produced excellent results for suspensions with hydrodynamic interactions [53].

For polymer solutions, we found a point-particle approach [8, 75] simpler and more efficient: While the solvent is run via the stochastic version of the LBM, the polymer system is simulated by MD augmented with friction and noise as in SD. However, the friction force is not $-\zeta \boldsymbol{v}$ ($\boldsymbol{v}$ particle velocity), but rather $-\zeta(\boldsymbol{v} - \boldsymbol{u})$, where $\boldsymbol{u}$ is the flow velocity at the position of the monomer, obtained via linear interpolation from the surrounding lattice sites. This determines the momentum transfer onto the particle which has come from the solvent. Momentum conservation requires that this momentum is subtracted from the fluid. Details of this latter subtraction are not important; we used a procedure where we distributed the momentum transfer onto the surrounding sites using the same weights as the initial interpolation procedure. On each site, we then updated the n_i by requiring that the distance to n_i^{eq} remained unchanged. It can be shown that the coupled system does satisfy the FDT. Since locality, mass conservation, and momentum conservation are fulfilled, this procedure simulates hydrodynamic interactions faithfully, while being roughly 20 times faster than the analogous MD system with hard solvent particles. The lattice spacing was set roughly equal to the bond length; this is necessary to resolve the hydrodynamic interaction down to the relevant scales. We have recently shown [76, 77] that this approach can also be used

to simulate colloidal particles, which are modeled as an arrangement of force centers like a "raspberry".

The friction coefficient ζ should be called "bare" friction coefficient, since the long-time single-particle mobility differs from $1/\zeta$ as a result of the long time tail. The correction can be quite strong, and actually depends on the lattice spacing a. This can be shown by the following consideration: We drag a particle with constant average velocity $\boldsymbol{v}$ and constant average force $\boldsymbol{F}$ through a fluid globally at rest. Our simulation procedure tells us that the force should be $\boldsymbol{F} = \zeta(\boldsymbol{v} - \boldsymbol{u})$, $\boldsymbol{u}$ being the flow velocity on the surrounding lattice sites, which are, on average, a distance of order a away. The Oseen tensor, in turn, tells us that u should be of order $u \sim F/(\eta a)$ or $u = F/(g\eta a)$, where g is some numerical coefficient. Combining these equations, we find for the mobility

$$\mu = \frac{1}{\zeta} + \frac{1}{g\eta a}; \tag{65}$$

this relation has been checked numerically [8]. The lattice thus provides a Stokes-like contribution to the mobility. It thus not only discretizes the hydrodynamics, but also regularizes it, i.e. it naturally cures the pathology that a point particle does not exist (note that in the continuum limit $a \to 0$ one would obtain an infinite mobility!). Since a is just a discretization parameter, the only conclusion is that ζ does not have any physical meaning. Rather, for comparing with experiments one should look at the "dressed" mobility μ.

3 Some Final Remarks

Although the presented material is highly selective, strongly reflecting my own research, I hope the present chapter has given a slight glimpse at the problems one encounters when simulating systems with hydrodynamic interactions, and also at the strategies which have been developed to cope with them. The development of so-called "mesoscopic" simulation methods ("somewhere between Molecular Dynamics and computational fluid dynamics (CFD)") for soft matter systems, with emphasis on hydrodynamics, is a quite active field of current research, and far from being closed. At the same time, applications are already quite broad and continue to grow. Typical systems are polymers, colloids, liquid crystals, membranes, multiphase flows (e.g. spinodal decomposition of binary mixtures), electro-rheological systems, microfluidic devices, and biophysical systems, like, e.g., models for swimming bacteria. Furthermore, these methods (in particular the LBM) enter more and more the field of classical CFD, with applications like fluid turbulence or automotive engineering.

When asked about a judgmental statement about the presented methods, I feel very reluctant to give a recommendation. Reducing the various approaches to their bare essentials, it turns out that they all are not fundamentally different from each other. Although detailed benchmark comparisons have, to my knowledge, so far never been done, it is hard to conceive that the methods

should differ very strongly in their computational performance. I personally like lattice methods, because they are easy to parallelize, and because thermal fluctuations can be turned on and off. However, for many interesting applications one may well get away with just a single processor, and for many soft matter systems, the inclusion of thermal fluctuations is anyways needed. Therefore, my personal opinion is that the choice of the method is, to a large extent, a matter of taste. The only recommendation which I can give is to try to understand as much physics as possible before running the simulation, and then ask the questions: (i) Which conservation laws (mass, momentum, energy) are needed to describe the physics correctly? (ii) Is the inclusion of thermal fluctuations really needed? (iii) Are large systems on a parallel machine needed? (iv) How much (molecular) structure of the fluid is needed? Starting from the answers to these questions, one should then be able to pick a suitable method and to construct a useful computational model.

Acknowledgments

The present paper is a corrected and extended version of a proceedings article which I wrote on occasion of the spring school *Computational Soft Matter: From Synthetic Polymers to Proteins*, organized by the Neumann Institute for Computing, Jülich (NIC) in 2004. I cordially thank the publisher (NIC) of the proceedings volume (same title, editors: N. Attig, K. Binder, H. Grubmüller, K. Kremer, ISBN 3-00-012641-4) for the kind permission to reuse this material.

Furthermore, I thank Patrick Ahlrichs, Ralf Everaers, Bo Liu, Igor Pasichnyk, Vladimir Lobaskin, Kurt Kremer, and Thomas Soddemann for a fruitful collaboration on the issues covered in the text, and Tony Ladd for stimulating discussions on mesoscopic simulation methods.

Last not least it is my pleasure to thank Christine Peter–Tittelbach, plus two further anonymous referees, for a careful and critical reading of the manuscript and useful suggestions.

References

1. M. Doi and S. F. Edwards (1986) *The Theory of Polymer Dynamics*. Oxford University Press, Oxford
2. P. E. Rouse (1953) A theory of the linear viscoelastic properties of dilute solutions of coiling polymers. *J. Chem. Phys.* **21**, pp. 1272–1280
3. B. H. Zimm (1956) Dynamics of polymer molecules in dilute solution: Viscoelasticity, flow birefringence and dielectric loss. *J. Chem. Phys.* **24**, pp. 269–278
4. C. Pierleoni and J. P. Ryckaert (1992) Molecular dynamics investigation of dynamic scaling for dilute polymer solutions in good solvent conditions. *J. Chem. Phys.* **96**, pp. 8539–8551

5. W. Smith and D. C. Rapaport (1992) Molecular dynamics simulation of linear polymers in a solvent. *Mol. Sim.* **9**, pp. 25–39
6. B. Dünweg and K. Kremer (1993) Molecular dynamics simulation of a polymer chain in solution. *J. Chem. Phys.* **99**, pp. 6983–6997
7. A. G. Schlijper, P. J. Hoogerbrugge, and C. W. Manke (1995) Computer simulation of dilute polymer solutions with the dissipative particle dynamics method. *J. Rheol.* **39**, pp. 567–579
8. P. Ahlrichs and B. Dünweg (1999) Simulation of a single polymer chain in solution by combining lattice Boltzmann and molecular dynamics. *J. Chem. Phys.* **111**, pp. 8225–8239
9. N. A. Spenley (2000) Scaling laws for polymers in dissipative particle dynamics. *Europhys. Lett.* **49**, pp. 534–540
10. A. Malevanets and J. M. Yeomans (2000) Dynamics of short polymer chains in solution. *Europhys. Lett.* **52**, pp. 231–237
11. K. Kremer and G. S. Grest (1990) Dynamics of entangled linear polymer melts: A molecular-dynamics simulation. *J. Chem. Phys.* **92**, pp. 5057–5086
12. W. Paul, K. Binder, D. W. Heermann, and K. Kremer (1991) Crossover scaling in semidilute polymer solutions: A Monte Carlo test. *J. Phys. II (Paris)* **1**, pp. 37–60
13. W. Paul, K. Binder, D. W. Heermann, and K. Kremer (1991) Dynamics of polymer solutions and melts. Reptation predictions and scaling of relaxation times. *J. Chem. Phys.* **95**, pp. 7726–7740
14. M. Pütz, K. Kremer, and G. S. Grest (2000) What is the entanglement length in a polymer melt? *Europhys. Lett.* **49**, pp. 735–741
15. P. Ahlrichs, R. Everaers, and B. Dünweg (2001) Screening of hydrodynamic interactions in semidilute polymer solutions: A computer simulation study. *Phys. Rev. E* **64**, p. 040501 (R)
16. R. Chang and A. Yethiraj (2001) Solvent effects on the collapse dynamics of polymers. *J. Chem. Phys.* **114**, pp. 7688–7699
17. C. F. Abrams, N. K. Lee, and S. P. Obukhov (2002) Collapse dynamics of a polymer chain: Theory and simulation. *Europhys. Lett.* **59**, pp. 391–397
18. N. Kikuchi, A. Gent, and J. M. Yeomans (2002) Polymer collapse in the presence of hydrodynamic interactions. *Europ. Phys. J. E* **9**, pp. 63–66
19. P. G. de Gennes (1979) *Scaling Concepts in Polymer Physics.* Cornell University Press, Ithaca
20. A. Y. Grosberg and A. R. Khokhlov (1994) *Statistical Physics of Macromolecules.* AIP Press, New York
21. B. Alder and T. Wainwright (1970) Decay of the velocity autocorrelation function. *Phys. Rev. A* **1**, pp. 18–21
22. A. Kopf, B. Dünweg, and W. Paul (1997) Dynamics of polymer "isotope" mixtures: Molecular dynamics simulation and Rouse model analysis. *J. Chem. Phys.* **107**, pp. 6945–6955
23. W. Paul, G. D. Smith, D. Y. Yoon, B. Farago, S. Rathgeber, A. Zirkel, L. Willner, and D. Richter (1998) Chain motion in an unentangled polyethylene melt: A critical test of the Rouse model by molecular dynamics simulations and neutron spin echo spectroscopy. *Phys. Rev. Lett.* **80**, pp. 2346–2349
24. W. Paul (2002) Anomalous diffusion in polymer melts. *Chemical Physics* **284**, pp. 59–66

25. V. G. Rostiashvili, M. Rehkopf, and T. A. Vilgis (1998) Dynamics of polymeric manifolds in melts: The Hartree approximation. *Europ. Phys. J. B* **6**, pp. 497–501
26. M. Guenza (2002) Cooperative dynamics in unentangled polymer fluids. *Phys. Rev. Lett.* **88**, p. 025901
27. B. Liu and B. Dünweg (2003) Translational diffusion of polymer chains with excluded volume and hydrodynamic interactions by Brownian dynamics simulation. *J. Chem. Phys.* **118**, pp. 8061–8072
28. S. F. Edwards and K. F. Freed (1974) Theory of the dynamical viscosity of polymer solutions. *J. Chem. Phys.* **61**, pp. 1189–1202
29. K. F. Freed and S. F. Edwards (1974) Polymer viscosity in concentrated solutions. *J. Chem. Phys.* **61**, pp. 3626–3633
30. P. G. de Gennes (1976) Dynamics of entangled polymer solutions. 2. Inclusion of hydrodynamic interactions. *Macromolecules* **9**, pp. 594–598
31. P. Wiltzius, H. R. Haller, and D. S. Carnell (1984) Dynamics of long-wavelength concentration fluctuations in solutions of linear polymers. *Phys. Rev. Lett.* **53**, pp. 834–837
32. P. Wiltzius and D. S. Carnell (1986) Wave-vector dependence of the initial decay rate of fluctuations in polymer solutions. *Phys. Rev. Lett.* **56**, pp. 61–64
33. D. Richter, K. Binder, B. Ewen, and B. Stühn (1984) Screening of hydrodynamic interactions in dense polymer solutions: A phenomenological theory and neutron scattering investigations. *J. Phys. Chem.* **88**, pp. 6618–6633
34. A. S. Sangani and G. Mo (1996) An O(N) algorithm for Stokes and Laplace interactions of particles. *Phys. Fluids* **8**, pp. 1990–2010
35. P. J. Hoogerbrugge and J. M. V. A. Koelman (1992) Simulating microscopic hydrodynamic phenomena with dissipative particle dynamics. *Europhys. Lett.* **19**, pp. 155–160
36. J. M. V. A. Koelman and P. J. Hoogerbrugge (1993) Dynamic simulations of hard-sphere suspensions under steady shear. *Europhys. Lett.* **21**, pp. 363–368
37. P. Español and P. Warren (1995) Statistical mechanics of dissipative particle dynamics. *Europhys. Lett.* **30**, pp. 191–196
38. P. Español (1995) Hydrodynamics from dissipative particle dynamics. *Phys. Rev. E* **52**, pp. 1734–1742
39. R. D. Groot and P. B. Warren (1997) Dissipative particle dynamics: Bridging the gap between atomistic and mesoscopic simulation. *J. Chem. Phys.* **107**, pp. 4423–4435
40. P. Español (1998) Fluid particle model. *Phys. Rev. E* **57**, pp. 2930–2948
41. I. Pagonabarraga, M. J. H. Hagen, and D. Frenkel (1998) Self-consistent dissipative particle dynamics algorithm. *Europhys. Lett.* **42**, pp. 377–382
42. J. B. Gibson, K. Chen, and S. Chynoweth (1999) The equilibrium of a velocity-Verlet type algorithm for DPD with finite time steps. *Int. J. Mod. Phys. C* **10**, pp. 241–261
43. G. Besold, I. Vattulainen, M. Karttunen, and J. M. Polson (2000) Towards better integrators for dissipative particle dynamics simulations. *Phys. Rev. E* **62**, pp. R7611–R7614
44. I. Vattulainen, M. Karttunen, G. Besold, and J. M. Polson (2002) Integration schemes for dissipative particle dynamics simulations: From softly interacting systems towards hybrid models. *J. Chem. Phys.* **116**, pp. 3967–3979

45. P. Nikunen, M. Karttunen, and I. Vattulainen (2003) How would you integrate the equations of motion in dissipative particle dynamics simulations? *Comp. Phys. Comm.* **153**, pp. 407–423
46. T. Shardlow (2003) Splitting for dissipative particle dynamics. *SIAM J. Sci. Comp.* **24**, pp. 1267–1282
47. A. Malevanets and R. Kapral (1999) Mesoscopic model for solvent dynamics. *J. Chem. Phys.* **110**, pp. 8605–8613
48. S. Succi (2001) The Lattice Boltzmann Equation for Fluid Dynamics and Beyond. Oxford University Press, Oxford
49. A. J. C. Ladd (1993) Short-time motion of colloidal particles: Numerical simulation via a fluctuating lattice-Boltzmann equation. *Phys. Rev. Lett.* **70**, pp. 1339–1342
50. A. J. C. Ladd (1994) Numerical simulations of particulate suspensions via a discretized Boltzmann equation. 1. Theoretical foundation. *J. Fluid Mech.* **271**, pp. 285–309
51. A. J. C. Ladd (1994) Numerical simulations of particulate suspensions via a discretized Boltzmann equation. 2. Numerical results. *J. Fluid Mech.* **271**, pp. 311–339
52. A. J. C. Ladd and R. Verberg (2001) Lattice-Boltzmann simulations of particle-fluid suspensions. *J. Stat. Phys.* **104**, pp. 1191–1251
53. A. J. C. Ladd, H. Gang, J. X. Zhu, and D. A. Weitz (1995) Temporal and spatial dependence of hydrodynamic correlations: Simulation and experiment. Phys. Rev. E. **52**, pp. 6550–6572
54. C. P. Lowe, D. Frenkel, and A. J. Masters (1995) Long-time tails in angular momentum correlations. J. Chem. Phys. **103**, pp. 1582–1587
55. M. H. J. Hagen, I. Pagonabarraga, C. P. Lowe, and D. Frenkel (1997) Algebraic decay of velocity fluctuations in a confined fluid. Phys. Rev. Lett. **78**, pp. 3785–3788
56. M. H. J. Hagen, D. Frenkel, and C. P. Lowe (1999) Rotational diffusion in dense suspensions. *Physica A* **272**, pp. 376–391
57. M. W. Heemels, M. H. J. Hagen, and C. P. Lowe (2000) Simulating solid colloidal particles using the lattice-Boltzmann method. *J. Comput. Phys.* **164**, pp. 48–61
58. T. Schneider and E. Stoll (1978) Molecular-dynamics study of a three-dimensional one-component model for distortive phase transitions. *Phys. Rev. B* **17**, pp. 1302–1322
59. S. Chandrasekhar (1943) Stochastic problems in physics and astronomy. *Rev. Mod. Phys.* **15**, pp. 1–89
60. H. Risken (1984) The Fokker–Planck Equation. Springer–Verlag, Berlin.
61. B. Dünweg (2003) Langevin methods. In B. Dünweg, D. P. Landau, A. I. Milchev, eds., Computer Simulations of Surfaces and Interfaces, Kluwer, Dordrecht. pp. 77–92
62. B. Dünweg (1993) Molecular dynamics algorithms and hydrodynamic screening. *J. Chem. Phys.* **99**, pp. 6977–6982
63. F. Müller-Plathe (1997) A simple nonequilibrium molecular dynamics method for calculating the thermal conductivity. *J. Chem. Phys.* **106**, pp. 6082–6085
64. F. Müller-Plathe (1999) Reversing the perturbation in nonequilibrium molecular dynamics: An easy way to calculate the shear viscosity of fluids. *Phys. Rev. E* **59**, pp. 4894–4898
65. D. J. Evans and G. P. Morriss (1990) Statistical Mechanics of Nonequilibrium Liquids. Academic Press, London

66. T. Soddemann, B. Dünweg, and K. Kremer (2003) Dissipative particle dynamics: A useful thermostat for equilibrium and nonequilibrium molecular dynamics simulations. *Phys. Rev. E* **68**, p. 046702
67. T. Ihle and D. M. Kroll (2003) Stochastic rotation dynamics. I. Formalism, Galilean invariance, and Green-Kubo relations. *Phys. Rev. E* **67**, p. 066705
68. T. Ihle and D. M. Kroll (2003) Stochastic rotation dynamics. II. Transport coefficients, numerics, and long-time tails. *Phys. Rev. E* **67**, p. 066706
69. N. Kikuchi, C. M. Pooley, J. F. Ryder, and J. M. Yeomans (2003) Transport coefficients of a mesoscopic fluid dynamics model. *J. Chem. Phys.* **119**, pp. 6388–6395
70. T. Ihle, E. Tüzel, and D. M. Kroll (2004) Resummed Green-Kubo relations for a fluctuating fluid-particle model. *Phys. Rev. E* **70**, p. 035701(R)
71. T. Ihle and D. M. Kroll (2000) Thermal lattice-Boltzmann method for non-ideal gases with potential energy. *Comp. Phys. Comm.* **129**, pp. 1–12
72. M. R. Swift, S. E. Orlandini, W. R. Osborn, and J. M. Yeomans (1996) Lattice Boltzmann simulations of liquid-gas and binary fluid systems. *Phys. Rev. E* **54**, pp. 5041–5052
73. L. D. Landau and E. M. Lifshitz (1959) *Fluid mechanics.* Addison-Wesley, Reading
74. R. Adhikari, K. Stratford, M. E. Cates, and A. J. Wagner (2005) Fluctuating lattice Boltzmann. *Europhys. Lett.* **71**, pp. 473–479
75. P. Ahlrichs and B. Dünweg (1998) Lattice Boltzmann simulation of polymer-solvent systems. *Int. J. Mod. Phys. C* **9**, pp. 1429–1438
76. V. Lobaskin and B. Dünweg (2004) A new model for simulating colloidal dynamics. *New J. Phys.* **6**, p. 54
77. V. Lobaskin, B. Dünweg, and C. Holm (2004) Electrophoretic mobility of a charged colloidal particle: A computer simulation study. *J. Physics: Cond. Matt.* **16**, pp. S4063–S4073

Polymer Dynamics: Long Time Simulations and Topological Constraints

K. Kremer

Max-Planck-Institut für Polymerforschung, Ackermannweg 10, 55128 Mainz, Germany
kremer@mpip-mainz.mpg.de



Kurt Kremer

K. Kremer: *Polymer Dynamics: Long Time Simulations and Topological Constraints*, Lect. Notes Phys. **704**, 341–378 (2006)
DOI 10.1007/3-540-35284-8_14
© Springer-Verlag Berlin Heidelberg 2006

1 Introduction 343

2 Polymer Dynamics and Network Elasticity 345

3 Theoretical Concepts 348

3.1 Unentangled Chains – Rouse Regime 348
3.2 Entangled Chains 350

4 Simulation Models, Equilibrated Melts 354

4.1 Preparing an Equilibrated Melt, Specific Systems 356

5 Entanglement Analysis 360

5.1 "Visual" Inspection of Melts and Networks 360
5.2 Analysis of Chain Motion 362
5.3 Structure and Property Relations: Specific Polymers 366

6 Primitive Path Analysis: PPA 369

7 Summary 374

8 Annotation 374

References 374

Simulations are very versatile tools to study the relaxation and dynamics of polymer melts and networks. The fact that polymer chains cannot pass through each other poses special difficulties for analytic theories, while on the other hand many experiments are dominated by this fact. The contribution discusses some basic concepts and conditions and ways to study such problems by computer simulations.

1 Introduction

This book deals with several aspects of soft matter science, which includes classical polymer physics problems as well as a variety of topics in biophysics, such as membranes or proteins. In all these cases we have to study and understand the properties of rather large molecules, which in the case of proteins and especially polymers can contain many thousands of atoms. In this context classical synthetic polymers are a prototypical class of systems, as their molecular structure in most cases is simpler than those of typical biopolymers. In the simplest case a polymer is a long chain molecule of identical repeat units (beads or monomers). Of course modern polymer chemistry goes well beyond that simple scheme. On the other hand many characteristic problems to treat such systems by computer simulations can be illustrated best for simple polymers. Especially the dynamical and relaxational properties of a dense melt of polymers still pose many scientific as well as technical problems. Therefore I will in the present contribution mostly restrict myself to the example of polymer melt dynamics, which is one of the most prominent classical polymer science problems, where significant progress has been made over the last years, especially also since the 1995 Como Summer School on Monte Carlo and Molecular Dynamics of Condensed Matter Systems.

In the present chapter I focus on dense polymeric systems, i.e. the consequences of the fact that chains cannot pass through but only along each other for the dynamics of such systems. To do this I will shortly review the background coming from short chains, the so called Rouse model, and then discuss similarities between polymer melts and networks, as they are observed experimentally and as they can be studied in different details in a simulation. This leads to the reptation or tube model. Since most of the information gained by simulation is complementary to typical (scattering) experiments an altogether rather coherent picture has emerged over the many years of research [1–3]. In the case of (semi-)dilute solutions hydrodynamic effects play a major role, which are completely screened in dense systems. These are discussed in a special chapter by B. Dünweg [4], who deals with the numerical methods and theoretical concepts needed there.

Dense polymer systems such as melts, glasses and crosslinked melts or solutions (networks such as rubber and gels) are very complex materials. Besides the local chemical interactions and density correlations, which are common to all liquids and disordered solids the global chain conformations and the chain

connectivity play a decisive role for many physical properties. Local interactions determine the liquid structure on the scale of a few Å or at most a few nm. To study such local properties the chemical details of the chains have to be taken into account in the simulations and atomistically detailed melts are considered. When we however look at the dynamics of a polymer chain in such a melt, local interactions determine the packing and the bead friction but not the generic properties [5]. It is the main focus of the present contribution to discuss generic aspects common to all polymers and then lateron go back to the question to what extent chemistry specific aspects play a role or make a difference. The consequences of the latter are also termed as structure property relations (SPR) in more applied research [5, 6].

To stick to simple situations we consider polymer melts or networks where the chains are all identical. They can be characterized by an overall particle density ρ and a number of monomers N per chain. The overall extension of the chains is well characterized by the properties of random walks [7–9]. With ℓ being the average bond length we then have (for $N \gg 1$) for the mean square end to end distance

$$\langle R^2(N)\rangle = \ell_K \ell(N-1) \approx \ell_K \ell N \tag{1}$$

and $\langle R_G^2(N)\rangle = \frac{1}{6}\langle R^2(N)\rangle$ for the radius of gyration, which is the mean squared distance of the beads from the center of mass of the chain, respectively. ℓ_K is the Kuhn length and a measure for the stiffness of the chain. This gives an average volume each chain covers of

$$V \propto \langle R^2(N)\rangle^{3/2} \sim N^{3/2} \tag{2}$$

leading asymptotically to a vanishing self density of the chains in a melt. In order to pack beads at a monomer density ρ, which is a typical density of a molecular liquid, $0(N^{1/2})$ other chains share the very same volume of the chain and their conformations strongly interpenetrate each other. These other chains effectively screen the long range excluded volume interaction, since the individual monomer cannot distinguish, whether a non-bonded neighbor monomer belongs to the same chain or not. This leads to the above mentioned random walk structure, unlike dilute solutions, where the chains are more extended and display the so called self avoiding walk behavior with different scaling exponents [7]. This general property is firmly established by experiment and many simulations [10].

On length scales much larger than R_G^2 polymers diffuse as a whole and the motion is well described by standard diffusion. However over distances up to the order of the chain size, the motion of a polymer chain is more complex, even though hydrodynamic interactions are screened and do not play a role. On smaller length scales, the random diffusive motion of a monomer is constrained by the chain connectivity and the interaction with other monomers. To a very good first approximation, the other chains can be viewed as providing a viscous background and a heat bath. This certainly is a drastic

oversimplification, which ignores all correlations due to the structure of the surrounding. The advantage of this simplification is that the Langevin dynamics of a single chain of point masses connected by harmonic springs can be solved exactly [1]. This was first done in a seminal paper by Rouse [11] and about the same time in a similar fashion by Bueche [12]. In this model, which is commonly referred to as the Rouse model, the diffusion constant of the chain $D \sim N^{-1}$, the longest relaxation time $\tau_d \sim N^2$ and the viscosity $\eta \sim N$. This describes the dynamics of a melt of relatively short chains, meaning molecular weights of e.g. $M \leq 20000$ for polystyrene [PS, $M_{mon} = 104$] or $M \leq 2000$ for polyethylene [PE, $M_{mon} = 14$], both qualitatively and quantitatively almost perfectly, though the reason is still not well understood. Only recently some deviations have been observed [13]. The effects are rather subtle and would require a detailed discussion beyond the scope of this chapter. For longer chains, the motion of the chains are observed to be significantly slower. Experiments show a dramatic decrease in the diffusion constant, $D \sim N^{-2.4}$ [14], and an increase of the viscosity towards $\eta \sim N^{3.4}$ [1]. The time-dependent elastic modulus $G(t)$ exhibits a solid or rubber-like plateau at intermediate times before decaying completely. Since the properties for all systems start to change at a chemistry- and temperature-dependent chain length N_e or molecular weight M_e in the same way, one is led to the idea that this can only originate from properties common to all chains, namely the chain connectivity and the fact that the chains cannot pass through each other.

2 Polymer Dynamics and Network Elasticity

Polymer melts and networks display a viscoelastic behavior with a characteristic time dependent elastic modulus $G(t)$, which can be derived from the time dependent restoring force of a polymer melt or network after a step strain. Experimentally usually an oscillatory shear is applied. What one finds is illustrated in Fig. 1. After a drastic fast initial decay the modulus $G(t)$ stays almost constant at a value G_N^o for a long time, if the chains are long enough, until $G(t)$ eventually decays to zero. Many related experimental findings can be found in Ferry's classical book [15]. Figure 1 also illustrates the similarities between cross-linked melts (rubber) and non-cross-linked melts. In both cases the value of the plateau modulus eventually becomes independent of N, which is either the chain length of the melt or the average strand length between two cross-links, if only N is large enough. These similarities lead to the famous reptation or tube concept by Edwards [16] and deGennes [17].

Edwards in his work on cross-linked networks introduced the concept of obstacles created by the other chains, resulting in a "tube" in which the monomers move. Figure 2 shows a "historical" sketch of the development of this concept. First consider a network. The figure shows one strand of the network in the center marked by a thick black chain and a rather crude sketch of the surrounding. Edwards discussed how the black center chain could move

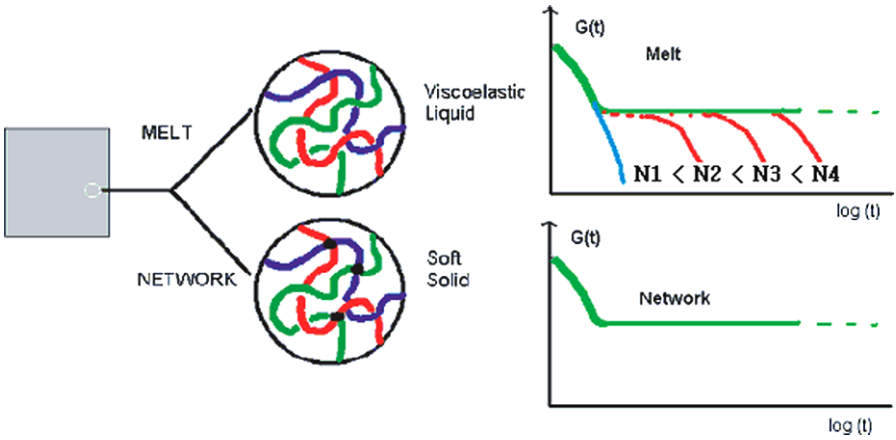


Fig. 1. Cartoon of the characteristic structure and response of a polymer melt (*top*) and a polymer network (*bottom*) after a step strain (typical log-log presentation). In melts, for short chains (length N1) the restoring force decays to zero very fast, while for the longer ones with increasing length N, as indicated, a plateau in the time dependent modulus occurs, which is independent N

around subject to obstacles created by all the other chains which in this case are part of the network. He noted that due to topological constraints the chain is much more localized than expected just by the fact that the two ends are connected to a cross-link. All loops and their links in the system are conserved; they cause the strand to be essentially confined to a tube-like region (Fig. 2, middle part). This hypothetical tube, built by all the other chains, follows the coarse-grained conformation of the chain. The length scale of this coarse graining is the tube diameter d_T. A tube segment of size d_T typically contains $d_T^2/ll_k \cong N_e$ monomers of the test chain, where N_e is called the entanglement length. Within this picture the network strand can perform a quasi one-dimensional Rouse relaxation along that tube. Later, deGennes realized that the motion and spatial fluctuations of long chains in melts should be governed by the same mechanism (Fig. 2, lower part). When the chains are very long, most of the monomers are far from the chain ends. Then, on intermediate time scales, these monomers do not realize that the ends are free. Since the density of chain ends is very small, $O(\rho/N)$, the topology of the surrounding does not change significantly on these intermediate time scales and a chain can only diffuse by reptating out of its original tube, leading to a longest relaxation time $\tau_d \sim N^3$, a diffusion constant $D \sim N^{-2}$, as well as a plateau modulus at intermediate time scales. Considering the simplicity of the concept, the model describes many experimental findings remarkably well. However, in spite of its successes, several open questions remain, including how to formulate the reptation concept on a more fundamental basis.

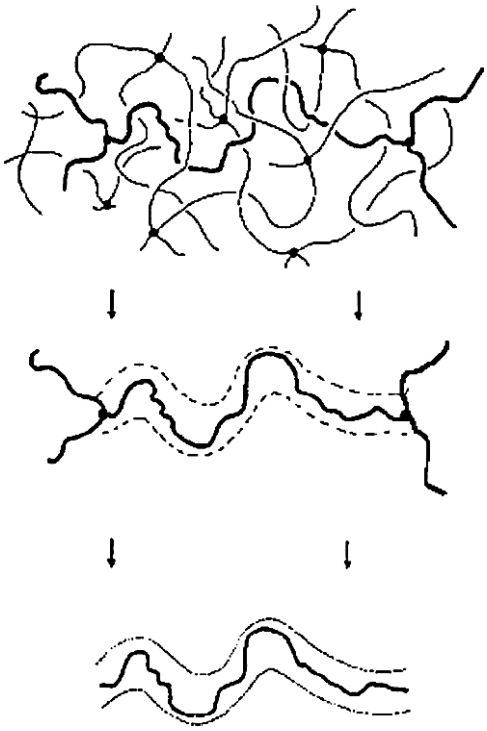


Fig. 2. Sketch of the historical development of the tube constraint and reptation concept. Starting from a network Edwards in 1967 defined the confinement to the tube, while deGennes in 1971 realized that for long chains the ends only play a small role for intermediate times

A quantitative structure based model or theory of what an entanglement really is, remained largely unsettled until some very recent progress [18–20]. Also the discrepancy between the observed viscosity of $\eta \propto N^{3.4}$ and the predicted power law $\eta \propto N^3$ by now is safely attributed to a very slow crossover towards the asymptotic regime [2]. This is a little different for the diffusion of constant D, where a collection of recent data finds [21] $D \propto N^{-2.4}$ instead of N^{-2}. It is however not yet settled whether this is the "same" crossover effect [22] or whether it can be explained by a correlation hole[1] effect. A detailed discussion however is beyond the scope of the present contribution.

[1] The correlation hole is defined by the self density of the chain in the melt. The density $\rho = \rho(r)_{other} + \rho_{self}(r)$ with $\rho_{self} \simeq N/R^3 \propto N^{-1/2}$. Viewing the chains as very soft spheres, they like in a liquid of ordinary spheres have to leave their cage when diffusing around. This leads to a back jump correlation and slows down the diffusion beyond the influence of the microscopic bead friction. Note that the center of gravity of a chain not necessarily has to move in order to relax the overall chain conformation.

It is clear that computer simulations can be a very versatile tool to investigate such problems, since they offer the unique opportunity to have full control over the chain conformations while simultaneously typical experimental observables can be "measured". Though CPU time intensive, simulations have played an important role over the years and will continue to do so.

Experimental quantities such as the viscosity, diffusion constant and modulus do not directly probe the microscopic motion of monomers on the chain. In contrast neutron spin-echo scattering covers the appropriate length scales, but the time range is rather limited. In a similar way pulsed field gradient spin-echo NMR is able to address the appropriate time and distance scales in a limited time regime as well. By performing the experiment at rather different temperatures, this can be extended, however creating new problems. Generally, an experiment typically probes one aspect only. Also samples are never really ideal and one must e.g. deal with polydispersity effects. Simulations do not suffer from such problems and can now be performed on melts of chains of 15–20 N_e, answering a number of unsolved questions.

3 Theoretical Concepts

3.1 Unentangled Chains – Rouse Regime

In the Rouse model, all the complicated interactions are absorbed into a monomeric friction and a coupling to a heat bath. It was originally proposed to model an isolated chain in solution (cf. chapter by B. Dünweg), though it actually only works well for short chains in a melt. For a detailed account I refer to the book of Doi and Edwards [1] and a recent review by McLeish [2]. The polymer is modelled as a freely jointed chain of N beads connected by $N-1$ springs, immersed in a continuum. Each bead experiences a friction, with friction coefficient ζ. The beads are connected by a Hookean spring with a force constant $k = 3\,k_BT/b^2$, where $b^2 = \ell\ell_K$. Each bead-spring unit is intended to model a short subchain, not just a monomer. The equation of motion of the beads is given by a Langevin equation. For monomer $i (i \neq 1, N)$ it reads,

$$\zeta\dot{\mathbf{r}}_i = k(2\mathbf{r}_i - \mathbf{r}_{i-1} - \mathbf{r}_{i+1}) + \mathbf{f}_i \tag{3}$$

The distribution of random forces $\mathbf{f}_i$ is Gaussian with zero mean and the second moment(k_B being Boltzmann's constant and T is the temperature):

$$\langle \mathbf{f}_i(t) \cdot \mathbf{f}_j(t') \rangle = 6\zeta k_BT\delta_{ij}\delta(t-t') \; . \tag{4}$$

Note that this model does not contain any specific interactions between monomers except those due to the chain connectivity. Since in a melt, the long-range hydrodynamic interactions are screened, it was suggested that this model could describe the motion of those chains, except that ζ arises from other chains rather than the solvent.

The Rouse model can be solved by transforming to normal coordinates $\mathbf{X}_p(t)$ of the chain, for details see [1, 2, 6, 7]. It leads to a longest relaxation time (Rouse time):

$$\tau_R = \frac{\zeta \ell \ell_K NN}{3\pi^2 k_B T} \ , \tag{5}$$

an exponentially decaying time dependent modulus

$$G(t) = \frac{\rho K_B T}{N} \sum_P \exp(-2tp^2/\tau_R) \tag{6}$$

and a melt viscosity

$$\eta = \int_0^\infty G(t)dt = \frac{\rho\zeta}{36}\langle R^2\rangle \sim N\zeta \ . \tag{7}$$

The self-diffusion constant D can be determined from the mean-square displacement of $\mathbf{r}_{cm}$, the center-of mass of the chain,

$$g_3(t) = \langle(\mathbf{r}_{cm}(t) - \mathbf{r}_{cm}(0))^2\rangle \tag{8}$$

Within the Rouse model $g_3(t) \sim t$ for all times and the diffusion constant $D(N) = lim_{t\to\infty} g_3(t)/6t$ is

$$D = \frac{k_B T}{N\zeta} \tag{9}$$

already for relatively short times compared to τ_R. In simulations often the mean-square displacements of monomers $g_1(t)$ as a function of time t

$$g_1(t) = \frac{1}{N}\sum_{i=1}^{N}\langle[\mathbf{r}_i(t) - \mathbf{r}_i(0)]^2\rangle \ . \tag{10}$$

is studied, with $r_i(t)$ being the position of bead i at time t. Using the fact that the chain structure is that of a Gaussian random walk, one gets

$$g_1(t) \sim \begin{cases} t^1, & t < \tau_0, & g_1(t) < \ell\ell_K \\ t^{1/2}, & \tau_0 < t < \tau_R, & \ell\ell_K \lesssim g_1(t) \lesssim \langle R^2\rangle \\ t^1, & t > \tau_R, & g_1(t) \gtrsim \langle R^2\rangle \end{cases} \ . \tag{11}$$

For very short times, when a monomer has moved less than its own diameter, it is affected little by its neighbors along the chain. This short time regime, $t < \tau_0$ is governed by the local chemical/model properties of the chains. For intermediate times, the motion of a monomer is slowed down because it is connected to other monomers. This can be viewed as the diffusion of a particle with increasing distance dependent mass. The actual mass at time t is just the number of monomers within a sphere of diameter $\sqrt{g_1(t)}$. This continues until the chain has moved a distance comparable to its size $\langle R^2\rangle^{1/2}$. After that one observes free diffusion with $D \sim N^{-1}$. It turns out experimentally

that this extremely simple model provides an excellent description of polymer dynamics, provided that the chains are short enough. Measurements of η [15], NMR [23,24]and neutron spin-echo scattering experiments [25] and results for molecular dynamics simulations on short chains also agree surprisingly well to that. For short chains, it turns out that the noncrossability of the chains, as well as the chains chemical structure mostly affect the prefactors in the diffusion coefficient through the monomeric friction coefficient ζ. Why these effects average to such a simple contribution still is not understood.

3.2 Entangled Chains

For chains which significantly exceed the length N_e, the motion is slowed down drastically. Clear Evidence for this slowing down comes from the diffusion constant D [14, 26, 27]. For $N \geq N_e$,

$$D \sim N^{-2} \ldots N^{-2.4} \tag{12}$$

Several forms for the prefactor of D have been discussed in the literature. Similarly the viscosity η increases from $\eta \sim N$ for short chains to

$$\eta \sim N^{3.4} \,. \tag{13}$$

In reptation theory the motion of the chains is viewed as Rouse motion of chains in a tube of diameter d_T. Since the chain is modelled as a random walk, forces at the ends have to keep a tube contour length $L_T \sim N$. In the original concept the tube is fixed and a chain had to completely move out of the tube to relax its conformation and any stress linked to the conformation. All other means of relaxation, such as constraint release due to moving chain ends of those chains, which form the tube, or fluctuation effects like contour length fluctuations of the tube modify this scheme only somewhat quantitatively, but do not alter the qualitative picture. I thus will here discuss the simplest case only [1].

For short time scales the motion of the monomers cannot be distinguished from that of the Rouse model, the motion of the monomer is isotropic and $g_1(t) \sim t^{1/2}$. Only after the motion reaches a distance of the $O(d_T^2 \cong \langle R^2(N_e)\rangle)$ the constraints from the tube are showing up. The corresponding time is the Rouse time of a subchain of N_e beads, namely $\tau_e \sim N_e^2$. After this time the monomers can diffuse along the tube only. By this forward and backward motion, the chain explores new space and very slowly destroys the original tube. The contour length of the tube L_T can be estimated to $L_T \approx d_T N/N_e \sim N/N_e^{1/2}$. For $t > \tau_e$, the chain performs essentially a one dimensional Rouse motion along the random walk like tube turning the $t^{1/2}$ power law for $g_1(t)$ into a $t^{1/4}$ power law. However, after the Rouse relaxation time τ_R, the monomers have only moved a distance of order $L_T^{1/2} \ll L_T$ for $N \gg N_e$ along the tube. Following this regime the overall diffusion along the

tube dominates and results in a second $t^{1/2}$ regime for the motion in space. The initial tube will be destroyed when one of the segments has visited $O(N)$ different contiguous sites. This requires a time $\tau_d \sim N^3/N_e$. In this time the chain has moved a distance comparable to its own size, therefore the diffusion constant D is expected to scale as N^{-2}. In summary the theory predicts the following general power-law sequence for the mean-square displacement in space, $g_1(t)$:

$$g_1(t) \sim \begin{cases} t^1, & t < \tau_0 \ ; \\ t^{1/2}, \tau_0 < t < \tau_e \sim N_e^2 \ ; \\ t^{1/4}, \tau_e < t < \tau_R \sim N^2 \ ; \\ t^{1/2}, \tau_R < t < \tau_d \sim N^3/N_e \ ; \\ t^1, & t > \tau_d \end{cases} \tag{14}$$

which is shown schematically in Fig. 3.

For the motion of the center-of-mass $g_3(t)$ one expects

$$g_3(t) \sim \begin{cases} t^1, & t < \tau_e \sim N_e^2; \\ t^{1/2}, \tau_e < t < \tau_R \sim N^2; \\ t^1, & \tau_R < t \end{cases} \tag{15}$$

Direct experimental evidence for these intermediate time regimes has been found by NMR [23,24,28], from diffusion experiments of polymers at an interface [29], and very recently from n-scattering [30], however simulations were the first to observe the crossover into the $t^{1/4}$ regime [3].

To understand the plateau modulus [1,15,31] G_N^0 we have to resort to the possible conformations of the chains. Assuming that in a melt as well as in

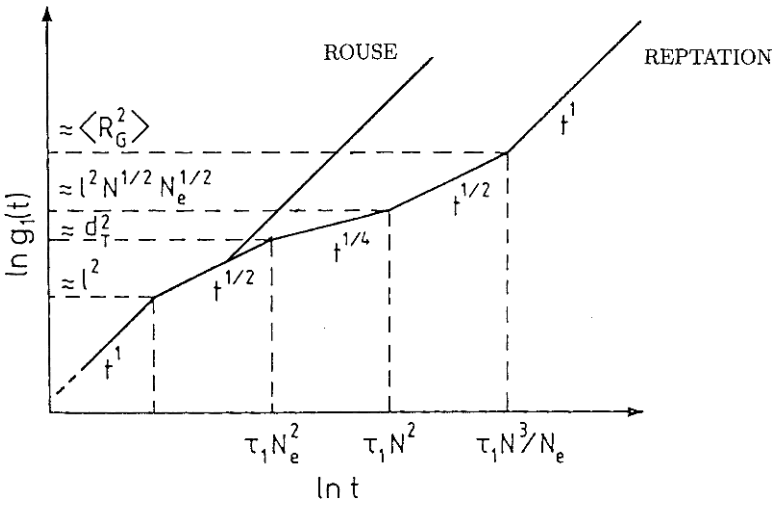


Fig. 3. Schematic plot of the mean-square displacement for a monomer in the Rouse and the reptation model

a network chains are unperturbed random walks, the distribution function of the end-to-end vector $\mathbf{R}$ is given by:

$$P(\mathbf{R}) = \left(\frac{3}{2\pi L\ell_K}\right)^{3/2} \exp\left(-3R^2/2L\ell_K\right) \tag{16}$$

Since no specific inter-chain interactions are considered, $P(\mathbf{R})$ is the probability of a chain in a melt or a strand between two crosslinks in a network to be separated by a vector $\mathbf{R}$. This defines the entropy for a given $\mathbf{R}$ as

$$S = k_B \ell n P(\mathbf{R}) = \text{const} - 3k_B R^2/2L\ell_K \tag{17}$$

with the energy $U = 0$ (chain is viewed as a free random walk)

$$F = U - TS = \frac{3k_B T R^2}{2L\ell_K} \tag{18}$$

for the conformation dependent contribution to the free energy of a chain with fixed end positions. This leads to an entropic force

$$\mathbf{f} = -\nabla F(\mathbf{R}) = -\frac{3k_B T}{L\ell_K}\mathbf{R}\,, \tag{19}$$

the force of a linear elastic spring. The whole chain acts like a bond in the Rouse model. A random walk at a temperature T is a Hookean spring with spring constant $\sim T/N$. This almost trivial fact is the basis of all theories of network elasticity as well as a basis for the view of the Rouse chain to be a string of beads with linear elastic springs in between. Now let us assume a perfect network with the crosslinks fixed in space. Taking a polymer network as a collection of hookean springs, the elastic modulus just is the density of elastically active strands.

$$G = k_B T \rho / N \tag{20}$$

For networks a variety of modifications are introduced, in order to properly count the elastically active strands. For short chains they all however modify the effective density of strands, but not the N dependency and the entropic origin of the modulus. Both aspects are well established from simulations and many experiments, where it is shown that for most polymers enthalpic contributions to the elasticity are negligible.

In the case of very long chains, $N \gg N_e$ the simple Rouse model assumption that the chains or strands act as non-interacting random walks is not applicable anymore. Following the scheme discussed before one can view the strands in the Edwards tube as a continuation of entropic springs of length N_e, which leads to [1, 15]

$$G_N^o = \frac{\rho k_B T}{N_e} \tag{21}$$

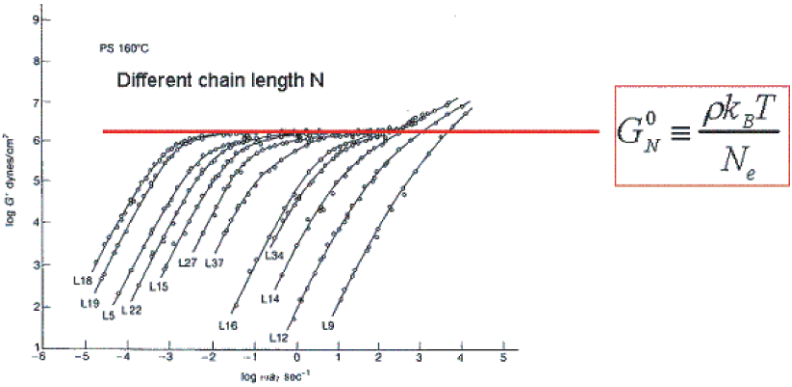


Fig. 4. Experimental determination of the frequency dependent modulus $G'(\omega)$ of polystyrene melts for many different chain lengths, from [1]. Note that the prefactor of 0.8, cf. 11, is omitted in the figure

for networks and taking additional relaxation mechanisms, such as tube length fluctuations, into account to

$$G_N^o = \frac{4}{5}\frac{\rho k_B T}{N_e} \tag{22}$$

for melts. This is the formula commonly used to determine N_e from measurements of the plateau modulus. A typical example in given in Fig. 4.

The following table summarizes our findings so far.

Table 1. Summary and comparison of the discussed quantities for the Rouse and the tube model

	Rouse Regime	Reptation Regime
	$N < N_e$, $M < M_e$	$N \gg N_e$, $M \gg M_e$
Diffusion	$D \propto N^{-1}$	$D \propto N^{-2}$
Relaxation Time	$\tau_R \propto N^2$	$\tau_d \propto N^3$
Viscosity	$\eta \propto N$	$\eta \propto N^3$
Modulus G_N^o	$G_N^o = 0$	$G_N^o \propto \frac{kT}{N_e}$
$(t \ll \tau_d)$		

Considering the previous discussion theory, simulations and experiment nowadays do not focus that much any more on the question whether entanglements exist and whether they are relevant but more on their consequences and on the problem of how to quantify them properly. Typical topics are:

- Crossover to asymptotic power laws for η and D, what additional relaxation mechanisms exist and what kind of additional constraints, mechanisms might slow down the diffusion?

- How and why do different measurements of N_e give rather different results?
- What are the effects of dilution and/or solvent quality in semidilute systems?
- What does the tube "look like" (under deformation)?
- Tube deformation and tube relaxation, quantitative and qualitative differences between networks and melts
- Swelling behavior with and without the influence of charges (polyelectrolyte gels).
- Effects of polydispersity, mixtures as well as the influence of branched additives
- Structure property relations, can we predict N_e from the chemical structure of the polymer or any static measurement?
- Nonequilibrium properties, like shear thinning etc.

This list is not comprehensive, but it shows that the consensus that entanglements dominate many crucial properties is not really the end of a long development but rather defines many new research opportunities. To follow this by simulation we first have to describe how to prepare "good" initial states.

4 Simulation Models, Equilibrated Melts

Most of the following will be discussed for a freely jointed bead spring chain. The extension to other models is straight forward [6,10,32]. At the beginning we have to have "well equilibrated samples" [33]. The most direct way would be to set up a system in an arbitrary state and equilibrate it by running for a few relaxation times, however requiring prohibitive amounts of CPU time. Since the required total CPU time for well optimized codes would scale with $(MN)N^{3.4}$, M being the number of chains, which should at least scale like $N^{1/2}$, also for modern computers it is important to start out with already well equilibrated systems. Even taking this into account, the simulation of an all atom system of entangled polymers for a long enough time is still out of question and also computers of the next few generations will not solve that problem.

Fortunately for equilibration there are ways out. This, however requires a precise knowledge of the average equilibrium conformation [34]. In a melt or dense solutions the chains assume random walk statistics for all contour lengths $L \gg l_k$. Once a "sample" which reaches that asymptotic regime is available, one can use this as a reference state even for much longer chains. Values for internal distances, $\langle \Delta r_{ij}^2 \rangle$, have to fall on a general master curve. For the internal distances $\langle \Delta r_{ij}^2 \rangle$ one finds

$$\langle \Delta r_{ij}^2 \rangle \propto \begin{cases} \mid i-j \mid^2, & \mid i-j \mid \ell < \ell_k \\ \mid i-j \mid, & \mid i-j \mid \ell \gg \ell_k \end{cases} \tag{23}$$

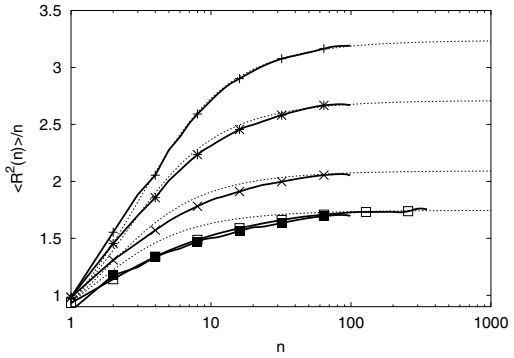


Fig. 5. Normalized internal distances $n = |i - j|$ for different chain lengths for a standard bead spring LJ polymer, for different ℓ_K ranging from about $\ell_K = 1.7\sigma \ldots 3.3\tau$. [34] The *dotted* line is a comparison to the so called worm like chain model, which approximately describes the crossover [34]. In all cases the particle density was fixed to $\rho = 0.85\sigma^{-3}$

with additional possible deviations for very small $\mid i - j \mid$ due to the chemical structure, if a more detailed model is treated. This function has to be determined in a complete simulation of shorter chains. Thus in a melt a system specific master plot gives $\langle \Delta r_{ij}^2 \rangle / (l \mid i - j \mid)$ vs $l \mid i - j \mid$, the corresponding contour length. An example is given in Fig (5),which also shows characteristic deviations from the theoretically expected curve for small distances, where the bead packing is dominant. These deviations for fully flexible chains extend out to contour distances of about 50 beads!

In a similar way one can apply the analysis of the chain form factor $S(q)$

$$S(q) = \frac{1}{N} \left\langle \left| \sum_{j=1}^{N} e^{i\boldsymbol{q}\boldsymbol{r}_j} \right|^2 \right\rangle_{|q|} \tag{24}$$

where the index $\mid q \mid$ denotes a spherical average. In a melt one finds

$$S(q) \propto \begin{cases} N(1 - \frac{1}{3} q^2 \langle R_G^2 \rangle), & \frac{2\pi}{q} \gg \langle R_G^2 \rangle \\ q,^{-2} & \frac{2\pi}{\ell_q} \ll \frac{2\pi}{q} \ll \langle R_G^2 \rangle \\ q,^{-1} & \frac{2\pi}{\ell_q} < \frac{2\pi}{q} < \frac{2\pi}{\ell} \\ O(1), & \frac{2\pi}{q} > \frac{2\pi}{\ell} \end{cases} \tag{25}$$

Besides the initial decay the form factor describes one characteristic, chain length independent curve. Figure 6 shows a typical example.

Often the so-called Kratky plot, $q^2 S(q)$ vs q, is shown. Deviations from slope zero in the q^{-2} regime are easily to detect, even by eye. This helps for a quick check. Master plots of $S(q)$ or $\langle \Delta r_{ij}^2 \rangle$ can be used to control the equilibration of melts.

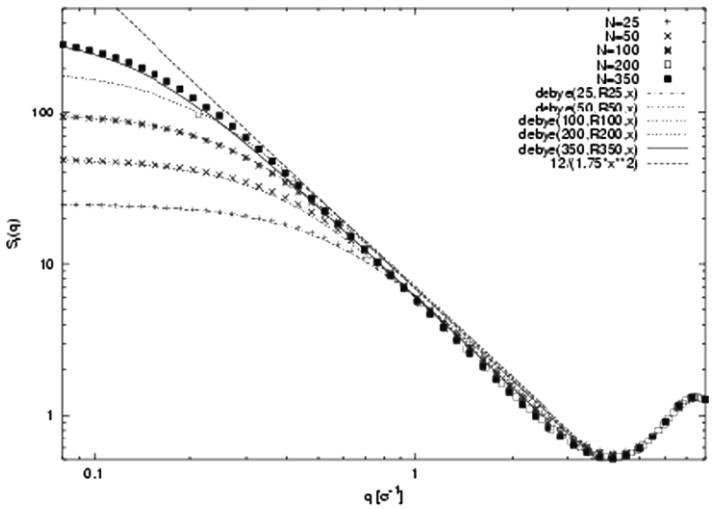


Fig. 6. Chain form factor $S(q)$ vs q for fully flexible LJ chains in a melt. Chain lengths range from $N = 25$ to $N = 350$. For fits the Debye function was used. The *straight line* indicates a slope of q^{-2}, from [35]

4.1 Preparing an Equilibrated Melt, Specific Systems

A rather straight forward approach is to run a system by a reptation or slithering snake algorithm. This beats the slow realistic dynamics by a bit more than $O(N)$, however, is not really applicable for dense continuum systems. For semidilute solutions or lattice polymers at moderate densities and for moderate chain lengths this however is appropriate. For dense continuum systems as well as systems with realistic chemical details such an approach fails. A well working strategy for many model systems is as follows [34]:

- Simulate a melt of many short, but long enough chains ($N\ell \gg \ell_K$) into equilibrium by a conventional method
- Use this melt to construct the master curve or target function for the melts of longer chains (cf. Figs. 5, 6)
- Create non reversal random walks of the correct bond length, which match the target function closely. They have to have the anticipated large distance conformations. Introduce, if needed beyond the intrinsic stiffness of the bonds, stiffness via a suitable second neighbor excluded volume potential along the chain. (This might be a bit larger than the one of the full melt!)
- Place these chains as rigid objects in the system and move them around by a Monte Carlo procedure (shifting, rotating, inverting..., but **not** manipulating the conformation itself) to minimize density fluctuations
- Use this state as starting state for melt simulations

- Introduce slowly but not too slowly the excluded volume potential by keeping the short range intra chain interactions, taking care that in the beginning the chains can easily cross through each other (see below)
- Run until the excluded volume potential is completely introduced. Control internal distances permanently to check for possible overshoots, deviations from the master curve.
- Eventually support long range relaxation by so called end bridging [36,37] or double pivot moves [34]

Independent on the details of the procedures, it is important to continuously compare the actual structure to the master or target curves. If one is stuck with this approach and deviations from the master curve occur and stay, one has to start all over again. Reference [34] also demonstrates some typical deviations from the master curve as they can occur during the set up.

As an example let us discuss a melt of simple bead spring chains [34,38,39]. All beads interact via a purely repulsive LJ potential, to model the excluded volume interaction.

$$U_{LJ}(r) = \begin{cases} 4\epsilon \left\{ (\sigma/r)^{12} - (\sigma/r)^6 + \frac{1}{4} \right\} & r \le r_c \\ 0 & r \ge r_c \end{cases} \tag{26}$$

with a cutoff $r_c = 2^{1/6}\sigma$. The beads are connected by a finite extensible nonlinear elastic potential (FENE)

$$U_{FENE}(r) = \begin{cases} -0.5 R_o^2 k \ell_n (1 - (r/R_o)^2) & r \le R_o \\ \infty & r > R_o \end{cases} \tag{27}$$

in addition to the Lennard Jones Potential. The parameters are usually taken as $k = 30\epsilon/\sigma^2$, $R_o = 1.5\sigma$ in melt simulations. The temperature $T = \epsilon/k_B$ and the basic unit of time is $\tau = \sigma(m/\epsilon)^{1/2}$, where all masses are set to one. The temperature was kept constant by coupling the motion to a Langevin thermostat. An alternative, which does not screen hydrodynamics is the DPD thermostat [40]. The bead density is set to $\rho = 0.85\sigma^{-3}$. The average bond length with these parameters is $\langle \ell^2 \rangle^{1/2} = 0.97\sigma$. The "samples" consist of M chains of N beads.[2]

For this model in a melt of density $\rho = 0.85\sigma^{-3}$ the bond length is $\langle l^2 \rangle^{1/2} = 0.97\sigma$ with $c_\infty l^2 = l l_k \simeq 1.7\sigma^2$. Since, in our case, there are no torsional barriers, the monomer packing, which locally depends on the ratio of bond length to effective excluded volume of the beads, relaxes quickly. This however does not tell anything about the global structure. The local packing only characterizes an equilibration on the smallest length scale considered and very small systematic deviations on that scale can lead to significant deviations

[2] In addition a bending stiffness can be introduced via an effective three body interaction. With $\cos\Theta_i = (\hat{\mathbf{r}}_{i,i-1} \cdot \hat{\mathbf{r}}_{i,i+1})$, with $\hat{\mathbf{r}}_{i,i\pm1} = (\mathbf{r}_i - \mathbf{r}_{i\pm1})/\mid \mathbf{r}_i - \mathbf{r}_{i\pm1} \mid$ the bending potential reads $U_{bend}(\Theta) = k_\Theta(1 - \cos\Theta)$. By variation of k_Θ the model can be tuned from very flexible to very stiff.

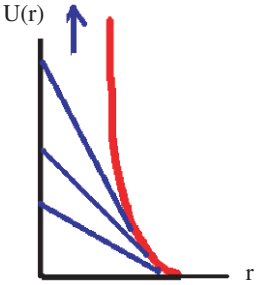


Fig. 7. Illustration of the "force capped" Lennard Jones interaction used to introduce step by step the full excluded volume of the chains

from equilibrium at larger length scales, which can change results significantly [33]. To test our approach, an extraordinary effort was taken by G. S. Grest, who equilibrated a melt of chains of $N = 350$ beads by MD simulations running in the background. Figure 5 also includes some data for stiffer chains. From these data we get the target function with:

$$\langle R^2(\mid i-j \mid, N)\rangle / \mid i-j \mid = \ell\ell_k, \mid i-j \mid \ell \gg \ell_k \tag{28}$$

Having this reference state, we can introduce the excluded volume and proceed as follows:

1. Use a force-capped-Lennard-Jones-potential
2. increase the push-off time to make the procedure "quasi static".

The force capped potential we use reads :

$$\mathrm{U_{FCLJ}}(r) = \begin{cases} (r - r_{fc}) * \mathrm{U_{LJ}}'(r_{fc}) + \mathrm{U_{LJ}}(r_{fc}) & r < r_{fc} \\ \mathrm{U_{LJ}}(r) & r \geq r_{fc} \end{cases} \tag{29}$$

In the present case, r_{fc} is gradually reduced from the Lennard-Jones cut-off radius $r_c = 2^{1/6}\sigma$ to 0.8σ which is significantly smaller than the relevant interparticle distances, as illustrated in Fig. 7.

Force-capping has the advantage that this form of the soft potential systematically approaches the true potential. The typical time we used to introduce the full potential was about 50000τ, which is of the order of twice the Rouse time of a chain of $N = 100$ beads. Such a procedure relatively safely equilibrates the internal conformations of the chains as Fig. 8 shows. There the results of a too fast and a proper push off procedure are shown for illustration. The too fast introduction shows the characteristic overshoot of the internal distances for small n. It is however also often rather important that the average end to end distance is close to the desired value, even though the overall number of chains might be relatively small. Then methods, which break and recombine chains, i.e. the double-pivot or bridging algorithms can be implemented in addition. Note that they very effectively relax and manipulate large distances, however affect small scales only extremely weakly!

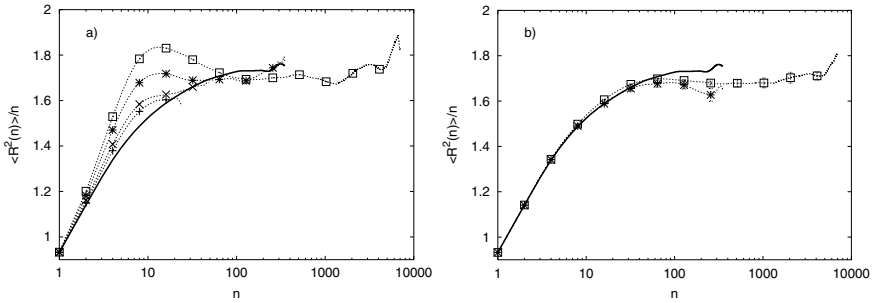


Fig. 8. Mean square internal distances for chains of length $N = 25$ (+), 50 (x), 350 (*) and 7000 (□) for a too fast pushoff procedure (**a**) and the slower version as described in the text (**b**), from [34]

Having prepared melts by the above procedure we now can either perform standard simulations to study the melt or crosslink them in order to create networks by a variety of methods.

The preparation of networks deserves another short comment. In the literature a number of different crosslinked systems has been studied extensively [35, 41–45]. A typical range of systems in illustrated in Fig. 9.

The randomly crosslinked system certainly is closer to standard experiments however encounters many problems [41]. Especially the dangling chain ends cause the relaxation to slow down dramatically (the time grows exponentially!), because it is linked to a retraction of the arms. Such effects are also known from other situations where branched polymers play an important role [2]. In addition the disorder is quenched. Thus simply running a system longer does not improve the data in comparison to experiment. For this one would need "many" medium sized systems or single extremely large ones. Thus we mostly stick to melts in the following.

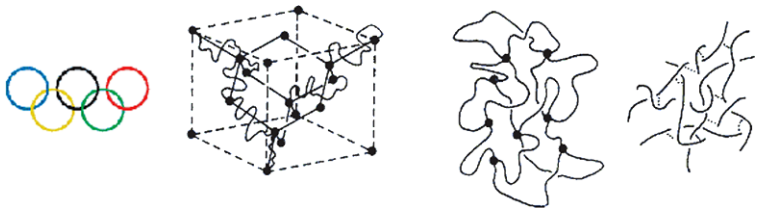


Fig. 9. Typical crosslinked systems ranging from randomly crosslinked melts via various versions of endlinked polymer networks to the idealized lattice structure systems and the theorist's dream of an "olympic gel"

5 Entanglement Analysis

5.1 "Visual" Inspection of Melts and Networks

Experimentally the determination of confinement is somewhat indirect and does not easily reveal the shape of the confining volume. Here simulations offer a unique way of visual inspection of the chain displacements and thus the tube. To do so one simply can plot the initial conformation of the chain and on top of that subsequent conformations, as is shown in Fig. 10 for a network.

Depending on the view onto the initial conformation (the typical conformation of a random walk is a flattened thick cigar with a ratio of principal moments of inertia of $R_{11}^2 : R_{22}^2 : R_{33}^2 :\cong 11.8 : 2.5 : 1$) one clearly can identify the confinement. The diameter of the confining tube actually nicely compares to results from earlier simulations on the mean square displacements. In a similar way one can also visualize the motion in polymer melts. There, however, due to a continuous release and creation of constraints, this is more difficult to visualize. Thus one can use a small trick. What we now plot is not the bare conformation, but the backbone of the statically averaged chain contour, namely

$$\mathbf{R}_i = \frac{1}{n+1} \sum_{j=1-n/2}^{i+n/2} \mathbf{r}_i \tag{30}$$

This smoothens the contour and by a proper choice of n should be close to the backbone of the tube. Here $n = 35$ was used. These contours, plotted vs time,

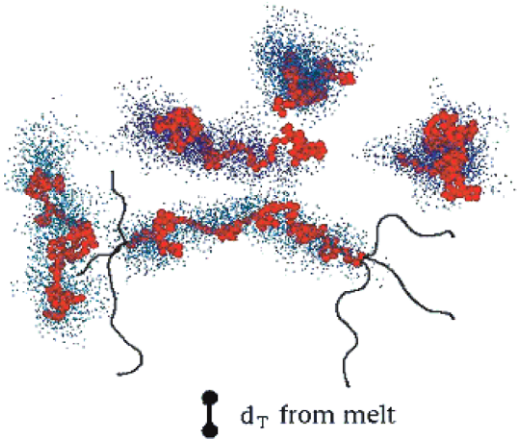


Fig. 10. Time evolution of the conformation of chains in an endlinked network of functionality $f = 4$, $N = 100$, where all network strands have the same length. The chains are, for clarity shown without the network. The indicated tube diameter was determined in an independent melt simulation of the very same polymer model, from [42]

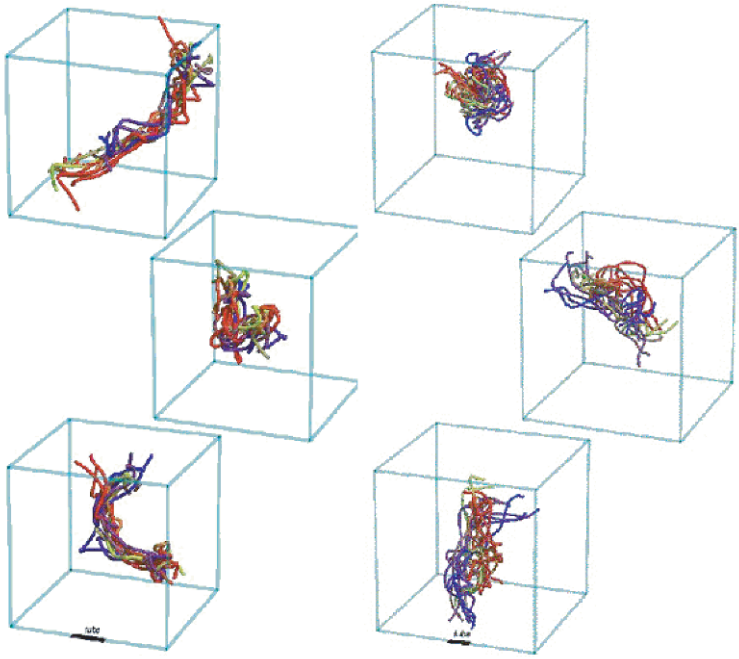


Fig. 11. Time evolution of the coarse grained backbone of the chains for a melt of $N = 350$ bead chains for times up to the Rouse time τ_R. The *left* panel shows chains with the full excluded volume, while for the *right* panel the crossing barrier is only a few kT, from [35]

are compared to a different system, where the chains can easily cross each other. For the full LJ interaction the barrier height is around $70\,kT$. One can however introduce a potential, where the chains can cut through each other, but the pressure, $\langle R^2 \rangle$ and the bead friction ζ remain unchanged. These two systems are compared in Fig. 11 for chains of $N = 350$. While the confinement is not as easily visible as for the networks, the difference between the two cases is striking, and nicely supports the tube idea.

Another direct test which shows the relevance of the noncrossability of the chains/network strands is shown in Fig. 12 for networks with diamond lattice connectivity. In order to fill space with the condition that the strands obey random walk statistics several interpenetrating networks are needed. This is the source of random linking of the different subnets. By this, "short topological paths" through the net are created, which under strong elongation carry most of the stress. Note that in this special example all "short chemical paths" have the same length, which allows us to reduce the problem to that of conserved loop-loop topology.

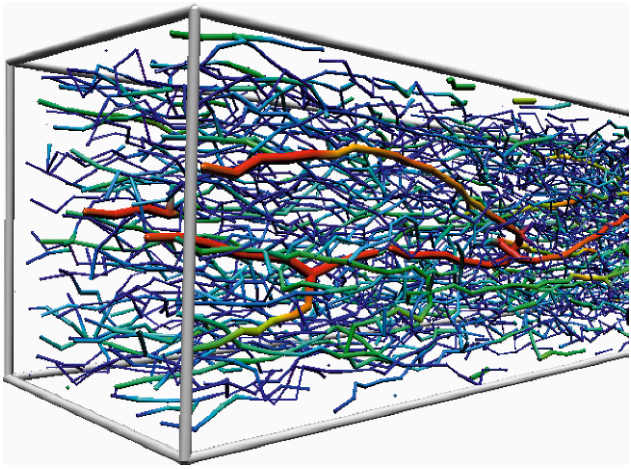


Fig. 12. Elongated diamond lattice networks, where the only source of disorder are random links between network loops. The strands are shown due to their stretching (similar to the stress they carry) from small stretching (*thin*) to strong stretching (*thick*). From [3]

5.2 Analysis of Chain Motion

After these pictures we move to a more quantitative data analysis, which demonstrates the confinement not that directly, but is rather closely related to possible experiments. The first look is at the mean square displacement of the chains and their middle beads. Many studies on this have been performed [39,46–49], for an overview see [3] and computer simulations actually were the first to observe the slowing down of the motion clearly [39,50]. Figure 13 shows results from a system of fully flexible polymers. More complicated models have been studied in [52, 53] and [47] for instance. Recently also a coarse grained model of polycarbonate [54] has been analyzed along the lines discussed here (see below). Figure 13 shows that the expected power laws are reproduced with a remarkable accuracy. From this a tube diameter of $7\sigma \leq d_T \leq 8\tau$ can be deduced, corresponding to an N_e of about $N_e \approx 32$, which is very small. To check the overall consistency the onset of the $t^{1/4}$ regime, the g_1, g_3 values at that time etc. are analyzed. In all cases the results roughly agree! Experiments also by now are able to analyze the chain motion directly and display the expected power laws [23,24,28,30]. They typically only cover a too small time window. To overcome this problem rather different temperatures are employed. It should however be noted that the shift in temperature not only affects the ratio k_BT/ζ but via conformational changes also N_e! Here simulations, despite of their many shortcomings are superior.

Besides the individual beads also the chain diffusion constant $D(N)$ can be used to estimate N_e. As mentioned earlier, a scaling plot of the ratio of the actual chain diffusion constant $D(N)$ and the hypothetical Rouse like diffusion

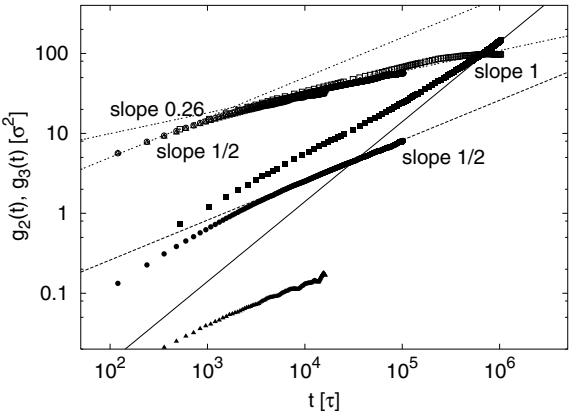


Fig. 13. Mean square displacements $g_2(t)$ (*open symbols*) and $g_3(t)$ (*closed symbols*) for chain length $N = 350(\square)$, 700 ($\circ$) and 10000 ($\triangle$). The *straight lines* show power laws as guide to the eye. The local reptation power laws $g_2(t) \propto t^{1/4}$ and $g_3(t) \propto t^{1/2}$ are verified with remarkable clarity [51]. g_2 actually is the same as g_1, but in the center of mass coordinate system of the respective chain

constant $D_R(N)$ vs the adjustable parameter N/N_e allows to fit N_e as well, once one has agreed on a master plot.

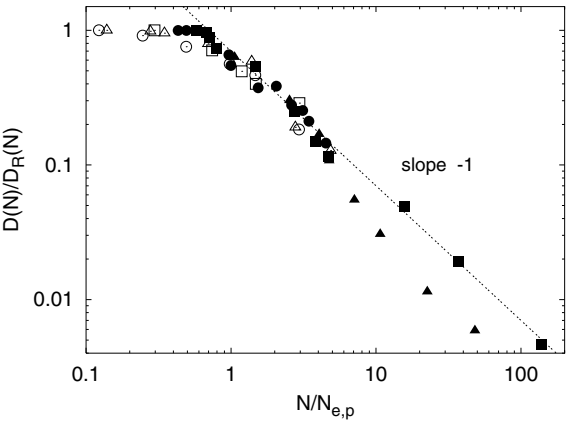


Fig. 14. Scaled diffusion constant $D(N)/D_R(N)$ vs. scaled chain length $N/N_{e,p}$ for polystyrene ($\bullet$) ($M_{e,p} = 14600$, $T = 485\,\mathrm{K}$), polyethylene ($\blacksquare$) ($M_{e,p} = 870$, $T = 448\,\mathrm{K}$), PEB2 ($\blacktriangle$) ($M_{e,p} = 990$, $T = 448K$), the present bead spring model ($\triangle$) ($N_{e,p} = 72$), the bond-fluctuation model for $\Phi = 0.5$ ($\square$) ($N_e = 30$) and tangent hard spheres at $\Phi = 0.45$ ($\circ$) ($N_e = 29$). All data are scaled with $N_{e,p}$ from the plateau modulus or with $2.25\,N_e$ from $g_1(t)$. From [51]

Figure 14 shows an example, where several very different systems are compared. This leads to a second estimate of $N_e \simeq 70$, significantly larger than the value from the bead motion.

Another important experimental technique is the neutron spin echo (NSE) method [25]. In NSE experiments the motion of the segments can be obtained by measuring the expectation value of the time-dependent single-chain structure function

$$S(\mathbf{k},t) = \frac{1}{N}\sum_{i,j}\langle \exp(i\mathbf{k}\cdot(\mathbf{r}_i(t)-\mathbf{r}_j(0)))\rangle \ . \tag{31}$$

For reptating chains deGennes calculated an expression for $S(k,t)$ [55]. In recent simulations a slightly modified version is applied [56, 57] which is identical to de Gennes result when a Gaussian model of the tube is explicitly assumed. We used this formula together with a correction for very long times (however, for the experimental NSE results or our present set of data this modification may be neglected):

$$\begin{aligned}\frac{S(k,t)}{S(k,0)} = &\left\{\left[1-\exp\left(-(kd/6)^2\right)\right]\cdot f\left(k^2b^2\sqrt{12Wt/\pi}\right)\right. \\ &\left.+\ \exp\left(-(kd/6)^2\right)\right\}\times\frac{8}{\pi^2}\sum_{p=1,odd}^{\infty}\frac{\exp\left(-tp^2/\tau_d\right)}{p^2},\end{aligned} \tag{32}$$

where $f(u) = \exp(u^2/36)\mathrm{erfc}(u/6)$. Note, that the derivation of deGennes does not take into account chain end effects. The above formula may only be applied in a narrow range of k-space $\frac{2\pi}{R_G} \lesssim k \lesssim \frac{2\pi}{d_T}$.

One finds a rather strong dependency of chain length N or the value of d_T deduced from the plot. Even for $N = 10000$ one finds N_e significantly larger than the result from the mean square displacements. While for the longer chains, there might be some problems with the statistics, the shorter chains and especially the studied fragments suffer from the fact, that they hardly fulfill the requirement of $\frac{2\pi}{R_G} \lesssim k \lesssim \frac{2\pi}{d_T}$; for the fully flexible systems shown here $6\langle R_G^2(N)\rangle \simeq \langle R^2(N)\rangle \simeq 1.7\,N\sigma^2$ giving $R_G \simeq 14.1\sigma$. Note that for these data the conformations of the longest chains were not prepared by the procedure discussed before, while all the shorter chains simply ran long enough!

The standard experimental method to determine $N_{e,p}$ (for clarity N_e determined from the plateau-modules) is by measuring the plateau modulus under oscillatory shear [1, 15]. Alternatively, it is also possible to measure the normal stress decay $\sigma_N(t)$ in a step strain elongation. Since the latter is much simpler to perform in a simulation, volume conserving step strain runs were performed for four different amplitudes $\lambda = 1.25$, 1.5, 1.75 and 2.0. The normal stress σ_N was determined by the microscopic virial-tensor. The

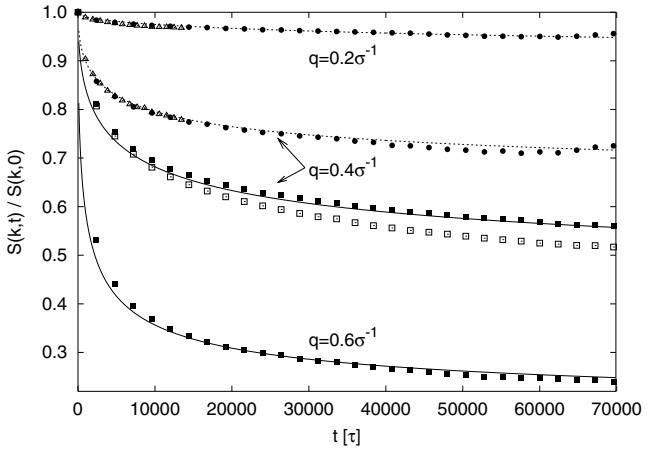


Fig. 15. $S(k,t)$ for different chain lengths: $N = 10000(\triangle)$, $N = 2000(\bullet)$, and $N = 700(\square)$ and the centered subchain of length 550 of the same $N = 700$ chains ($\blacksquare$). Continuous curves are simultaneous fits to the $N = 2000$ data corresponding to $d_T = 9.6\sigma$. The dotted curve is a simultaneous fit to the $N = 700$, 550 monomer subchain with a tube diameter of $d_t = 12.9\sigma$. [51, 58]

plateau-values of the stress were fitted to the stress-strain formulas for classical rubber elasticity (CRE) [59] $\sigma_N = G\left(\lambda^2 - \frac{1}{\lambda}\right)$ and to the Mooney-Rivlin (MR) formula [60] $\sigma_N = 2G_1\left(\lambda^2 - \frac{1}{\lambda}\right) + 2G_2\left(\lambda - \frac{1}{\lambda^2}\right)$ to determine G_N^0. MR gives $G_N^0 = 0.0105 k_B T \sigma^{-3}$ while CRE gives $G_N^0 = 0.008 k_B T \sigma^{-3}$. It is known experimentally that MR slightly overestimates the modulus while CRE always underestimates it. The standard formula [1] to calculate $N_{e,p}$,

$$G_N^0 = \frac{4}{5}\frac{\rho k_B T}{N_{e,p}} \,, \tag{33}$$

gives $N_{e,p} \approx 65$ for the MR and $N_{e,p} \approx 80$ for the CRE fit. Both values are much higher than our previous estimate, $N_e = 32$ from the mean square displacement, while the other results from the diffusion constant and the scattering functions give $N_e \simeq 70$ and $N_e \simeq 55$ respectively. Some first attempts to study the viscosity[3] by non equilibrium molecular dynamics (NEMD) simulations [61] of an almost identical model give $2N_e \simeq M_e = 110 \pm 10$. From a theoretical point of view the discrepancy might not be relevant, since the prefactors in the reptation model are not rigorously determined. For practical issues, however, like comparing results from various experiments and simulations the deviations are relevant. It also however turns out, that the ratios of the different values for N_e are strongly dependent on the simulation model employed.

[3] The crossover chain length N_c for the viscosity is experimentally typically about a factor of 2 to 2.2 larger than N_e as determined from the modulus, which here seems not really to be the case.

5.3 Structure and Property Relations: Specific Polymers

With increasing stiffness, the value of N_e decreases, however, the ratio of the different values obtained from the mean square displacements, scattering and modulus does not remain constant. This is one difficulty, when trying to apply this to specific systems. One measurement obviously is not sufficient to characterize the dynamic properties sufficiently. In addition N_e or M_e respectively varies significantly for different chemistries. The table shows the results for a variety of commodity polymers, as determined from rheology:

	PE (Polyethylene)	PS (Polystyrene)	PDMS (Polydimethyl-Siloxane)	BPA PC (Polycarbonate)
N_e	100	170	135	5 – 7
M_e	1400	18000	10000	≈1500

Even within one class of materials, e.g. different modifications of polycarbonates, significant differences can be observed, even though they have almost identical Kuhn lengths [62, 63].

These results demonstrate a characteristic challenge for a successful link between the more basic, generic phenomena oriented research and modern materials science. Our understanding eventually has to become more quantitative. Because of the huge relaxation times a factor of 2 in the entanglements change the viscosity also by a factor of $2^{2.4} \approx 5.2$ and similar reduces the diffusion constant. Uncertainties of this order are relatively small, however, can be of significant experimental and technological relevance.

As mentioned before that the simulation of a melt of entangled polymers on an all atom or even united atom level all the way into the diffusion limit is not feasible. The required CPU time simply is prohibitively large. Recently a different ansatz was used, by employing a systematic coarse graining procedure. There one maps the polymer structure onto a more coarse grained model, which still retains essential aspects of the chemical structure, but is much simpler and computationally much more efficient. Figure 16 illustrates this method as it was used for polycarbonate (BPA-PC). The chemical repeat unit is mapped onto four beads, which are connected by springs. The intramolecular potentials, such as bond length, bond angle and torsion potentials are derived, without any adjustable parameter, from a Monte Carlo sampling of isolated all atom chains. There, based on the atomistic structure, distribution functions for the coarse grained model are sampled and the potential ist determined by a Boltzmann inversion of these distributions. The nonbonded interactions then are just modelled by a repulsive Lennard Jones potential, which is adjusted to the molecular volume of the original chains. The density of these simulations is then set to the experimentally available melt density. This method has been proven to very well represent the static

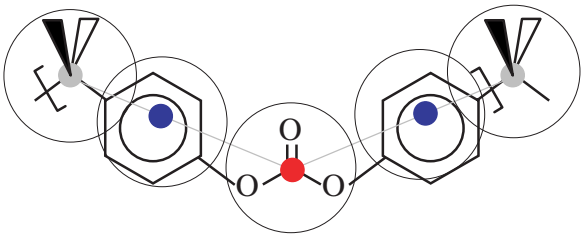


Fig. 16. Schematic representation of the coarse-grained 4:1 scheme for BPA-PC. Each chemical repeat unit is represented by four beads, corresponding to the isopropylidene, carbonate and the two phenylene groups. Because of the chain ends, a coarse-grained chain of a polymer of N repeat units contains $4N + 3$ spheres [54, 64]

structure of the polymer melts, e.g. by comparing to the experimentally determined R_G. In addition, since we are very close to the chemical structure, one can reintroduce the chemical details and compare the results to scattering experiments. Also this shows excellent agreement [64]. Currently this inverse mapping is used to generate long time atomistically resolved trajectories, to allow for a very detailed comparison with experiment. For details of the mapping I refer to [54, 64].

While the very nature of the mapping fixes the length scaling to 4.41 Å = 1σ in our Lennard Jones units, the mapping of the simulations time to a real physical time is not obvious at all. One can try to compare the viscosity obtained via the Rouse model from the chain diffusion constant to experimental viscosities in order to derive a rough estimate. This approach, however, contains a number of problems both on the experimental and simulation side. A better and direct way, is to take a short chain system and run it in both, the all atom and the coarse grained representation. This was done for the 5-mer melt of BPA-PC at the typical process temperature of T = $570K$ and is shown in Fig. 17. This reveals two important aspects. First the displacements in time have the same "shape" down to rather small distances, so that the curves overlap, after an appropriate fit of τ, already at motion distances of the size of the repeat unit (and most probably somewhat below as well). This is a clear indication, that the coarse grained model closely resembles the dynamics of the real systems. Second, by overlapping the curves we can fit the value of τ to real times (note that the lengths are fixed due to the mapping procedure!). This leads to $1\tau \approx 3.0 \cdot 10^{-11}$sec, allowing for overall simulations times of up to 10^{-4}sec of entangled polymers ($N \leq 120$, note that $N_e \simeq 6$ from rheology), several orders of magnitude more than in any other atom based simulation [64, 65]!

Based on this the same set of data analysis has been performed as for simple bead spring chains. An example is given in Fig. 18, where the dynamic structure function for the longest chains is shown, which yields values for N_e between 9.5 and 10.5 repeat units.

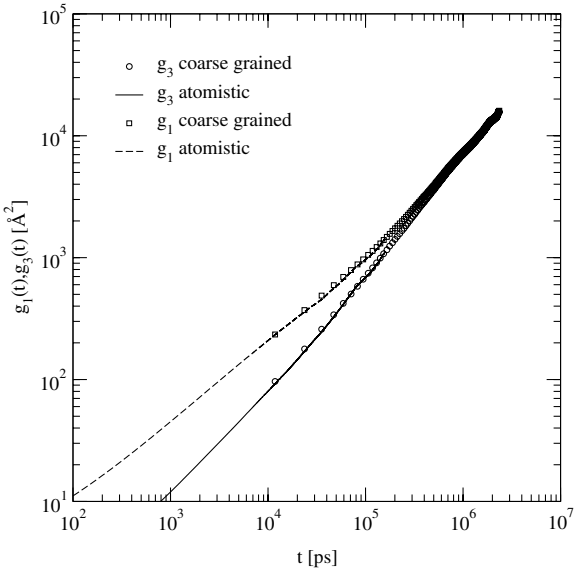


Fig. 17. Mean square displacements of the centers of mass and averaged over all repeat units or beads respectively for an N=5 melt of 50 chains in an all atom simulation and a bead spring simulation, leading to the absolute time scaling of the coarse grained bead spring simulation, from [54, 65]

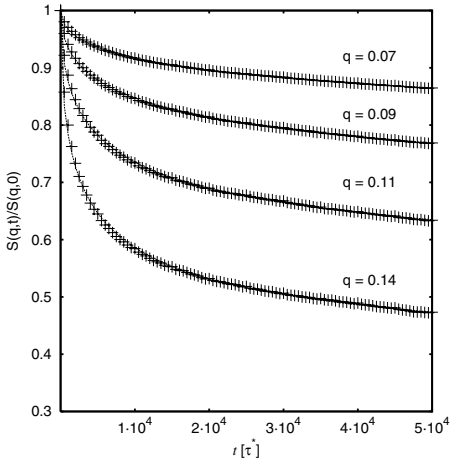


Fig. 18. Scaled dynamic structure factor $S(q,t)/S(q,0)$ as a function of time for the melt with chain length $N = 120$ repeat units. For each q-value, *dotted lines* show the fit to the expression of equation 32. Note that these curves are obtained for $4N + 3 = 483$ scattering beads per chain [54]

Table 2. Values obtained for N_e using the different approaches discussed in the text for fully flexible bead spring chains [51] and a coarse grained model of polycarbonate (BPA-PC) [54]. The more recent experiments on BPA-PC, which use better characterized samples and longer chains, clearly tend towards the lower value of N_e, for details see [54]. The maximum chain lengths studied are also indicated

	N_e^{BPA-PC}	$N_e^{bead-spring}$
$D(N)$	≈ 15	≈ 70
$g(t)$	6–10	32
$S(q,t)$	10	55
G_N^0	4	72
N_{max}	120	2000
experiment	5–10	

Taking the results form the simulation of BPA-PC and those from the simulations of flexible bead spring chains together illustrates the characteristic problem mentioned already before:

There seems to be no clear systematics behind the different result, indicating that one cannot use one experiment to propose the outcome of another experiment looking at a different quantity. It is thus highly desirable to obtain alternative methods to estimate N_e, which preferably should be just based on the conformations, rather than the time evolution of the melts or networks themselves.

6 Primitive Path Analysis: PPA

At the beginning of the whole chapter was the idea of topological constraints, which goes back to S. F. Edwards in 1967 [16] and let to the formulation of the tube model. This turned out to be an extremely successful concept, which significantly shaped our understanding of polymer and network dynamics. While the undisputable fact that the chains can pass along but not through each other leads to the described phenomena, the characteristic quantity of this concept, namely d_T or N_e remained essentially an adjustable parameter for the different means of investigation. Because of that there were already from the very beginning attempts to link the plateau modulus and N_e to the chain conformation and the packing of the chains in the melt. One set of publications deals with an analysis of the topological structure of a network and related approximations for melts [66]. For this the problem however is that multi chain/multi ring effects in all cases (Gauss integrals, Alexander polynomials ...) can only be taken into account up to a finite order. How important many chain effects can be is illustrated by Fig. 19 for an example of three chains only [44, 67].

In principle one would need for long chains an "infinite" hierarchy of such interactions. On the other hand, in order to observe that N_e is independent

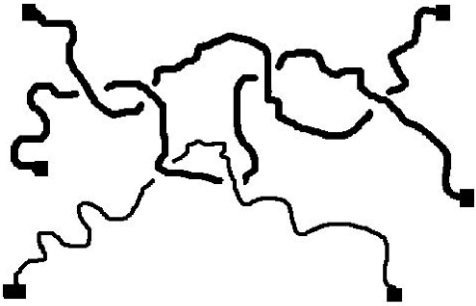


Fig. 19. Constraints between the two thick black chains depend on the presence of the 3rd (*thin*) chain

of N for $N \gg N_e$, a constant average density of motion constraints along the backbone of the chain, which provide the tube, is needed. Iwata and Edwards [67, 68] tried to overcome this by looking at the local values of the Gauss integral upon numerical integration. This suggests that the local chain-chain packing plays a crucial role. In parallel relatively early first attempts were made to relate N_e to the number of different chains within the volume d_T^3 [69,70]. All these attempts were not really satisfactory. More recently, based on a collection of many moduli data of different polymers, Fetters et al. [18,19] suggest that the modulus is related to the so called chain packing length $p = N/\rho\langle R\rangle$ in a way that $G \sim p^{-3}$. p is a measure for the spatial packing of the chains. Many polymers seem to follow that within a $\pm 20\%$ corridor. This however still does not explain how this packing is able to produce the entanglement mesh.

In order to understand that, we go back to the original idea of Edwards [16] who identified the random walk-like axis of the tube with what he called the "primitive" path: the shortest path between the endpoints of the original chain into which its contour can be contracted without crossing any obstacle. Similar to the tube, the primitive path is usually discussed without specifying the relation between the obstacles and the melt structure. However the obstacles themselves are chains, which have to be taken into account, when constructing the primitive path [20, 71, 72]. Of course Edwards himself was aware of this severe simplification. In [20] is introduced a *primitive path analysis* (PPA) where all polymers in the system are contracted simultaneously in a way, that a linear stress strain relationship is kept. This allows to establish the microscopic foundations of the tube model and to endow a highly successful phenomenological model with predictive power for structure-property relations. The primitive path analysis is performed on a variety of different simple bead spring chain melts [39] with variable intrinsic stiffness [34] and semi dilute solutions [73]. All these cases include the crucial ingredients characteristic for polymeric systems: connectivity, flexibility, local liquid-like monomer packing

and mutual uncrossability of the chain backbones. The parameters used for the PPA are given in [71].

Monodisperse polymer melts of $M = 200 - 500$ chains of length $50 \leq N \leq 700$ at a bead density of $\rho = 0.85\ \sigma^{-3}$ (in Lennard Jones units) are studied. By introducing a small intrinsic bond bending potential, the Kuhn length, $l_K = \langle R^2 \rangle / L$, is varied between $1.80\ \sigma$ and $3.34\ \sigma$. $\langle R^2 \rangle$ is the mean squared end to end distance of the chain and L is the chain contour length. For details see [34]. In addition, we present results for a coarse-grained model for polycarbonate (BPA-PC) [54, 64]. In this case, we analyzed melt configurations for $M = 100$ chains of $N = 60$ chemical repeat units which are represented by four beads each. All the melt samples studied are tabulated in [71]. Figure 20 shows a characteristic example, of how such a primitive path network looks like for the case of an entangled BPA-PC melt. The implementation of the PPA is straightforward: First the chain-ends are fixed in space. Then, all intrachain interactions other than the FENE bond interaction, which has its minimum at $r = 0$, are switched off. Finally the total energy is minimized by cooling the system slowly to $T \to 0$. Without thermal fluctuations and intra-chain EV interactions, the bond springs try to reduce the bond length to zero and pull the chains taut. The interchain EV interactions provide an energy barrier to prevent chain crossings. Needless to say, a crucial ingredient for the success of the whole procedure is the availability of well equilibrated

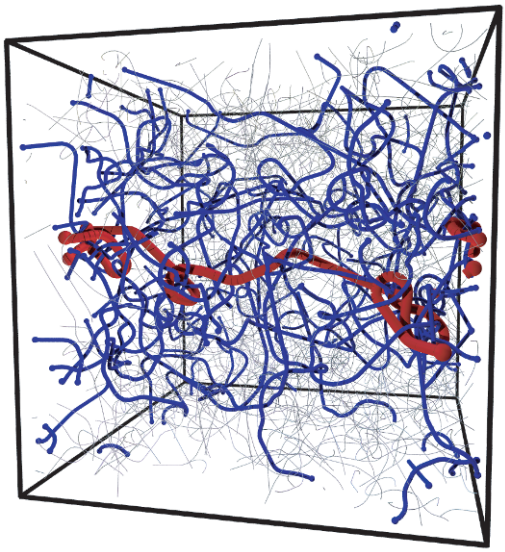


Fig. 20. Result of the primitive path analysis for all chains in the melt of 100 BPA-PC chains of $N = 60$ repeat units (243 beads). We show the primitive path of one chain (*thick, pale*) together with all of those it is entangled with (*thin black*). The primitive paths of all other chains in the system are shown as thin lines [54]

conformations [33,34]! For details of the PPA protocol we refer to [71]. The algorithm introduced so far does not account for self-entanglements. In a recent variant of the approach this is included [71]. Recently some variants have been proposed,which do not keep the excluded volume of the chain, but reduce the diameter of the strands to zero [74–76]. They however lead to different results and do not fit to the experiment as well as the present approach(see below).

In order to understand the relationship between the melt structure and properties like the plateau modulus for many different polymers and simple model systems we need a dimensionless way to compare data. To render the G_N^0 values dimensionless, we need an energy scale and a length scale. As the dominant contribution to G_N^0 is of entropic origin, k_BT as the energy scale suggests itself. However, there are essentially two independent length scales that characterize the local structure of polymer melts. One, the Kuhn length, l_K [1] is defined for random walks as the length of an individual step of a freely jointed chain with the same mean-square end-to-end distance $R^2 = l_K L$ and contour length L. Random walks however do not densely fill space. Locally most monomers will belong to the same chain. The length at which the polymers start to interpenetrate is given by the packing length, $p = (\rho_{chain} R^2)^{-1}$. The packing length can be visualized as the average strand strand distance. It is relevant to note that the product of the number density of chains, ρ_{chain}, and of R^2 is independent of chain length for a fixed monomer number density, ρ. Following the standard convention [77] based on the chemical structure of the polymers, we choose l_K as the unit of length. With this choice, instead of G_N^0 we can consider the dimensionless quantity $G_N^0 l_K^3/k_BT$. This quantity has to be a function of the only remaining dimensionless parameter: the ratio of Kuhn and packing length l_K/p.

The available experimental data for dense melts are the result of a long term substantial experimental effort [19, 78]. Rheological measurements were performed on these samples and in many cases the $\langle R^2 \rangle$ values were determined using small angle neutron scattering. References. [19, 78] provide the values for the plateau modulus, G_N^0, the mass density, ρ_m, the ratio of the mean-square end-to-end distance to the molecular weight of the polymer, R^2/M, and the packing length, $p = M/(R^2 \rho_m N_A)$, where N_A is the Avagadro number. All the melt data points obey the empirical relation $G_N^0 = 0.00226 k_BT/p^3$, indicated as the dashed line in Fig. 21. The way to determine l_k for experimental systems, depends on the definition of the contour length. It is however worthwhile to note, if the empirical relation is valid, then any choice for l_K will preserve the scaling relationship between G_N^0 and p. The actual choice for l_K will determine the position of the data points in Fig. 21 and will just move them along the dashed line in the figure. All the melt data shown in the figure are tabulated in [71], where also a more detailed discussion of the mapping and the length scales can be found.

The aforementioned mapping can be also be used to identify bead spring polymer models to individual chemical species. The standard model [39] with fully flexible polymer chains used in molecular dynamics simulations corre-

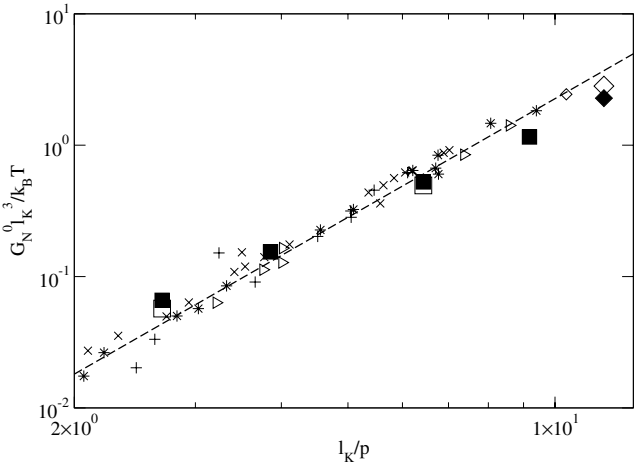


Fig. 21. Dimensionless plateau moduli $G_N^0 l_K^3/k_B T$ as a function of the dimensionless ratio l_K/p of Kuhn length l_K and packing length p. Inset (a) contains (i) experimentally measured plateau moduli for polymer melts [18] ($*$ Polyolefins, $\times$ Polydienes, $+$ Polyacrylates and $\triangleright$ miscellaneous, (ii) plateau moduli inferred from the normal tensions measured in computer simulation of bead-spring melts [20, 51] ($\square$) and a semi-atomistic polycarbonate melt [54] ($\lozenge$) under an elongational strain, and (iii) predictions of the tube model 34 based on the results of our PPA for bead-spring melts ($\blacksquare$), and the semi-atomistic polycarbonate melt ($\blacklozenge$). The line indicates the best fit to the experimental melt data for polymer melts by Fetters et al. [19]. Errors for all the simulation data are smaller than the symbol size

sponds to $l_K/p = 2.68$. Among the available experimental data, the chemical species with the closest l_K/p value is natural rubber (cis-Polyisoprene) with $l_K/p = 2.72$. This suggests that the elastic properties of the usually studied bead spring polymer model corresponds most closely to that of natural rubber [20], the prototypical experimental system for elastic behavior.

Having provided a method to determine the primitive paths for melt configurations of entangled bead-spring model polymers, we can now test the predictive power of the tube model. To this end, we use the standard expression [1, 20]

$$G_N^0 = \frac{4}{5}\frac{k_B T}{p\, d_T^2} = \frac{4}{5}\frac{\rho k_B T}{N_e}\,, \tag{34}$$

which relates the plateau modulus to the Kuhn length of the primitive path. The d_T values can be obtained from the contour length of the primitive path, L_{pp} (which is the sum of the lengths of all the back bone bonds of a chain averaged over all the chains in the melt) using $d_T = R^2/L_{pp}$. For details refer [20, 71]. Figure 21 shows an explicit comparison of the dimensionless plateau moduli $G_N^0 l_K^3/k_B T$ of experimental systems and bead-spring model polymers with identical ratios l_K/p. The good agreement of the two data

sets confirms the insensitivity of entanglement effects to atomistic details and provides the necessary validation of our generic bead-spring models. Based on the above analysis Everaers et al. [20] developed a scaling scheme, which actually showed that within this primitive path picture, $G_N^0 \sim p^{-3}$.

7 Summary

This contribution tried to introduce the reader to the current status of computer simulations for one central aspect of dense polymer systems, namely melt and network dynamics and relaxation. Over the years there has been significant progress in theory, simulation and experiments. This leads us to the situation, where the basic concepts are fairly well understood. We now can go on and study new phenomena. Especially simulations reach a stage, where they more and more are of predictive power with clear advantages to analytic theory but also experiment. It is probably safe to expect, that this kind of development will continue for the coming years. Thus we are a step closer to the aim of material characterization and design by intelligent simulation techniques.

8 Annotation

The present review is an updated and in some places extended version of [79].

Acknowledgements

Over the years the author has had very enjoyable collaborations with many students and colleagues, who cannot all be mentioned here. During all these years however the very fruitful collaboration with G. S. Grest and R. Everaers remained active.

References

1. M. Doi and S. F. Edwards (1986) *The Theory of Polymer Dynamics.* Clarendon, Oxford
2. T. C. B. McLeish (2002) Tube Theory of Entangled Polymer Dynamics. *Adv. Phys.* **5**, pp. 1379–1527
3. K. Kremer and G. S. Grest (1995) In K. Binder, ed., *Monte Carlo and Molecular Dynamics Simulations in Polymer Science*, p. 194, Oxford University Press, New York
4. B. Dünweg (2006) Mesoscopic Simulations for Problems with Hydrodynamics, with Emphasis on Polymer Dynamics. *Lect. Notes Phys.* **704**, pp. 313–342, Springer-Verlag

5. C. F. Abrams and K. Kremer (2002) Effects of excluded volume and bond length on the dynamics of dense bead-spring polymer melts. *J. Chem. Phys.* **116**, p. 3162
6. K. Kremer (2000) *Soft and Fragile Matter, Nonequilibrium Dynamics, Metastability and Flow.* NATO ASI Workshop, St. Andrews, M. E. Cates, M. R. Evans eds., Institute of Physics, London
7. P. G. de Gennes (1979) *Scaling Concepts in Polymer Physics.* Cornell University Press, Ithaca NY
8. M. Rubinstein and R. H. Colby (2003) *Polymer Physics.* Oxford University Press, Oxford
9. A. R. Khokhlov and A. Yu Grosberg (1994) *Statistical Physics of Macromolecules.* AIP Press, New York
10. K. Binder (1995) In K. Binder, ed., *Monte Carlo and Molecular Dynamics Simulations in Polymer Science*, p. 356. Oxford University Press, New York
11. P. E. Rouse (1953) A Theory of the Linear Viscoelastic Properties of Dilute Solutions of Coiling Polymers. *J. Chem. Phys.* **21**, p. 1272
12. F. Bueche (1954) The Viscoelastic Properties of Plastics. *J. Chem. Phys.* **22**, p. 603
13. W. Paul, G. D. Smith, D. Y. Yoon, B. Farago, S. Rathgeber, A. Zirkel, L. Willner, and D. Richter (1998) Chain motion in an unentangled polyethylene melt: A critical test of the rouse model by md simulations and neutron spin echo spectroscopy. *Phys. Rev. Lett.* **80**, p. 2346
14. H. Tao, T. P. Lodge, and E. D. von Meerwall (2000) Diffusivity and Viscosity of Concentrated Hydrogenated Polybutadiene Solutions. *Macromolecules* **33**, p. 1747
15. J. D. Ferry (1994) *Viscoelastic Properties of Polymers.* Wiley, New York
16. S. F. Edwards (1967) The statistical mechanics of polymerized material. *Proc. Phys. Soc.* **92**, pp. 9-16
17. P. G. de Gennes (1971) Reptation of a Polymer Chain in the Presence of Fixed Obstacles. *J. Chem. Phys.* **55**, p. 572
18. L. J. Fetters, D. J. Lohse, S. T. Milner, and W. W. Graessley (1999) Packing Length Influence in Linear Polymer Melts on the Entanglement, Critical, and Reptation Molecular Weights. *Macromolecules* **32**, p. 6847
19. L. J. Fetters, D. J. Lohse, and W. W. Graessley (1999) Chain dimensions and entanglement spacings in dense macromolecular systems. *J. Poly. Sci. B: Pol. Phys.* **37**, p. 1023
20. R. Everaers, S. K. Sukumaran, G. S. Grest, C. Svaneborg, A. Sivasubramanian, and K. Kremer (2004) Rheology and Microscopic Topology of Entangled Polymeric Liquids. *Science* **303**, p. 823
21. T. P. Lodge (1999) Reconciliation of molecular weight dependence of diffusion and viscosity in entangled polymer. *Phys. Rev. Lett.* **83**, p. 3218
22. S. T. Milner and T. C. B. McLeish (1998) Star polymers and failure of time-temperature superposition. *Macromolecules* **31**, p. 8623
23. M. Appel and G. Fleischer (1993) Investigation of the chain length dependence of self-diffusion of poly(dimethylsiloxane) and poly(ethylene oxide) in the melt with pulsed field gradient NMR. *Macromolecules* **26**, p. 5520
24. P. T. Callaghan and A. Coy (1992) Evidence for reptational motion and the entanglement tube in semidilute polymer solutions Paul T. Callaghan and Andrew Coy. *Phys. Rev. Lett.* **68**, p. 3176

25. B. Ewen and D. Richter (1997) Neutron Spin Echo Investigations on the Segmental Dynamics of Polymers in Melts, Networks and Solutions. *Adv. Pol. Sci.* **134**, pp. 1-129
26. D. S. Pearson (1987) Recent Advances in the Molecular Aspects of Polymer Viscoelasticity. *Rubber Chem. Tech.* **60**, p. 439
27. D. S. Pearson, L. J. Fetters, W. W. Graessley, G. ver Strate, and E. von Meerwall (1994) Viscosity and self-diffusion coefficient of hydrogenated polybutadiene. *Macromolecules* **27**, p. 711
28. M. Appel, G. Fleischer, J. Kärger, F. Fujara, and I. Chang (1994) Anomalous Segment Diffusion in Polymer Melts. *Macromolecules* **27**, p. 4274
29. T. P. Russell, V. R. Deline, W. D. Dozier, G. P. Felcher, G. Agrawal, R. P. Wool, and J. W. Mays (1993) Direct observation of reptation at polymer interfaces. *Nature* **365**, p. 235
30. A. Wischnewski, M. Monkenbusch, L. Willner, D. Richter, and G. Kali (2003) Direct observation of the transition from free to constrained single-segment motion in entangled polymer melts. *Phys. Rev. Lett.* **90**, p. 058302
31. S. F. Edwards and T. A. Vilgis (1988) The tube model theory of rubber elasticity. *Rep. Prog. Phys.* **51**, p. 243
32. K. Kremer (1996) *Computer simulation methods for polymer physics. Ital. Phys. Soc.* Bologna
33. R. S. Hoy and M. O. Robbins (2005) Effect of equilibration on primitive path analyses of entangled polymers. *Phys. Rev. E* **72**, p. 061802
34. R. Auhl, R. Everaers, G. S. Grest, K. Kremer, and S. J. Plimpton (2003) Equilibration of long chain polymer melts in computer simulations. *J. Chem. Phys.* **119**, p. 12718
35. M. Pütz (1999) *Dynamik von Polymerschmelzen und Quellverhalten ungeordneter Netzwerke.* Ph.D. Thesis, University of Mainz
36. V. G. Mavrantzas, T. D. Boone, E. Zervopoulou, and D. N. Theodorou (1999) End-Bridging Monte Carlo: A Fast Algorithm for Atomistic Simulation of Condensed Phases of Long Polymer Chains. *Macromolecules* **32**, p. 5072
37. A. Uhlherr, S. J. Leak, N. E. Adam, P. E. Nyberg, M. Doastakis, V. G. Mavrantzas, and D. N. Theordorou (2002) Large scale atomistic polymer simulations using Monte Carlo methods for parallel vector processors. *Comp. Phys. Comm.* **144**, p. 1
38. G. S. Grest and K. Kremer (1986) Molecular dynamics simulation for polymers in the presence of a heat bath. *Phys. Rev. A* **33**, p. 3628
39. K. Kremer and G. S. Grest (1990) Dynamics of entangled linear polymer melts: A molecular-dynamics simulation. *J. Chem. Phys.* **92**, p. 5057
40. T. Soddemann, B. Dünweg, and K. Kremer (2003) Dissipative particle dynamics: A useful thermostat for equilibrium and nonequilibrium molecular dynamics simulations. *Phys. Rev. E* **68**, p. 046702
41. E. R. Duering, K. Kremer, and G. S. Grest (1991) Relaxation of randomly cross-linked polymer melts. *Phys. Rev. Lett.* **67**, p. 3531
42. E. R. Duering, K. Kremer, and G. S. Grest (1994) Structure and relaxation of end-linked polymer networks *J. Chem. Phys.* **101**, p. 8169
43. M. Pütz, R. Everaers, and K. Kremer (2000) Self-similar chain conformations in polymer gels. *Phys. Rev. Lett.* **84**, p. 298
44. R. Everaers and K. Kremer (1996) Topological interactions in model polymer networks. *Phys. Rev. E* **53**, p. R37

45. R. Everaers (1999) Entanglement effects in defect-free model polymer networks *New J. Phys.* **1**, pp. 12–1
46. W. Paul, K. Binder, D. W. Heermann, and K. Kremer (1991) Dynamics of polymer solutions and melts. Reptation predictions and scaling of relaxation times. *J. Chem. Phys.* **95**, p. 7726
47. W. Paul, K. Binder, K. Kremer, and D. W. Heermann (1991) Structure Property Correlation of Polymers, a Monte-Carlo Approach. *Macromolecules* **24**, p. 6323
48. S. W. Smith, C. K. Hall, and B. D. Freeman (1996) Molecular dynamics study of entangled hard-chain fluids. *J. Chem. Phys.* **104**, p. 5616
49. M. Kröger and H. Voigt (1994) On a quantity describing the degree of entanglement in linear polymer systems. *Macromol. Theory Simul.* **3**, p. 639
50. K. Kremer, G. S. Grest, and I. Carmesin (1988) Crossover from Rouse to Reptation Dynamics: A Molecular-Dynamics Simulation. *Phys. Rev. Lett.* **61**, p. 566
51. M. Pütz, K. Kremer, and G. S. Grest (2000) What is the entanglement length in a polymer melt? *Europhys. Lett.* **49**, p. 735
52. R. Faller, F. Müller-Plathe, and A. Heuer (2000) Local Reorientation Dynamics of Semiflexible Polymers in the Melt. *Macromolecules* **33**, p. 6602
53. R. Faller and F. Müller-Plathe (2001) Chain stiffness intensifies the reptation characteristics of polymer dynamics in the melt. *Chem. Phys. Chem.* **2**, p. 180
54. S. Leon, L. D. Site, N. van der Vegt, and K. Kremer (2005) Bisphenol A Polycarbonate: Entanglement Analysis from Coarse-Grained MD Simulations. *Macromolecules* **38**, p. 8078
55. P. G. de Gennes (1981) Coherent Scattering by One Reptating Chain. *J. Phys. (Paris)* **42**, p. 735
56. K. Kremer and K. Binder (1984) Dynamics of polymer chains confined into tubes: Scaling theory and Monte Carlo simulations. *J. Chem. Phys.* **81**, p. 6381
57. M. Doi (1988) An estimation of the tube radius in the entanglement effect of concentrated polymer solutions. *J. Phys. A* **8**, p. 959
58. M. Pütz, K. Kremer, and G. S. Grest (2000) Reply to the Comment by A. Wischnewski and D. Richter on "What is the entanglement length in a polymer melt? *Europhys. Lett.* **52**, p. 721
59. L. R. G. Treloar (1986) *The Physics of Rubber Elasticity.* Clarendon Press, Oxford
60. M. Mooney (1940) A theory of large elastic deformation. *J. Appl. Phys.* **11**, p. 582
61. M. Kröger and S. Hess (2000) Rheological evidence for a dynamical crossover in polymer melts via nonequilibrium molecular dynamics. *Phys. Rev. Lett.* **85**, p. 1128
62. K. Sommer, J. Batoulis, W. Jilge, L. Morbitzer, B. Pittel, R. Plaetschke, K. Reuter, R. Timmermann, K. Binder, W. Paul, F. T. Gentile, D. W. Heermann, K. Kremer, M. Laso, U. W. Suter, and P. J. Ludivice (1991) Correlation between primary chemical structure and property phenomena in polycondensates. *Adv. Mat.* **3**, p. 590
63. J. Baschnagel, K. Binder, P. Doruker, A. A. Gusev, O. Hahn, K. Kremer, W. L. Mattice, F. Müller-Plathe, M. Murat, W. Paul, S. Santos, U. W. Suter, and V. Tries, *Advances in Polymer Science: Viscoelasticity, Atomistic Models, Statistical Chemistry.* Springer-Verlag, Heidelberg
64. C. F. Abrams and K. Kremer (2003) Combined Coarse-Grained and Atomistic Simulation of Liquid Bisphenol A-Polycarbonate: Liquid Packing and Intramolecular Structure. *Macromolecules* **36**, p. 260

65. B. Hess. S. Leon, N. van der Vegt, and K. Kremer (2006) Long time atomistic polymer trajectories from coarse grained simulations. Bisphenol-A polycarbonate. *Softmatter* **2**, p. 409
66. W. Michalke, M. Lang, S. Kreitmeier, and D. Göritz (2001) Simulations on the number of entanglements of a polymer network using knot theory. *Phys. Rev. E* **64**, p. 012801
67. K. Iwata and S. F. Edwards (1989) New model of polymer entanglement: Localized Gauss integral model. Plateau modulus GN, topological second virial coefficient A and physical foundation of the tube model. *J. Chem. Phys.* **90**, p. 4567
68. K. Iwata (1991) Topological origin of reptation: a collective motion of local knots. *Macromolecules* **24**, p. 1107
69. T. A. Kavassalis and J. Noolandi (1987) New View of Entanglements in Dense Polymer Systems. *Phys. Rev. Lett.* **59**, p. 2674
70. T. A. Kavassalis and J. Noolandi (1989) Entanglement scaling in polymer melts and solutions. *Macromolecules* **22**, p. 2709
71. S. K. Sukumaran, G. S. Grest, K. Kremer, and R. Everaers (2004) Identifying the primitive path mesh in entangled polymer liquids. *J. Polym. Sci. Part B: Polym. Phys. Ed.* **41**, p. 917
72. M. Rubinstein and E. Helfand (1985) Statistics of the entanglement of polymers: Concentration effects. *J. Chem. Phys.* **82**, p. 2477
73. P. Ahlrichs, R. Everaers, and B. Dünweg (2001) Screening of hydrodynamic interactions in semidilute polymer solutions: A computer simulation study. *Phys. Rev. E* **64**, p. 040501
74. D. Theodorou (2005) *Comp. Phys. Comm.* **169**, p. 82
75. Q. Zhou and R. G. Larson (2005) Primitive path identification and statistics in molecular dynamics simulations of entangled polymer melts. *Macromolecules* **38**, p. 5761
76. M. Kröger (2005) Shortest multiple disconnected path for the analysis of entanglements in two- and three-dimensional polymeric systems. *Comp. Phys. Comm.* **168**, p. 209
77. W. W. Graessley and S. F. Edwards (1981) Entanglement interactions in polymers and the chain contour concentration. *Polymer* **22**, p. 1329
78. L. J. Fetters, D. J. Lohse, D. Richter, T. A. Witten, and A. Zirkel (1994) Connection between Polymer Molecular Weight, Density, Chain Dimensions, and Melt Viscoelastic Properties . *Macromolecules* **27**, p. 4639
79. K. Kremer (1992) In H. Grubmüller, K. Kremer, N. Attig, and K. Binder, eds., *Computational Soft Matter: From Synthetic Polymers to Proteins, Lecture Notes*, p. 400, NIC, FZ Jülich, Jülich

Reaction Kinetics of Coarse-Grained Equilibrium Polymers: A Brownian Dynamics Study

C.-C. Huang[1,2], H. Xu[2], F. Crevel[3], J. Wittmer[3], and J.-P. Ryckaert[1]

[1] Physique des polymères, CP223, Université Libre de Bruxelles, Bv du Triomphe, 1050 Brussels, Belgium
jryckaer@ulb.ac.be
[2] LPMD, Inst. Physique-Electronique, Univ. Paul Verlaine-Metz, 1bd Arago, 57078 Metz cedex 3, France
hongxu@univ-metz.fr
[3] Institut Charles Sadron, 6 Rue Boussingault,67083 Strasbourg, France
jwittmer@ics.u-strasbg.fr



Jean-Paul Ryckaert

C.-C. Huang et al.: *Reaction Kinetics of Coarse-Grained Equilibrium Polymers: A Brownian Dynamics Study*, Lect. Notes Phys. **704**, 379–418 (2006)
DOI 10.1007/3-540-35284-8_15 © Springer-Verlag Berlin Heidelberg 2006

1 Introduction 381

2 Theoretical Framework 384

2.1 Statistical Mechanics Derivation of the Distribution of Chain Lengths 384
2.2 A Kinetic Model for Scissions and Recombinations 389

3 The Mesoscopic Model and Its Brownian Dynamics Implementation 392

4 Characterization of Static Properties 398

4.1 List of Experiments and Chain Length Distributions 398
4.2 Chain Length Conformational Analysis 401

5 Kinetics Analysis 403

5.1 Distribution of First Recombination Times 404
5.2 Cumulative Hazard Analysis 406
5.3 The Reactive Flux Correlation Function and the Transmission Coefficient 409
5.4 Estimation of Rate Constants: Comparison of the Various Methods 411
5.5 Analysis of the Monomer Diffusion 412
5.6 Analysis of the First Recombination Times Distribution and Corresponding Diffusive Steps 414

6 Summary and Conclusions 416

References 417

Self-assembled linear structures like giant cylindrical micelles or discotic molecules in solution stacked in flexible columns are systems reminiscent of polydisperse polymer solutions, ranging from dilute to concentrated solutions as the overall monomer density and/or the chain length increases. These supramolecular polymers have an equilibrium length distribution, the result of a competition between the random breakage of chains and the fusion of chains to generate longer ones. This scission-recombination mechanism is believed to be responsible of some peculiar dynamical properties like the Maxwell fluid rheological character for entangled micelles. Simulations employing simple mesoscopic models provide a powerful approach to test mean-field theories or scaling approaches which have been proposed to rationalize the structural and kinetic properties of these soft matter systems. In the present work, we review the basic theoretical concepts of these "equilibrium polymers" and some of the important results obtained by simulation approaches. We propose a new version of a mesoscopic model in continuous space based on the bead and FENE spring polymer model which is treated by Brownian Dynamics and Monte-Carlo binding/unbinding reversible changes for adjacent monomers in space, characterized by an attempt frequency parameter ω. For a dilute and a moderately semi-dilute state-points which both correspond to dynamically unentangled regimes, the dynamic properties are found to depend upon ω through the effective life time τ_b of the average size chain which, in turn, yields the kinetic reaction coefficients of the mean-field kinetic model proposed by Cates. Simple kinetic theories seem to work for times $t \geq \tau_b$ while at shorter time, strong dynamical correlation effects are observed. Other dynamical properties like the overall monomer diffusion and the mobility of reactive end-monomers are also investigated.

1 Introduction

In the field of self-assembling structures, supramolecular polymers are attracting nowadays much attention [1]. It is well known and schematically illustrated in Fig. 1 that some surfactant molecules in solution can self-assemble and form wormlike micelles [2].

Such micellar solutions exhibit fascinating rheological behaviour, such as shear-banding [3], shear-thickenning [4], Maxwell fluid behaviour [2] or anomalous diffusion (Levy flight) [5]. Self-assembled stacks of discotic molecules and chains of bifunctional molecules are other examples of supramolecular polymers. All these examples differ by the nature of the intermolecular forces involved in the self-assembling of the basic units, but they lead to a similar physical situation bearing much analogy with a traditional system of polydisperse flexibles polymers when their length becomes sufficiently large with respect to their persistence length. The specificity and originality of these supramolecular polymers comes from the fact that these chains are continuously subject to scissions at random places along their contour and subject

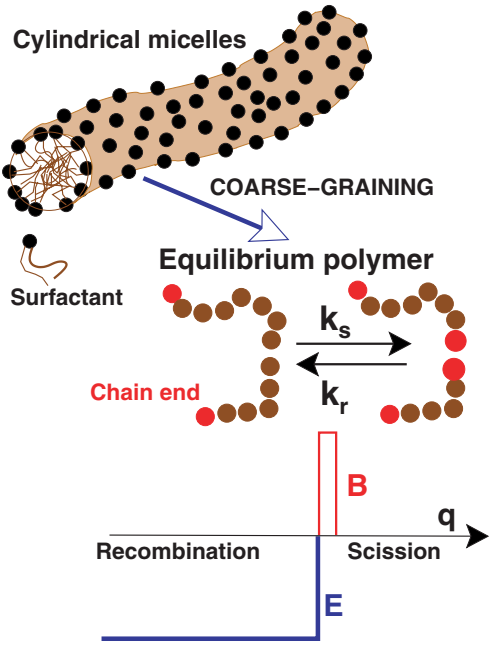


Fig. 1. Some surfactant molecules in solution self-assemble and form long wormlike micelles which continously break and recombine. Their mass distribution is, hence, in thermal equilibrium and they present an important example of the vaste class of systems termed "equilibrium polymers" [1].The free energy E of the (spherical) end cap of these micelles has been estimated [2] to be of order of $10\,k_BT$. This energy penalty (together with the monomer density) determines essentially the static properties and fixes the ratio of the scission and recombination rates, k_s and k_r. Additionally, these rates are influenced by the barrier height B which has been estimated to be similar to the end cap energy. Both important energy scales have been sketched schematically as a function of a generic reaction coordinate q (see Chap. 8 of reference [19]). Following closely the analytical description [2, 13] these micellar systems are represented in this study by coarse-grained effective potentials in terms of a standard bead-spring model. The end cap free energy becomes now an energy penalty for scission events, i.e., the creation of two unsaturated chain ends. The dynamical barrier is taken into account by means of an attempt frequency $\omega = \exp(-B/k_BT)$. If ω is large, successive breakage and recombination events for a given chain can be assumed to be uncorrelated and the recombination of a newly created chain ends will be of standard mean-field type. On the other hand, the (return) probability that two newly created chain ends recombine immediately must be particulary important at large ω. These highly correlated "diffusion controlled" [12] recombination events do not contribute do the effective macroscopic reaction rates which determine the dynamics of the system. The key task of this paper is to compare different methods to compute the effective rates as a function of the attempt frequency ω for a weakly semi-dilute state point in the non-entangled dynamical regime

to end to end recombinations, leading to a dynamical equilibrium between different chain lengths species. These supramolecular polymers are typical soft matter systems and the chain length distribution which determines their properties, is very sensitive to external conditions (temperature, concentration, external fields, salt contents, etc...).

Different times scales and length scales are involved in these systems, as in usual polymer systems. Unlike the latter, which have been much studied during the last 50 years [6–8], theoretical work on micellar solutions are quite recent [1, 2]. While building up a molecular scale theory is clearly too heavy for such complex systems, theoretical approaches, based on mean-field concept and non-local phenomenological approaches, can explore with success some equilibrium, and rheological properties [9–11], under reasonable but often drastic approximations. Concerning specifically micelle kinetics, recently, O'Shaughnessy and Yu [12] suggested that there are two possible kinds of kinetics associated with scission/recombination: diffusion controlled and mean-field, which may be distinguished by the time dependence of the first recombination times distribution function of the chains.

In parallel, mesoscopic scale computer simulations can shed much light on these rich but complex systems, thanks to techniques borrowed from simulations of polymers. Systems of wormlike micelles can be modelled by "Equilibrium polymers" (EP), sometimes called "living polymers", which are polymer chains endowed with scission/recombination (S-R) processes taking place in them. The advantage of numerical studies is their ability to make links either to experiments or to mean-field type theories, given that those systems are difficult to characterize by experiments and the results are not simple to interpret. The bond fluctuation model (BFM) has been the object of intensive studies on the statics and dynamics of living polymers [13, 14].

Brownian dynamics studies of a similar model [15] have been used to study the scaling predictions on the dependence of the average chain length upon the overall monomer density. In a recent study, Padding and Boek [16] showed that a mesoscopic model of wormlike micelle known as the FENE-C model [17], seems to obey the diffusion-controlled kinetics.

In our work [18] which is the basis for this course, we investigate the structural and kinetic properties of equilibrium polymers in solution by means of a Brownian Dynamics algorithm applied to a standard polymer model [20] coupled to a time-reversible EP algorithm for polymer scission-recombinations [13,14,21–23]. This particular continuous space model has the advantage that its kinetic properties are well separated from the structural and thermodynamical aspects. Our model does not allow for cyclic structures and hence, we simplify the computational load and, more importantly, the theoretical interpretation of the static and dynamical properties of our samples. Anyway, for cylindrical micelles, cyclic structures are expected to exist in very low concentrations [2].

We study the kinetics of equilibrium polymers solutions at two state points, exploring in each case a range of rates of the scission-recombination process.

The two state points correspond to a dilute solution and a semi-dilute solution regime close to the cross-over region. A direct measure of chain overlap can be provided by the ratio of the mean distance between chain centres of mass over the radius of giration of the average size chain (respectively ≈ 10 and ≈ 50 for the dilute and semi-dilute cases). This gives ratios of 1.4 for the dilute case and 0.6 for the semi-dilute case, which suggests that even in the semi-dilute case, the chain relaxation is still typically Rouse like and thus yet kinetically entangled.

The main focus is to link the microscopic model to the more macroscopic kinetics theories, e.g. by Cates [2], and to discuss in a well defined case the concepts of "diffusion controlled" and "mean-field", defined by authors of [12,16]. In particular, we want to estimate the effective reaction rates for the scission and recombination processes comparing different methods. Furthermore, we shall study the diffusion of monomers and discuss their link to the scission/recombination kinetics.

2 Theoretical Framework

2.1 Statistical Mechanics Derivation of the Distribution of Chain Lengths

To treat a system of supramolecular polymers theoretically, it is convenient to work at the mesoscopic scale using a model of linear flexible polymers made of L monomers of size b linked together by a non permanent bonding scheme. Within the system, individual chain lengths fluctuate by bond scission and by fusion of two chain ends of different chains. Statistical mechanics can be employed to predict the equilibrium distribution of chain lengths [2]. In terms of the equilibrium chain number density $c_0(L)$, the average chain length L_0 and the total monomer density ϕ are given by

$$L_0 = \frac{\sum_{L=0}^{\infty} L c_0(L)}{\sum_{L=0}^{\infty} c_0(L)} \tag{1}$$

$$\phi = \sum_{L=0}^{\infty} L c_0(L) = \frac{M}{V} \tag{2}$$

where M denotes the total number of monomers in the system.

Conceptually, we consider the Helmholtz free energy $F(V, T; \{N(L)\}, N_s)$ of a mixture of chains molecules of different length L in solvent where, in addition to the temperature T, the volume V and the number of solvent molecules N_s, the number of chains of each specific length $N(L)$ is fixed. Let $F(V, T; M, N_s)$ be the Helmholtz free energy of a similar system where only the total number of solute monomers M is fixed. The equilibrium chain length distribution $c_0(L) = N(L)/V$ will result from the set $\{N(L)\}$ which satisfies the condition

$$F(V,T;M,N_s) = min_{\{N(L)\}}\left[F(V,T;\{N(L)\},N_s) + \mu\sum_{L=0}^{\infty} LN(L)\right] \quad (3)$$

The parameter μ is the Lagrange multiplier associated with the constraint that individual numbers of chains $N(L)$ must keep fixed the total number of monomers $M = \sum_0^\infty LN(L)$. Minimisation requires that the first derivative with respect to any $N(L')$ variable (L'=1,2, ...) is zero, giving

$$\frac{\delta F(V,T;\{N(L)\},N_s)}{\delta N(L')} + \mu L' = 0; L' = 1,2,\ldots . \quad (4)$$

We expect the entropic part of the total free energy $F(V,T;\{N(L)\},N_s)$ to be the sum of translational and chain internal configurational contributions which both depend upon the way the M monomers are arranged into a particular chain size distribution. For the translation part, the polydisperse system entropy is estimated as the ideal mixture entropy S_{id}

$$S_{id}(V,T;\{N(L)\},N_s) = -k_B\sum_L N(L)\ln\left(CN(L)\right) + S^{solv} \quad (5)$$

where $C = b^3/V$ is a dimensionless constant independent of L and where S^{solv} is the solvent contribution, independent of the $N(L)$ distribution. The configurational entropy of an individual chain with L monomers is written as $S_1(L) = k_B \ln \Omega_L$, in terms of Ω_L, the total number of configurations of the chain. Adding the configurational contributions to S_{id} as given by(5), the total entropy becomes

$$S(V,T;\{N(L)\},N_s) = -k_B\sum_L N(L)\left[\ln\left(CN(L)\right) - \ln\Omega_L\right] + S^{solv} \quad (6)$$

We now turn to the energy $E(V,T;\{N(L)\},N_s)$ of the same system. If $E_1(V,T;L)$ represents the internal energy of a chain of L monomers and E_s the energy of a solvent molecule, the energy can be written as

$$E(V,T;\{N(L)\},N_s) = \sum_L N(L)E_1(V,T;L) + N_sE_s(V,T) \quad (7)$$

The key contribution in E_1 is the chain end-cap energy E which corresponds to the chain end energy penalty required to break a chain in two pieces. We will suppose that $E_1(L) = E + L\tilde{\epsilon}$ where $\tilde{\epsilon}$ is an irrelevant energy per monomer as $M\tilde{\epsilon}$, its total contribution to the system energy, is independent of the chain length distribution.

The present approximation of the total free energy of the system is thus given by incorporating in the general expression (3) the expressions (6) and (7), giving

$$\beta F(V,T;\{N(L)\},N_s) = \sum_L N(L)\left[\ln N(L) + \ln C - \ln\Omega_L + \beta E + \beta\tilde{\epsilon}L\right] \quad (8)$$

where irrelevant constant solvent terms have been omitted as we only need the first derivative of the free energy with respect to $N(L)$, which now takes the form

$$\frac{\delta\beta F}{\delta N(L)} = \ln N(L) + \ln C - \ln \Omega_L + \beta E + \beta\tilde{\epsilon}L + 1 \; . \tag{9}$$

With this expression, the minimisation condition on $N(L)$ becomes

$$\ln N(L) + \ln C - \ln \Omega_L + \beta E + 1 + \mu' L = 0 \tag{10}$$

where $\mu' = (\beta\mu + \beta\tilde{\epsilon})$. We note at this stage that the second derivative of $\beta F(V,T;\ \{N(L)\}) + \mu' \sum_L LN(L)$ with respect to $N(L)$ and $N(L')$ variables gives the non negative result $\frac{\delta_{LL'}}{N(L)}$, indicating that the extremum is indeed a minimum.

Solving for $N(L)$ in(10), we get

$$N(L) = {C'}^{-1} \exp -(\mu' L + \beta E - \ln \Omega_L) \tag{11}$$

where $C' = eC$ while, according to (2)), μ' must be such that

$$\sum_L L \exp -(\mu' L + \beta E - \ln \Omega_L) = MC' \tag{12}$$

The equilibrium $N(L)$ variables are also related to the equilibrium chain length average L_0 (see (1)), so that

$$\sum_L \exp -(\mu' L + \beta E - \ln \Omega_L) = \frac{MC'}{L_0} \tag{13}$$

To progress, we now need to specify the explicit L dependence of Ω_L. The traditional single chain theories of polymer physics provide universal expression of Ω_L in terms of the polymer size, the environment being simply taken into account through the solvent quality and the swollen blob size in the semi-dilute (good solvent) case.

The Case of Mean Field or Ideal Chains

The basic mean-field or ideal chain model for a L segments chain gives

$$\Omega_L^{id} = \left[C_1 z^L\right] \tag{14}$$

where z is the single monomer partition function and C_1 a dimensionless constant. Adapting(11), one has

$$N(L) = \frac{C_1}{C'} \exp -(\beta E) \exp\left(-\mu" L\right) \tag{15}$$

where $\mu" = \mu' - \ln z$ must, according to (12), be such that

$$\sum_L L \exp -(\mu" L) = \frac{1}{\mu"^2} = \frac{MC'}{C_1 \exp -(\beta E)} \tag{16}$$

while (13) takes the form

$$\sum_L \exp -(\mu" L) = \frac{1}{\mu"} = \frac{MC'}{L_0 C_1 \exp -(\beta E)} \tag{17}$$

In (16) and (17), sums over L from 1 to ∞ have been approximated by the result of their continuous integral counterparts.

Combining (15), (16) and (17), one gets the final expression for the chain number densities

$$c_0(L) = \frac{\phi}{L_0^2} \exp\left(-\frac{L}{L_0}\right) \tag{18}$$

with the average polymer length given by

$$L_0 = B^{1/2} \phi^{\frac{1}{2}} \exp\left(\frac{\beta E}{2}\right). \tag{19}$$

where $B = eb^3/C_1$ is a constant depending upon the monomer size b and the prefactor in the number of ideal chain configurations in (14).

The Case of Dilute Chains in Good Solvent

Polymer solutions in good solvent are in a dilute regime when chains do not overlap and in semi-dilute regime when chains do strongly overlap while the total monomer volume fraction is still well below its melt value. In the semi-dilute regime, chains remain swollen locally over some correlation length, known as the swollen blob size χ, but they are ideal over larger distances as a result of the screening of excluded volume interactions between blobs.

Specifically, for a given monomer number density ϕ, the blob size is given by the condition that the blob volume times ϕ must be equal to the number of monomers L^* in the swollen blob. This gives in terms of the reduced number density $\phi' = b^3\phi$

$$L^* \sim \phi'^{\left(\frac{1}{1-3\nu}\right)} \tag{20}$$

$$\chi \sim b\phi'^{\left(\frac{\nu}{1-3\nu}\right)} \tag{21}$$

where $\nu = 0.588$ in present good solvent conditions [8].

In living polymers characterized by a monomer number density ϕ and some averaged chain length L_0, the semi-dilute conditons correspond to the case $L_0 \gg L^*$. We discuss in this subsection the theory for the dilute case where $L_0 \ll L^*$. We will come back to the semi-dilute case in the next subsection.

Self-avoiding walks statistics apply to dilute chains in good solvent, and we thus adopt the number of configurations [6, 7](See especially page 128 of the book of Grosberg and Khokhlov [7])

$$\Omega_L^{EV} = C_1 L^{(\gamma-1)} z^L \tag{22}$$

where γ is the (entropy related) universal exponent equal to 1.165.

Incorporating expression (22) in (11), one gets

$$N(L) = \frac{V}{B} \exp -(\beta E) L^{(\gamma-1)} \exp -(\mu" L) \tag{23}$$

where B was introduced in(19) and where $\mu" = \mu' - \ln z$ must be fixed by(2)

$$\sum_L L^\gamma \exp -(\mu" L) = B\phi \exp (\beta E) \tag{24}$$

while (13) takes here the form

$$\sum_L L^{(\gamma-1)} \exp -(\mu" L) = \frac{B\phi}{L_0} \exp (\beta E) \tag{25}$$

If L is treated as a continuous variable, (24) and (25) can be rewritten in terms of the Euler Gamma function satisfying $\Gamma(x) = x\Gamma(x-1)$ as

$$\int_0^\infty L^\gamma \exp -(\mu" L) dL = \frac{\Gamma(\gamma+1)}{\mu"^{(\gamma+1)}} = B\phi \exp (\beta E) \tag{26}$$

$$\int_0^\infty L^{(\gamma-1)} \exp -(\mu" L) dL = \frac{\Gamma(\gamma)}{\mu"^{\gamma}} = \frac{B\phi}{L_0} \exp (\beta E) \tag{27}$$

From (26) and (27), one gets

$$\mu" = \frac{\gamma}{L_0} \tag{28}$$

$$B\phi \exp (\beta E) = \frac{\Gamma(\gamma+1)}{\gamma^{(\gamma+1)}} L_0^{(\gamma+1)} \tag{29}$$

These results lead then finally to the Schulz-Zimm distribution of chain lengths, namely

$$c_0(L) = \frac{\exp (-\beta E)}{B} L^{(\gamma-1)} \exp \left(-\gamma \frac{L}{L_0} \right) \tag{30}$$

and an average polymer length given by

$$L_0 = \left(\frac{\gamma^\gamma}{\Gamma(\gamma)} \right)^{\frac{1}{1+\gamma}} B^{\frac{1}{1+\gamma}} \phi^{\frac{1}{1+\gamma}} \exp \left(\frac{\beta E}{(1+\gamma)} \right) . \tag{31}$$

The Semi-Dilute Case

We consider here the semi-dilute case in good solvent where the average length of living polymers L_0 is much larger than the blob length L^*. The usual picture of a semi-dilute polymer solution is an assembly of ideal chains made of blobs of size χ. Using this approach, Cates and Candau [2] and later J.P. Wittmer et al [13] derived the relevant equilibrium polymer size distribution. In this subsection, we adapt their derivation to the theoretical framework presented above.

Let Ω_b be the number of internal configurations per blob and z' some coordination number for successive blobs. As there are $n_b = L/L^*$ blobs for a chain of L monomers, we write the total number of internal configurations of a chain of size L as

$$\Omega_L^{SD} = C_1 L^{*(\gamma-1)} \Omega_b^{L/L^*} z'^{L/L^*} \tag{32}$$

where γ is the universal exponent in the excluded volume chain statistics met earlier for chains in dilute solutions. The important factor $L^{*(\gamma-1)}$ can be seen as an entropy correction for chain ends just like E was an energy correction to $L\tilde{\epsilon}$. This entropic term which involves the number of monomers per blob, is needed to take into account that when a chain breaks, its two ends are subject to a reduced excluded volume repulsion. The other factors in (32) will lead to terms linear in L after taking the logarithm and thus will be absorbed in the Lagrange multiplier definition, as seen earlier in similar cases for ideal and dilute chains.The resulting expression of $N(L)$ in terms of the Lagrange multiplier (cfr (15)) can then be written by analogy as

$$N(L) = \frac{C_1}{C'} \exp - [\beta E - (\gamma - 1) \ln L^*] \exp(-\mu L) \tag{33}$$

Proceeding as in the ideal case (simply replacing at every step the constant βE by $\beta E - (\gamma - 1) \ln L^*$, one recovers in the semi-dilute case the simple exponential distribution

$$c_0(L) = \frac{\phi}{L_0^2} \exp\left(-\frac{L}{L_0}\right) \tag{34}$$

with a slightly different formula for the average polymer length

$$L_0 \propto \phi^\alpha \exp\left(\frac{\beta E}{2}\right) \tag{35}$$

where $\alpha = \frac{1}{2}(1 + \frac{\gamma-1}{3\nu-1})$ is about 0.6.

2.2 A Kinetic Model for Scissions and Recombinations

The interest for wormlike micelles dynamics came from the experimental observation that entangled flexible supramolecular polymers display, after an

initial strain, a simple exponential stress relaxation which is qualitatively different from the behaviour of usual entangled melts. In the latter system, individual chains must leave by a reptation mechanism the strained topological tube created by the entangled temporarily network, in order to relax the shearing forces. A plausible scission-recombination model, allowing for such an extra relaxation mechanism, was shown [2] to lead to an exponential decay of the shearing forces with a decay Maxwell time equal to $\tau \approx \sqrt{\tau_b \tau_{rep}}$ where τ_{rep} is the chain reptation time and τ_b is the mean life time of a chain of average size in the system.

We will assume in the following that the Cate's scission-recombination model governing the population dynamics, originally deviced to explain entangled equilibrium polymer melt rheology, should also apply to the kinetically unentangled regime which is explored in the present work.

This kinetic model [9] assumes that

- the scission of a chain is a unimolecular process, which occurs with equal probability per unit time and per unit length on all chains. The rate of this reaction is a constant k_s for each chemical bond, giving

$$\tau_b = \frac{1}{k_s L_0} \tag{36}$$

  for the lifetime of a chain of mean length L_0 before it breaks into two pieces.
- recombination is a bimolecular process, with a rate k_r which is identical for all chain ends, independent of the molecular weight of the two reacting species they belong to. It is assumed that recombination takes place with a new partner respect to its previous dissociation as chain end spatial correlations are neglected within the present mean field theory approach. It results from detailed balance that the mean life time of a chain end is also equal to τ_b.

Let $c(t, L)$ be the number of chains per unit volume having a size L at time t. Cates [9] writes the kinetic equations as

$$\frac{dc(t,L)}{dt} = -k_s L c(t,L) + 2k_s \int_L^\infty c(t,L')dL' \tag{37}$$

$$+ \frac{k_r}{2} \int_0^L c(t,L')c(t,L-L')dL' - k_r c(t,L) \int_0^\infty c(t,L')dL' \tag{38}$$

where the two first terms deal with chain scission (respectively disappearence or appearence of chains with length L) while the two latter terms deal with chain recombination (respectively provoking the appearence or disappearence of chains of length L).

It is remarkable that the solution of this empirical kinetic model leads to an exponential distribution of chain lengths. Indeed, direct substitution of solution $c_0(L)$ in the above equation leads to the detailed balance condition:

$$\phi \frac{k_r}{2k_s} = L_0^2 \tag{39}$$

the ratio of the two kinetic constant being thus restricted by the thermodynamic state. Detailed balance means that for the equilibrium distribution $c_0(L)$, the number of scissions is equal to the number of recombinations. The total number of scissions and recombinations per unit volume and per unit time, denoted respectively as n_s and n_r, can be expressed as

$$n_s = k_s \frac{\phi}{L_0^2} \int_0^\infty L \exp(-L/L_0) dL = k_s \phi \tag{40}$$

$$n_r = \frac{k_r}{2} \frac{\phi^2}{L_0^4} \int_0^\infty dL' \int_0^\infty dL" \exp(-\frac{L'}{L_0}) \exp(-\frac{L"}{L_0}) = \frac{\phi^2 k_r}{2L_0^2} \tag{41}$$

and it can be easily verified that detailed balance condition implies $n_s = n_r$.

Mean field theory assumes that a polymer of length L will break on average after a time equal to $\tau_b = (k_s L)^{-1}$ through a Poisson process. This implies that the distribution of first breaking times (equal to the survival times distribution) must be of the form

$$\Psi(t) = \exp\left(-\frac{t}{\tau_b}\right) \tag{42}$$

for a chain of average size. Detailed balance then requires that the same distribution represents the distribution of first recombination for a chain end [2]. Accordingly, throughout the rest of this chapter, the symbol τ_b will represent as well the average time to break a polymer of average size or the average time between end chain recombinations. In the same spirit, we stress that among the different estimates of τ_b proposed in this work, some are based on analyzing the scission statistics while others are based on the recombination statistics.

Two additional points may be stressed at this stage:

- The mean field model in the present context has been questionned [12] because in many applications, there are indications that a newly created chain end often recombines after a short diffusive walk with its original partner. In that case, a possibly large number of breaking events are just not effective and the kinetics proceeds thus differently.
- Given the statistical mechanics analysis in the previous subsection, we see that the equilibrium distribution of chain lengths resulting from the simple empirical kinetic model is perfectly compatible with the equilibrium distribution in polymer solutions at the θ point (ideal chains) or for semi-solutions (ideal chains of swollen blobs). The kinetic model is still pertinent, given its simplicity, in the case of dilute solutions as the static chain length distribution based on statistical mechanics, although non exponential, is not very far from it.

3 The Mesoscopic Model and Its Brownian Dynamics Implementation

In order to test the theoretical understanding summarized in the previous section, we have devised a new microscopic model bearing resemblances with an off-lattice Monte-Carlo model proposed earlier [21], but which uses Brownian Dynamics to follow the monomer dynamics for a standard polymer chain model [20]. To be able to clearly unravel static and dynamic aspects, our model contains a kinetic parameter which allows to vary the rates of the reactions without affecting the static and thermodynamic properties. As already mentioned in the introduction, two thermodynamic states corresponding to equilibrium polymers dissolved in a good solvent are considered, namely a dilute and a semi-dilute solution.

We consider a set of micelles consisting of (non-cyclic) linear assemblies of Brownian particules. Within such a linear assembly, the bounding potential $U_1(r)$ acting between adjacent particles is expressed as the sum of a repulsive Lennard-Jones (shifted and truncated at its minimum) and an attractive part of the FENE type [20]. The pair potential $U_2(r)$ governing the interactions between any unbounded pair (both intramicellar and intermicellar) is a pure repulsive potential corresponding to a simple Lennard-Jones potential shifted and truncated at its minimum. This choice of effective interaction between monomers implies good solvent conditions.

Using the Heaviside function $\Theta(x) = 0$ or 1 for $x < 0$ or $x \geq 0$ respectively, explicit expressions (see Fig. 2) are

$$U_2(r) = 4\epsilon \left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6} + \frac{1}{4} \right] \Theta(2^{1/6}\sigma - r)) \tag{43}$$

$$U_1(r) = U_2(r) - 0.5kR^2 \ln\left(1 - \left[\frac{r}{R}\right]^2\right) - U_{min} - E \tag{44}$$

In the second expression valid for $r < R$, $k = 30\epsilon/\sigma^2$ is the spring constant and $R = 1.5\sigma$ is the value at which the FENE potential diverges. U_{min} is the minimum value of the sum of the two first terms of the second expression (occuring $r_{min} = 0.96094\sigma$ for the adopted parameters) while E is a key parameter which corresponds, when pair potential exchanges between bounded and unbounded situations will be allowed, to the typical energy gain (loss) when an unbounded (bounded) pair is the object of a recombination (scission).

We deal with M monomers at temperature T enclosed in a volume V, defining a monomer number concentration $\phi = M/V$. The relevant canonical partition function is a sum over all microscopic states specified not only by monomer positions (and momenta) but also by a set of variables defining unambiguously the connectivity scheme for the particular microscopic state. Disregarding the irrelevant momentum variables in Brownian Dynamics, we write the partition function as

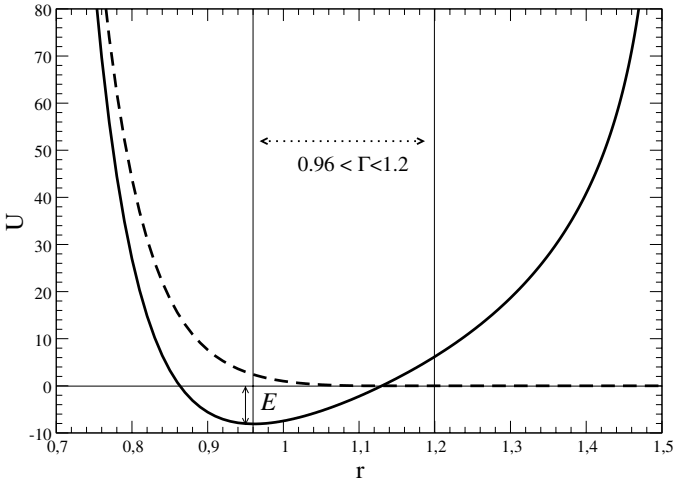


Fig. 2. Bounded potential $U_1(r)$ (*continuous curve*) and unbounded potential $U_2(r)$ (*dashed line*) between a pair of monomers. Each monomer can participate to at most two bounded links (with potential $U_1(r)$). All unbounded monomer pairs interact via the potential $U_2(r)$. E is a parameter tuning the energy required to open the bond. The figure also shows the Γ region where potential swaps (equivalent to bond scissions or bond recombinations) are allowed during the brownian dynamics simulation

$$Q_{MVT} \sim \sum_{\{r\}} \sum_{m} \exp\left(-\frac{U(\{r\},m)}{k_B T}\right) \tag{45}$$

where the double sum runs over all possible arrangements of M distinguishable monomers and over all related bounding schemes. The first sum runs over all spatial positions $\{r\}$ of the M monomers, each configuration being characterized by a list of K monomer-monomer pairs with relative distance shorter than R. The second sum runs, for a given spatial arrangement with a subset of K pairs available for bounding, over all allowed distinct bounding schemes, namely those among the 2^K distinct possibilities which satisfy two restricting rules, namely

- no monomer can be engaged into more than two bounding pairs (= no branching).
- no cyclic bounding structure is allowed.

The potential energy associated to a particular configuration can be written

$$U(\{r\},m) = \frac{1}{2}\sum_{i}\sum_{j\neq i}\left(s_{ij}U_1(r_{ij}) + (1-s_{ij})U_2(r_{ij})\right) \tag{46}$$

where $s_{ij} = 0$ for all pairs with $r_{ij} > R$ while $s_{ij} =$ either 0 or 1 for the K pairs with $r_{ij} < R$. Index m in the two previous expressions corresponds to a

particular bounding network representing a particular set of 0 and 1 for the K s_{ij} values.

In order to discuss topological changes (change in bounding network) during the phase space exploration, it is useful to consider that each monomer possesses two arms available for bounding. If both arms are free, we have an isolated monomer with degree of polymerisation $L = 1$. If at least one arm of a particular monomer is linked through potential $U_1(r)$ to an arm of another monomer, these two monomers belong to the same linear structure (micelle) of degree of polymerisation $L \geq 2$. A terminal or an interior monomer will correspond to a monomer engaged respectively in one or two bonds with other monomers.

The phase space exploration will be performed by a Brownian Dynamics (BD) scheme (controling r positions updates) coupled to a single bond creation (replacement of potential $U_2(r)$ by $U_1(r)$)/single bond annihilation (replacement of potential $U_1(r)$ by $U_2(r)$) scheme satifying microscopic reversibility.

The BD is standard and samples the set of positions of the M monomers according to a canonical ensemble corresponding to a particular potential energy (equivalent in the present case to a particular bonding network m).

The single BD step of particule i subject to a total force $\overline{F}_i$ is simply

$$\overline{r}_i(t + \Delta t) = \overline{r}_i(t) + \frac{\overline{F}_i}{\xi}\Delta t + \overline{R}_i(\Delta t) \tag{47}$$

where the last term corresponds to a vectorial random gaussian quantity with first and second moments given by

$$\langle R_{i\alpha}(\Delta t)\rangle = 0 \tag{48}$$

$$\langle R_{i\alpha}(\Delta t) R_{j\beta}(\Delta t)\rangle = 2\left(\frac{k_B T}{\xi}\right)\Delta t \delta_{\alpha\beta}\delta_{ij} \tag{49}$$

for arbitrary particles i and j and where $\alpha\beta$ stand for the x, y, z cartesian components.

At this stage, it is useful to fix units. In the following, we will adopt the size of the monomer σ as unit of length, the ϵ parameter as energy unit and we will adopt $\xi\sigma^2/(3\pi\epsilon)$ as time unit. We also introduce the reduced temperature $k_B T/\epsilon = T^*$, so that in reduced coordinates (written with symbol*) the algorithm becomes

$$\langle R^*_{i\alpha}(\Delta t^*)\rangle = 0 \tag{50}$$

$$\langle R^*_{i\alpha}(\Delta t^*) R^*_{j\beta}(\Delta t^*)\rangle = \frac{2}{3\pi}\Delta t^* \delta_{\alpha\beta}\delta_{ij}\, T^* \tag{51}$$

which means that each particle, if isolated in the solvent, would diffuse with a RMSD of $\sqrt{2T^*/\pi}$ per unit of time.

The bonding network is itself the object of random instantaneous changes provided by a Monte-Carlo algorithm which is built according to the standard Metropolis scheme.

The probability $P_{m,n}$ to go from a bounding network m to a different one n is written as

$$P_{m,n} = A^{trial}_{m,n} P^{acc}_{m,n} \tag{52}$$

where $A^{trial}_{m,n}$ is the trial probability to reach a new bounding network n starting from the old one m, within a single MC step. This trial probability is chosen here to be symmetric as usually adopted in Metropolis Monte Carlo schemes. This trial probability is chosen to be different from zero only if both bounding networks n and m differ by the status of a single bond, say the pair of monomers (ij). To satisfy microreversibilty, the acceptance probability $P^{acc}_{m,n}$ for the trial $(m \longrightarrow n)$ must be given by

$$P^{acc}_{m,n} = Min\left[1, \exp\left(-\frac{(U(\{r\},n) - U(\{r\},m))}{k_B T}\right)\right] . \tag{53}$$

In the present case, as a single pair (ij) changes its status, the acceptance probability takes the explicit form

$$P^{acc}_{m,n} = Min\left[1, \exp\left(-\frac{(U_2(r_{ij}) - U_1(r_{ij}))}{k_B T}\right)\right] \tag{54}$$

$$P^{acc}_{m,n} = Min\left[1, \exp\left(-\frac{(U_1(r_{ij}) - U_2(r_{ij}))}{k_B T}\right)\right] \tag{55}$$

for bond scission and bond recombination respectively.

The way to specify the trial matrix $A^{trial}_{m,n}$ starts by defining a range of distances, called Γ and defined by $0.96 < r < 1.20$ within the range $r < R$. For $r \in \Gamma$, a change of bounding is allowed as long as the two restricting rules stated above are respected. Consider the particular configuration illustrated by Fig. 3 where $M = 7$ monomers located at the shown positions, are characterized by a connecting scheme made explicit by representing a bounding potential by a continuous line. The dashed line between monomers 5 and 6 represents the changing pair with distance $r \in \Gamma$ which is a bond $U_1(r)$ in configuration m but is just an ordinary intermolecular pair $U_2(r)$ in configuration n.

For further purposes, we have also indicated in Fig. 3 by a dotted line all pairs with $r \in \Gamma$ which are potentially able to undergo a change from a non bounded state to a bounded one in the case where the (56) pair, on which we focus, is non bounded (state n). Note that bond (35) is not represented by a dotted line eventhough the distance is within the Γ range: a bond formation in that case is not allowed as it would lead to a cyclic conformation. Note also that in state n, monomer 5 could thus form bonds either with monomer 2 or with monomer 6 while monomer 6 can only form a bond with monomer 5.

We now state the algorithm and come back later on the special $(m \longrightarrow n)$ transition illustrated in Fig. 3.

During the BD dynamics, with an attempt frequency ω per arm and per unit of time, a change of the chosen arm status (bounded $U_1(r)$ to unbounded

$U_2(r)$ or unbounded to bounded) is tried. If it is accepted as a "trial move" of the bounding network, it obviously implies the modification of the status of a paired arm belonging to another monomer situated at a distance $r \in \Gamma$ from the monomer chosen in the first place.

The trial move goes as follows: a particular arm is chosen, say arm 1 of monomer i, and one first checks whether this arm is engaged in a bounding pair or not.

- If the selected arm is bounded to another arm (say arm 2 of monomer j) and the distance between the two monomers lies within the interval $r_{ij} \in \Gamma$, an opening is attempted with a probability $1/(N_i + 1)$, where the integer N_i represents the number of monomers available for bonding with monomer i, besides the monomer j (N_i is thus the number of monomers with at least one free arm whose distance to the monomer carrying the originally selected arm i lies within the interval Γ, excluding from counting the monomer j and any particular arm leading to a ring closure). If the trial change consisting in opening the (ij) pair is refused (either because the distance is not within the Γ range or because the opening attempt has failed in the case $N_i > 0$), the MC step is stopped without bonding network change (This implies that the BD restarts with the (ij) pair being bonded as before).
- If the selected arm (again arm 1 of monomer i) is free from bonding, a search is made to detect all monomers with at least one arm free which lie in the "reactive" distance range $r \in \Gamma$ from the selected monomer i (Note that if monomer i is a terminal monomer of a chain, one needs to eliminate from the list if needed, the other terminal monomer of the same chain in order to avoid cyclic micelles configurations). Among the monomers of this "reactive" neighbour list, one monomer is then selected at random with equal probability to provide an explicit trial bonding attempt between monomer i and the particular monomer chosen from the list. Note that if the list is empty, it means that the trial attempt to create a new bond involving arm 1 of monomer i has failed and no change in the bonding network will take place.

In both cases, if a trial change is proposed, it will be accepted with the probability $P^{acc}_{m,n}$ defined earlier. If the change is accepted, BD will be pursued with the new bonding scheme (state n) while if the trial move is finally rejected, BD restarts with the original bonding scheme corresponfing to state m.

Coming back to Fig. 3, we now show that the MC algorithm mentioned above garantees that the matrix A^{trial}_{mn} is symmetric, an important issue as it leads to the micro-reversibility property when combined with the acceptance probabilities described earlier. Let us define as $P_{arm} = 1/2N$ the probability to select a particular arm, a uniform quantity.

If configuration m with pair (56) being "bounded" is taken as the starting configuration, the number of available arms to form alternative bonds with

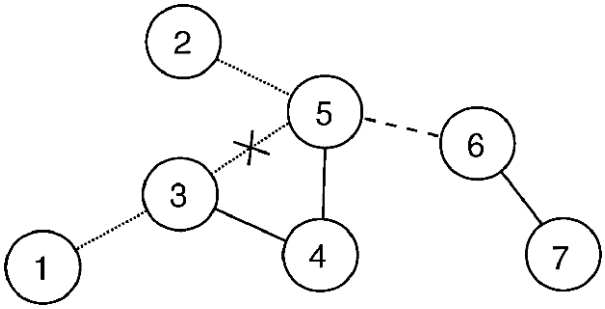


Fig. 3. Exemplary configuration of a 7 monomers system in state n where monomers 3,4 and 5 form a trimer and monomers 6 and 7 a dimer. All pairs of monomers with mutual distances within the Γ region are indicated by a *dotted* or a *dashed line*. In the text, we consider the Monte Carlo scheme for transitions between states n and m which only differ by the fact that in state n and m the 5-6 pair is respectively open or bounded. The n $\rightarrow$ m transition corresponds to the creation of a pentamer by connecting a dimer and a trimer while the m $\rightarrow$ n transition leads to the opposite scission. The cross symbol on link 3-5 indicates that in state n, when looking to all monomers which could form a new link with monomer 5, monomer 3 is excluded because it would lead to a cyclic polymer which is not allowed within the present model

monomer 5 and monomer 6 are respectively $N_5 = 1$ and $N_6 = 0$. Therefore, applying the MC rules described above, the probability to get configuration n where the pair (56) has to be unbounded is given by the sum of probabilities to arrive at this situation through selection of the arm of monomer 5 engaged in the bond with monomer 6 or through selection of the arm of monomer 6 engaged in a bond with monomer 5. This gives

$$A_{m,n}^{trial} = P_{arm} * \frac{1}{N_5 + 1} + P_{arm} \frac{1}{N_6 + 1} = \frac{3}{2} * P_{arm} \tag{56}$$

If configuration n with pair (56) being "unbounded" is taken as the starting configuration, the application of the MC rules lead to the probability to get configuration m where the pair (56) has to be bounded is given by the sum of probabilities to arrive at this situation through selection of the free arm of monomer 5 (which has two bounding possibilities, namely with monomers 2 and 6) or through selection of the free arm of monomer 6 which can only form a bond with monomer 5. This gives

$$A_{n,m}^{trial} = P_{arm} * \frac{1}{2} + P_{arm} = \frac{3}{2} * P_{arm} \tag{57}$$

showing the required matrix symmetry.

4 Characterization of Static Properties

In the following, all quantities will systematically be expressed in reduced units without an explicit $*$ symbol.

4.1 List of Experiments and Chain Length Distributions

We consider a system of $M = 1000$ monomers interacting via the potential model defined by (44) at $T = 1$. This is a rather small system by nowadays standards but we were only interested here to a preliminary demonstrative study before going in the future to larger systems and a more efficient description of solvent effects.

This system is studied at two state points:

- A dilute solution at the number density $\phi = 0.05$ and an energy parameter $E = 8$.
- A semi-dilute solution at the number density $\phi = 0.15$ and an energy parameter $E = 10$.

Each system evolves according to the Brownian Dynamics algorithm with time step $\Delta t = 0.001$ and is subject to random trials of bond scission/recombination with arm trial attempts frequencies in the range $0.1 < \omega < 5$. This choice was motivated by the necessity to speed up the kinetics and hence, getting reasonably good statistics on dynamical properties. Obviously, all static properties are independent of ω and therefore, all data can be cumulated over all ω values.

The systems have been equilibrated for a time of $2\ 10^4$. Production runs have been followed for a similar time $T_{run} = 2\ 10^4$. The equilibrium values of the mean chain length L_0 and the total number of chains $\langle N_{ch} \rangle$ are related by $M = L_0 \langle N_{ch} \rangle$.

The Dilute Case

The first experiment deals with a dilute solution as the average chain length $L_0 = 10.4 \pm 0.1$ is much smaller than the crossover value at that monomer number density as calculated by (21), $L^* = 50.5$.

The populations of bonds ready to open within the Γ range is $\langle N_{1\Gamma} \rangle = 551.6$ and represents 61% of the bonded pairs $\langle N_1 \rangle$, while the population of free arms pairs ready to close, again within the same Γ range, is only $\langle N_{2\Gamma} \rangle = 1.16$. Figures 4 and 5 show respectively the pair correlation function for chain ends (unsaturated arms) and the distribution of the distance between bounded monomers, in particular within the Γ region where potential changes do occur. The "free arm" fraction, namely $1 - N_1/M = 0.0957$, is close to $L_0^{-1} = \langle N_{ch}/M \rangle = 0.0964$.

Dilute solution conditions are confirmed by the chain length distribution shown in Fig. 6. For dilute conditions, a distribution given by (30) is expected.

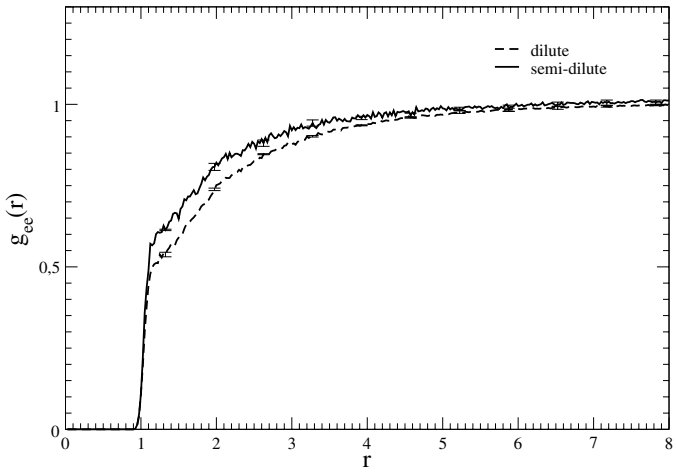


Fig. 4. Chain end pair distribution function $g_{ee}(r)$ in the dilute and semi-dilute experiments. Note that the Γ region where the bounding changes take place correspond to the region of first (fast) increase of the distribution function around $r = \sigma$

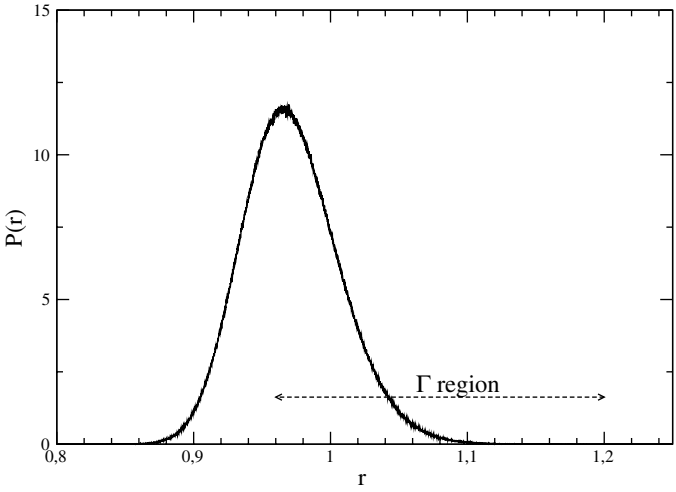


Fig. 5. Normalized distribution function $P(r)$ of the distance between bounded monomers in the semi-dilute experiment. The corresponding function in the dilute solution experiment is marginally different from the semi-dilute case and is therefore not shown in the figure

Accordingly, we have fitted our data with a single parameter (prefactor) fit function, $c(L) = A_0 L^{\gamma-1} \exp(-\gamma L/L_0)$ where L_0 was given its computed average value and where γ was given its expected value, $\gamma = 1.165$. This curve is significantly better than the simple exponential distribution expected

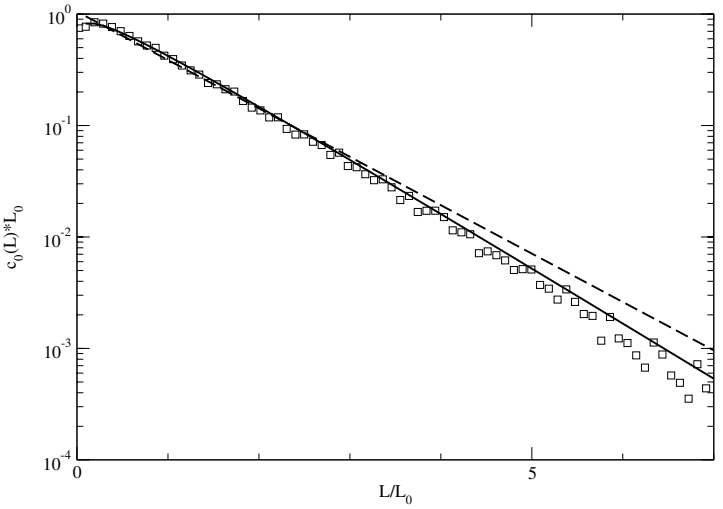


Fig. 6. Distribution of chain length number densities for the dilute case in good solvent. The data (*squares*) are fitted using (30) with imposed value $\gamma = 1.165$ (continuous curve). The *dashed line* shows the exponential distribution $C_0(L) \propto \exp(-\frac{L}{L_0})$ which does not fit data as well

for ideal or semi dilute chains. If γ is left as a second free parameter in the fit, it takes an even larger value 1.187 and the fit (not shown) is slightly better especially at high L.

The Semi-Dilute Case

In the second state point experiments, the average chain length is found to be $\langle L \rangle = 49.9 \pm 0.1$, a value which is much larger than the crossover value $L^* \approx 12$ at $\phi = 0.15$ according to (20).

The populations of bonds ready to open in the Γ region is $\langle N_{1\Gamma} \rangle = 597 \pm 1$, while the population of free arms pairs ready to close is $\langle N_{2\Gamma} \rangle = 0.183$. Distance distributions of unbounded and bounded pairs are given in Figs. 4 and 5 together with the dilute case data. We note that $g(r)$ function between free ends is not very different between the two cases. The "free arm" fraction $1 - N_1/M = 0.021$ is 5 times smaller than in the dilute case as the chains are 5 times longer.

Semi-dilute solution conditions are confirmed by the observation of an exponential distribution of the reduced chain length (see Fig. 7). In the latter figure, we show the best fit by a one parameter fitting function $c^{fit}(L) = A \exp(-L/L_0)$.

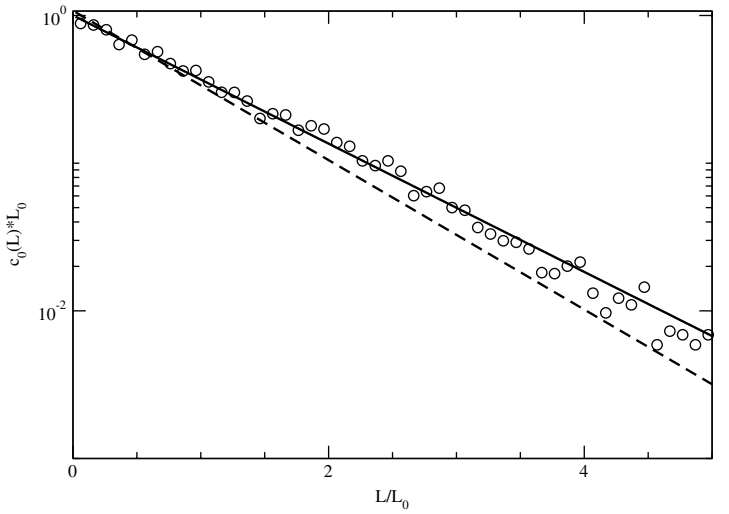


Fig. 7. Distribution of chain lengths number densities for the semi-dilute case. The data (*circles*) are fitted with a simple exponential function $A \exp(-L/L_0)$ (*continuous curve*), in reference to (34) with imposed average chain size $L_0 = 49.9$. *Dashed lines* show a function $\exp(-1.165 \frac{L}{L_0})$ which does not fit the data as well

4.2 Chain Length Conformational Analysis

Figures 8 and 9 show for dilute and semi-dilute state points respectively, the average of the square of the radius of gyration and the average of the square of the end-to-end vector as a function of the chain length within the polydisperse sample.

In Fig. 8 relative to the dilute case, standard power law scaling $L^{2\nu}$ law with $\nu = 0.588$ is indicated for long chains. For our polydisperse system with an average chain length of $L_0 \approx 10$, our data on chains with $L > L_0$ are affected by too large statistical error bars to test critically the validity of this asymptotic regime and its exact range of application, limited in the low L regime by usual short chain effects clearly visible below $L = 25$.

The conformational properties of chains in the semi-dilute system are compatible with ideal chains for $L > L^* \approx 12$ where L^* is estimated from (20) with $\phi = 0.15$, and with a larger exponent power law for lower L values but poor statistics and small chain effects prevent a firm test of scaling laws for these semi-dilute chains.

When $\langle R^2 \rangle$ data from both state points are plotted versus L on a unique graph (not shown), the short chains behaviours are found to be undistinguishable, while divergences between the two state points are observed for long chains. The power laws shown in Figs. 8 and 9 intersect at the blob length value of $L^* = 37$ monomers, giving a blob size or a correlation length of $\chi = \langle R^2 \rangle^{1/2} = 20$. Note that in this analysis, we implictly assume that

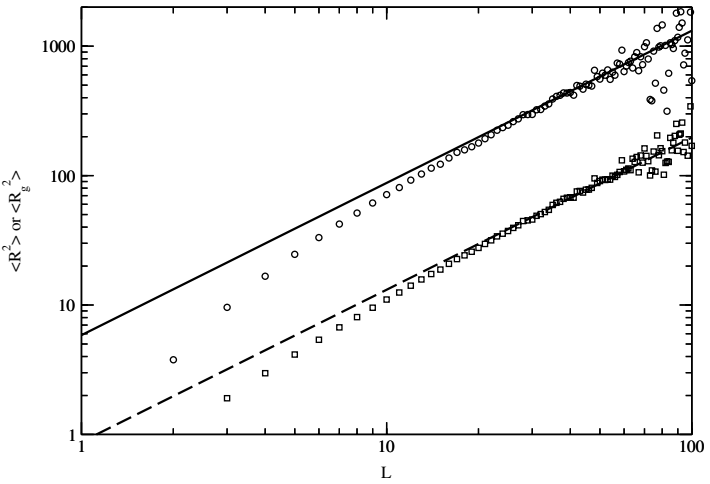


Fig. 8. Conformational properties in the dilute case: $\langle R^2 \rangle$ (*circles*) and $\langle R_g^2 \rangle$ (*squares*) versus chain length L. The fitting functions which are power law with exponent fixed to $2\nu = 1.176$ (prefactors 0.876 for $\langle R_g^2 \rangle$ and 5.854 for $\langle R^2 \rangle$) fit the data for longer chains only

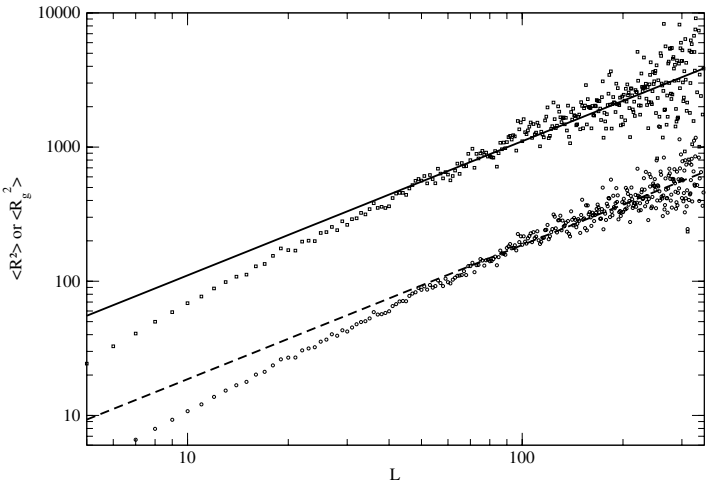


Fig. 9. Conformational properties in the semi-dilute case: $\langle R^2 \rangle$ (*squares*) and $\langle R_g^2 \rangle$ (*circles*) versus chain length L. The fitting functions (linear in L) assume ideal chain statistics with prefactors 1.862 for $\langle R_g^2 \rangle$ and 11.06 for $\langle R^2 \rangle$. The fit works for the $L > L^*$ behaviour

the prefactor of the dilute chains power law is independent of ϕ while we consider the ϕ dependent prefactor in the ideal chain power law corresponding to the semi-dilute state at $\phi = 0.15$. We thus see that the blob size as determined from the intersection of asymptotic power laws (neglecting short

chains effects) suggest a numerical prefactor of about 3 in the scaling relations (20). On this basis, our dilute system at $\phi = 0.05$ would be characterized by a blob length of $L^* = 50$ and thus a ratio $L_0/L^* = 0.2$ suggesting a rather clear dilute solution character. The same ratio is $L_0/L^* = 1.4$ for the more concentrated solution showing that the usual concept of chains of blobs in a semi-dilute solution applies only marginally to the longer chains in the sample. To conclude, this analysis indicates that, even in the semi-dilute case, the chain dynamics should not be significantly affected by the weak entanglements between chains. Let us stress that the semi-dilute character of the sample is however well illustrated by static properties such as the exponential chain length distribution and the ideal chain power laws for chain sizes.

5 Kinetics Analysis

The number of Monte-Carlo "accepted" scissions (or "accepted" recombinations) are obtained by simple counting during the simulations. In Table 1 (dilute case) and in Table 2 (semi-dilute case), we list the number n_s of "accepted" transitions per unit time and unit volume, for the two state points and for the different attempt frequencies ω investigated. For each state point, we observe that the number of scissions per unit of time is, as expected, proportional to ω and we have also verified that the number of recombinations differs from the number of scissions by marginal amounts (0.05%), which shows that the chain length distributions are well equilibrated.

The quantity α mentioned in the tables is the fraction of scissions which lead to a recombination with a new partner. In our analysis, the fate of a scission has been initially followed for a maximum time of $T_{max} = 1000$. In our analysis, we suppose that chain ends, not get recombined at time $t = \tau_{max}$, would ultimately recombine with a new partner. The product αn_s is the number of scissions per unit of time and per unit volume which lead to a recombination with a new partner, that is those which lead to real changes in the distribution of chain lengths.

In the explored ω range, the value of α indicates that a large fraction of elementary scissions are followed by recombinations of the same original chain fragments. Using α as a measure of the percentage of effective reactions, we observe a decrease of this percentage as the rate of attempted transitions ω increases. Actually, uneffective transitions may also come from more complex particular transition sequences. Consider a chain end monomer (say monomer i) lying close in space to another chain at the level of two adjacent monomers j and k. At high ω, a scission of the bond jk may be immediately followed by a recombination between j or k with monomer i. In turn, that ik or ij bond may reopen and recombine to restaure the very first situation, ending with no effective transition without being detected through the criterium of a successive recombination with the same partner. The occurence of such

Table 1. Kinetic data for the dilute case at four scission-recombination attempt frequencies ω. n_s is the total number of Monte-Carlo scissions (recombinations) per unit time and per unit volume. α is the fraction of scissions where newly created free arms recombine with a new partner. Superscript (i) refers to the methodology used to get various estimates of a given quantity, namely $i = 1$ for the method based on the distribution of first recombination times, $i = 2$ for the cumulative hazard method and $i = 3$ for the "reactive flux" approach. τ_b is the average first recombination time and also the typical life time of a chain of average length L_0. The estimate $\tau_b^{(i)}$ is obtained by the long time behaviour of the relevant dynamical function used in methodology (i). $\kappa^{(i)}$ is the fraction of transitions which are effective. Such transitions can be seen as those which do not belong to sequences of correlated transitions (chain scission followed by almost immediate recombination ending into no change in chains topology). An additional estimate of τ_b (not listed in th etable) may be obtained on the basis of n_s and $\kappa^{(i)}$ computed with methodology i using $\tau_b = \frac{\phi}{L_0 \kappa^{(i)} n_s}$

ω	$\kappa^{(3)}$	α	$n_s * 10^3$	$\tau_b^{(1)}$	$\tau_b^{(2)}$	$\tau_b^{(3)}$
0.1	.89	0.89	0.00759	677.	763.	625.
0.5	0.60	0.68	0.038	215	212.	185.
1.0	0.45	0.55	0.076	138	146.	122.
5.0	0.20	0.28	0.38	69	67.0	63.3

Table 2. Kinetic data for the semi-dilute case at four scission-recombination attempt frequencies.All quantities are defined in the previous table

ω	$\kappa^{(2)}$	$\kappa^{(3)}$	α	$n_s * 10^3$	$\tau_b^{(1)}$	$\tau_b^{(2)}$	$\tau_b^{(3)}$
0.1	0.85	0.85	0.85	0.0042	823	885.	813.
0.5	0.52	0.44	0.53	0.022	263	265.	282.
1.0	0.34	0.35	0.40	0.042	191	190.	184
5.0	0.12	0.12	0.16	0.21	107	108.	106.

events is proven indirectly in the following by noting almost immediate chain end recombination with another partner (see Fig. 11).

To extract estimates of the rate constants defined by the kinetic model of Cates, it is thus important to eliminate the spurious transitions from the effective ones. The best route for this is certainly to approach the problem via a time scale separation between scission-recombination processes taking place on a fast time scale and effective transitions taking place on the reaction time scale τ_b.

5.1 Distribution of First Recombination Times

The first approach is to compute the histogram of first recombination times $t = t_2 - t_1$ for an arm which became free by scission at time t_1 and which recombined with another free arm at time t_2. Figure 10 shows the result for the semi-dilute case for the four values of ω. In Fig. 11 relative to the case $\omega = 0.5$

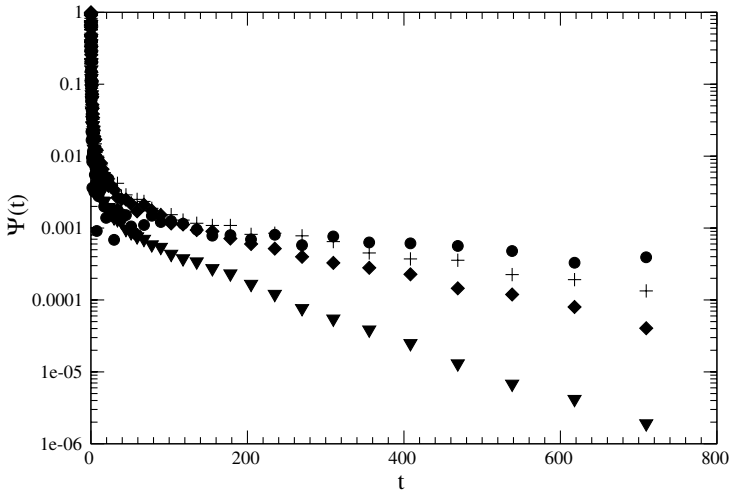


Fig. 10. Distribution of first recombination times $\Psi(t)$ for the semi-dilute case. Data are shown for $\omega = 0.1$ (*filled circles*), $\omega = 0.5$ (+ *symbols*), $\omega = 1.0$ (*filled lozenges*) and $\omega = 5.0$ (*filled triangles*). Estimates of $\tau_b^{(1)}$ are obtained from the slope of a linear fit of $\ln \Psi(t)$ versus time in the long time region

specifically, we show the contribution to the distribution of first recombination times from events implying either the same or a different partner. In the latter figure, one observes that at short time, the recombination with the same partner dominates. However, even at short times, recombinations with another partner are possible as discussed earlier. The two curves corresponding to the two types of contributions cross each other around $t = 25$ and at later times, the recombination with another partner progressively dominates. We note however that at times as long as $t = 200$ where monomers have diffused by a distance corresponding to three times their size (see Fig. 9), there is still 10% of recombinations at that time which take place with the same partner.

Data of Fig. 10 have been analyzed by considering that $\Psi(t)$ is exponential at long times. The slope is interpreted as the inverse of the average life time of a chain end which is denoted as $\tau_b^{(1)}$. Specific values for the different ω's are indicated in Tables 1 and 2, respectively for dilute and semi-dilute cases.

The short time behaviour of the distribution can be analyzed by plotting the function versus time on a log-log scale. This is done in Fig. 11 where we distinguish contributions from recombinations with the same partner or with a new partner. The theory of diffusion controlled recombination kinetics [12] predicts an algebraic decay $At^{-5/4}$ for $\Psi(t)$ and we observe that it is perfectly satisfied by our data for the self recombination part for times larger than $t = 1$. The picture shows how self-recombinations dominate at short times while recombinations with new partners dominate beyond $t = 30$. At very

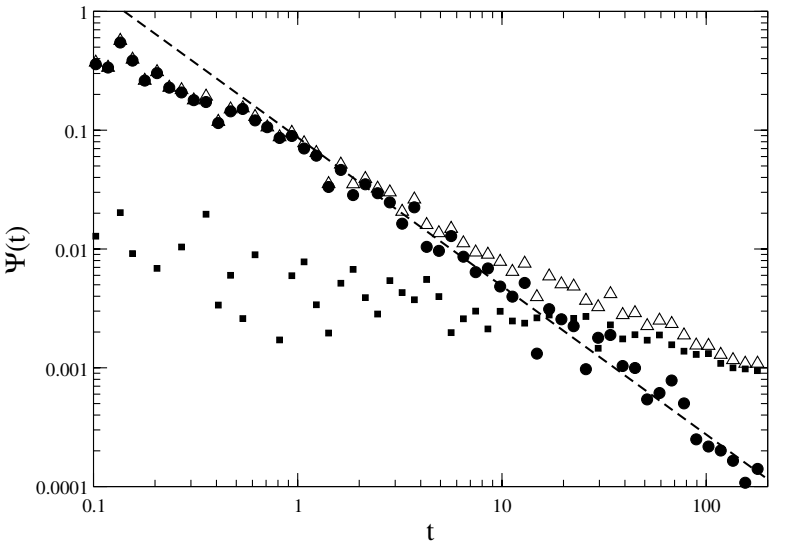


Fig. 11. Distribution of first recombination times $\Psi(t)$ for the semi-dilute case for attempt frequencies $\omega = 0.5$ (*open triangles*). The partition between contributions from recombinations with the same partner (*filled circles*) or with a new partner (*filled square*) are distinguished. The power law $0,087(t)^{-5/4}$ shows that the "diffusion-controlled" theoretical prediction of $-5/4$ for the power law exponent is perfectly satisfied for the self recombination part for times larger than $t = 1$. The picture shows how self-recombinations dominate at short times while recombinations with new partners dominate beyond $t = 30$

short times ($t < 1$), a significative part (about 10%) of the $\Psi(t)$ function implies recombinations with anorther partner.

5.2 Cumulative Hazard Analysis

To analyse the rates of conformational transitions in butane and other short alkane molecules, Helfand [24] suggested to exploit hazard rate plotting. We have adapted this technique to the present case. We first summarize the theoretical foundations of the method, using the particular case of equilibrium polymer kinetics to directly illustrate the concepts.

Let $h(t)dt$ be the probability that a free arm, created at time $t = 0$ and which is still free at time t, undergoes a first recombination in the interval $[t, t + dt]$. Let $P(t)$ be the probability that a free arm, created at time $t = 0$, has undergone a (first) recombination between 0 and t.

Using these definitions, the following steps can be written

$$[1 - P(t + dt)] = [1 - P(t)]\,(1 - h(t)dt) \tag{58}$$

$$\frac{d}{dt}\,[1 - P(t)] = -h(t)\,[1 - P(t)] \tag{59}$$

$$[1 - P(t)] = \exp\left(-\int_0^t h(t')dt'\right) = \exp\,[-H(t)] \tag{60}$$

where $H(t) = \int_0^t h(t')dt'$ is the cumulative hazard. In the present case, we anticipate a complex process involving correlated events at short times $h(t) = h^{sh}(t)$ and a Poisson process emerging at long times with uniform frequency $lim_{t\to\infty}h(t) = \lambda$ where λ is the recombination rate constant. On the relevant time scale of the kinetics, one thus should find for the cumulative hazard function and the $P(t)$ probability

$$H(t) = H_I + \lambda t \tag{61}$$

$$P(t) = 1 - \exp\,(-H_I)\exp\,(-\lambda t) \tag{62}$$

where H_I, the ordinate intercept of the function $H(t)$ versus t, can be seen as the time integral of $(h^{sh}(t) - \lambda)$ from 0 to ∞. If a good time scale separation exists,(62) implies that $c \equiv P(0^+) = 1-\exp\,(-H_I)$ is the fraction of correlated transitions, $\kappa^{(2)} = 1 - c$ being the estimate according to the present analysis of the average probability for a newly created free arm to recombine by the Poisson "mean field" kinetic process postulated by Cates in his theory [9].

We now explain how the estimate of the cumulative hazard function is constructed. We start by extracting from the BD trajectory a collection of times $\{\hat{t}\}$ where each member corresponds to the elapsed time between a scission of a particular arm at time t_1 and its next recombination at time t_2, so $\hat{t} = t_2 - t_1$. All the $2M$ arms of the system contribute to the $\hat{t}$ data sample which, for each arm, can contain several times of this kind between the start (at $t = 0$) and the end (at $t = T_{max}$) of the BD trajectory.

Moreover, if the first change of status of a particular arm since the beginning of the BD simulation is a recombination taking place at time t, we can say that this time is a lower bound $\hat{T}$ of an additional unknown elapsed time $\hat{t}$ between a scission (out of our reach) and the next recombination (we observed). Also, if the last change of arm status before the end of the BD trajectory is a scission taking place at time t, then, the time $\hat{T} = T_{max} - t$ is a similar lower bound of yet another time of interest. The analysis thus furnishes a set of K times $\{\hat{t}\}$ and M lower bounds times $\{\hat{T}\}$) which are then separately ordered from the shortest time up to the longer one and indexed accordingly as $(\hat{t}_1, \hat{t}_2, \hat{t}_3, ..\hat{t}_K)$ and $(\hat{T}_1, \hat{T}_2, \hat{T}_3, \ldots \hat{T}_M)$. Let us consider successively all individual arm life times $\{\hat{t}\}$. For any time $\hat{t}_i$ of that collection, the probability that a free arm which had survived up to the previous time $\hat{t}_{i-1}$ changes its status and becomes engaged in the formation of new bond between times $\hat{t}_{i-1}$ and $\hat{t}_i$ is given with our available statistics by $1/N(\hat{t}_{i-1})$ where the denominator is the number of cases (including both types of times $\hat{t}$ and $\hat{T}$) where recombination takes place at a time longer than $\hat{t}_{i-1}$, i.e.

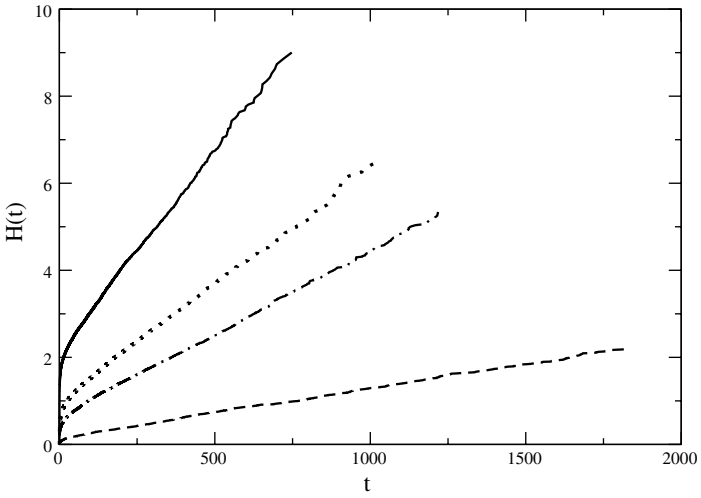


Fig. 12. Cumulative Hazard curves of recombination times for the semi-dilute case. Data are shown for various attempt frequencies, namely $\omega = 0.1$ (*dashed line*), $\omega = 0.5$ (*dot-dashed line*), $\omega = 1.0$ (*dotted line*) and $\omega = 5.0$ (*continuous line*)

$$N(\hat{t}_{i-1}) = [K - (i-1)] + [M - m_{(i-1)}] \tag{63}$$

where $m_{(i-1)}$ is the index of largest $\hat{T}$ value which is still inferior to $\hat{t}_{i-1}$. In terms of these definitions, the cumulative hazard function $H(t)$ can be evaluated at each time $\hat{t}$

$$H(\hat{t}_i) = \sum_{j=1}^{i-1} \frac{1}{N(t_j)} \tag{64}$$

If the function is linear in time, it implies a Poisson process with a rate of transitions given by the slope of that linear portion.

Figure 12 shows the resulting cumulative hazard functions for the four ω values in the semi-dilute case. A linear behaviour starts around 100–150 in all cases and it lasts up to times of the order of 1000 where statistical noise sets in. The inverse slopes of the linear portions provide estimates $\tau_b^{(2)}$ of the mean life time of a chain end. Specific values for both thermodynamic states are gathered in Table 1 and Table 2. They are compatible with the values extracted from the histogram of first recombination times. Using the procedure described at the beginning of this section, one gets from the ordinate intercepts of $H(t)$ the $\kappa^{(2)}$ estimate indicated in Table 2 for the semi dilute case.

The cumulative hazard plot has the advantage that it avoids the necessity of bining recombination time data which become scarse at long times. It also exploits statistically some additional information (useful for low rates) from portions of the $2M$ trajectories between a scission and the next recombination in a way which is truncated either at the beginning or at the end of the BD run.

5.3 The Reactive Flux Correlation Function and the Transmission Coefficient

We now discuss still another way to isolate short time transitional effects in order to estimate the rate of "effective" transtions. It relies on the way so-called "Transition State Theory" (TST) values of chemical reactions rates can be corrected by a multiplicative transmission coefficient $0 < \kappa < 1$ which takes into account the fraction of events where the products formed by the forward reaction are effective transitions. By effective, it means that after such a transition, the product is able to relax its energy excess (which was accumulated to cross the forward barrier) sufficiently quickly to avoid an immediate recrossing of the barrier in the backward direction.

In our case, the analogy is clear. If a bond scission is followed by the same bond recombination within a short microscopic time scale (too short for relative diffusion to take place), it basically means that no effective scission has happened to the chain to which the bond belongs on time scales larger than the relaxation time of a chain fragment of a few monomers.

We have thus estimated $\kappa^{(3)}$ by exploiting the concept of the reactive flux correlation function, as explained by Chandler in the framework of an equilibrium between forward and backward unimolecular reactions [19]. We assume here that for an equilibrium between forward unimolecular reactions (chain breaking) and bimolecular backward reactions (end-chain recombinations), the same conceptual expression can still be applied for the forward reaction.

Let us consider one specific arm (among the two of one particular monomer) which is labelled by an index i and to each arm, we associate a signature variable $S_i(t)$ which is equal to unity if the ith arm is free at time t and equal to zero if the arm is engaged in a binding with another monomer at that time. Crossings are then characterized by a step change of the $S_i(t)$ values, $S_i(t)$ going from 0 to 1 or from 1 to 0 for scissions and recombinations respectively. A statistics over all crossing events is defined: for each individual crossing, we set the time at crossing to zero, denoting times very shortly after or before the transition by $t = 0^+$ and $t = 0^-$. The reactive flux correlation function is then defined by [25]

$$C_{RF}(t) = 2\langle \left[S_i(0^+) - S_i(0^-)\right] S_i(t)\rangle_{cross} \tag{65}$$

where the average is performed over all crossings (both scissions and recombinations). If we treat separately the averages on scissions and recombinations, we get the equivalent alternative formulation

$$C_{RF}(t) = \langle S_i(t)\rangle_{scissions} - \langle S_i(t)\rangle_{recombinations} \tag{66}$$

This function starts from unity at very short (positive) times with the first term being one and the second being zero. On very long times (long respect to the relevant inverse kinetic rate constants), this function goes to zero as it is the difference of two averages which ultimately converge to the same equilibrium average value $\langle S_i\rangle$.

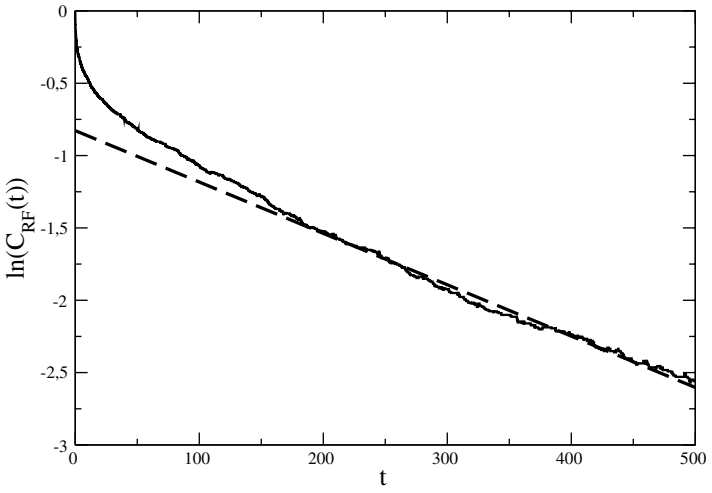


Fig. 13. Reactive flux correlation function in the semi-dilute case for $\omega = 0.5$. The *dashed line* corresponds to a fit by a function $\kappa \exp(-t/\tau_b)$ in the time region $150 < t < 500$. Best parameters found are $\kappa^{(3)} = 0.48$ and $\tau_b^{(3)} = 282$

If all transitions are effective, the reactive flux function will decrease smoothly from 1 to zero reflecting only the kinetics of these effective recombination and scissions. If transitions are often followed by opposite transitions within a short time scale τ_{micro} or if successive transitions are correlated within such a short time, this will be reflected in the reactive flux by a fast initial decay over that microscopic time. At later times, the macrocopic kinetics (the one considerd by mean field theoretical expresions) will be reflected by the $C_{RF}(t)$ evolution towards zero on a macroscopic time scale, τ_{macro}. The transmission coefficient $\kappa^{(3)}$ is thus estimated as the initial value of the "slow" decay part of that $C_{RF}(t)$ function. This can be expressed as

$$\kappa^{(3)} = C_{RF}(\tau_{micro}) \tag{67}$$

with the requirement of a good time scale separation, that is $\tau_{micro} \ll \tau_{macro}$.

Figure 13 shows the $\omega = 0.5$ case for the semi-dilute case already discussed in Fig. 11 in the context of the distribution of first recombination after a scission. This curve is more difficult to analyse because it is difficult on the basis of this sole function to decide when the behaviour of $ln(C_{RF}(t))$ really becomes close from a linear behaviour. We had to rely on our experience with the distribution $\Psi(t)$ and the cumulative hazard function $H(t)$ for that state point and ω value. As both functions suggest that the simple "mean field" process emerges at times later than 150, we have fitted the $C_{RF}(t)$ data by an exponential function $C_{RF}^{fit}(t) = \kappa^{(3)} * \exp(-t/\tau_b^{(3)})$ in the accessible time window above $t = 150$. In the present case, we see that there is no real time scale separation between a short time behaviour, observed in the

$0 < t < 150$ regime, and the macroscopic time scale $\tau_b^{(3)}$ which is equal to ≈ 280. Nevertheless, the value of $\tau_b^{(3)}$ turns out to be rather close to the other estimates of a chain end average life time and the $\kappa^{(3)}$ value will be shown in the next subsection to lead to a reasonable correction of n_s, the total number of transitions per unit of time and unit volume quoted in Tables 1 and 2. Results of $\kappa^{(3)}$ and $\tau_b^{(3)}$ for all ω's and both state points are included in the Tables 1 and 2.

5.4 Estimation of Rate Constants: Comparison of the Various Methods

Explicitly, the scission rate constants k_s and k_r of the "mean-field" kinetic model can be estimated from our data in two distinct ways.

- On the basis of the total number of transitions n_s per unit time and per unit volume, and on the basis of an estimate of κ by one of the methods discussed above, the rates can be estimated from a trivial modification of (40, 41), namely
$$\kappa n_s = k_s \phi = \frac{k_r \phi^2}{2L_0^2} \tag{68}$$
- On the basis of the long time behaviour of a dynamical function relevant to the particular methodology adopted, the "chain end" life time can be directly estimated. As the latter is equivalent to the typical life time of a chain of average length denoted as τ_b, one gets rate constants through the equivalence
$$\tau_b = \frac{1}{k_s L_0} = \frac{2L_0}{k_r \phi} \tag{69}$$
where the last equality follows from detailed balance requirements.

Tables 1 and 2 provide all needed data. We find that the implicit time scale separation inherent to all analyses provide compatible estimates of the effective rate constants. This can be verified by applying (68) and (69) with the particular estimate of the left hand side.

Looking at all values in the tables, we get an overall consistency between all methodologies. It indicates that all our strategies to extract the "macroscopic" rate constants do work. Among them, the cumulative hazard analysis appears the most straightforward to analyse once $H(t)$ is known and particularly robust as noise influence seems minimal.

We are now in a position to compare the percentage of effective transitions κ based on a time scale separation and the α value based on the percentage of recombinations involving a new partner respect to the one left by scission. We observe in the two tables that both estimates are reasonably consistent and quickly decrease for ω increasing. Fine analysis suggests however that some distinction exists between both quantities in the present case, that is

a free Rouse relaxation type of dynamics in solution for moderately long chains. It requires further analysis and the consideration of more entangled cases to make progress on that point. Finally, it should be stressed that while diffusion controlled and mean field kinetics have sometimes been opposed to each other [16], it appears in our analysis that they consistently describe short time correlation effects dominated by self recombinations and long time kinetics dominated by recombinations with new partners.

5.5 Analysis of the Monomer Diffusion

The mean squared displacement (MSD) of all monomers as a function of time is shown in Fig. 14 for all ω's in the semi-dilute case. As these monomers belong to a polydisperse set of chains which continuously break and recombine, the interpretation of the MSD is rather complex.

- Being a single monomer property, a relevant quantity is the average fraction of monomers belonging to chains of length L, namely a distribution $\propto L{*}c_0(L)$. Figure 15 shows the resulting curve for the adopted semi-dilute conditions with $\langle L\rangle = 49.9$.
- The time scale for the scission/recombination process has been estimated so far in terms of the average survival time $1/(k_s\langle L\rangle)$ for the polymer of

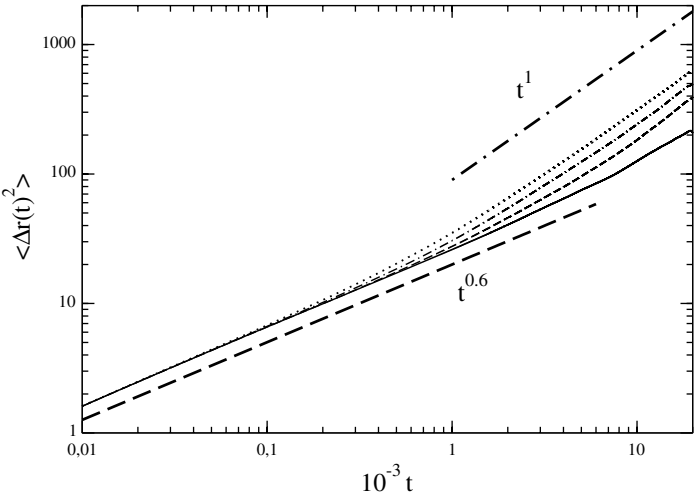


Fig. 14. Mean square displacement of monomers (semi-dilute case). The continuous curve has been obtained for the relevant polydisperse sample but with ω set to zero, averaging over many independent initial configurations. It is a reference from which the other curves depart at decreasing times when increasing values of ω are considered. The cases $\omega = 0.1$ (*dashed lines*), $\omega = 0.5$ (*dot-dashed lines*) and $\omega = 5.0$ (*dotted lines*) are shown together with two indicative power laws which suggest the evolution of the subdiffusive scaling regime to the normal diffusion limit

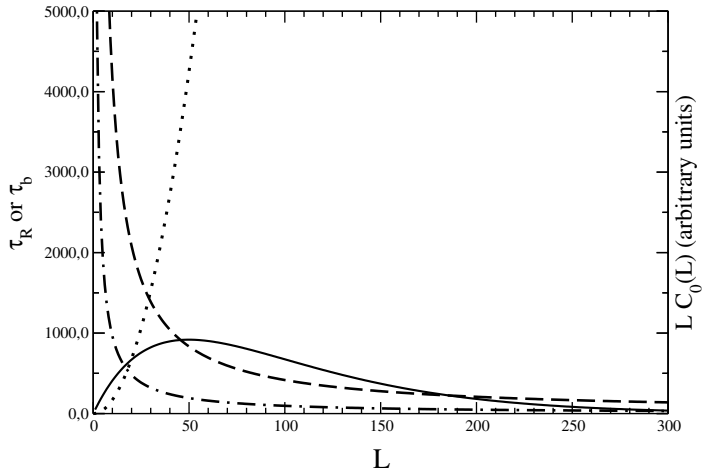


Fig. 15. Characteristic times as a function of polymer size. The Rouse relaxation time $\tau_R = 1.70 * L^2$, shown as *dotted curve*, must be compared to the life time of a polymer of size L, as given $\tau_b = (k_s L)^{-1}$. Two frequencies are illustrated, namely $\omega = 1.0$ (*dot-dashed line*) and $\omega = 0.1$ (*dashed line*) using effective k_s values. We show on the same graph the distribution of monomers (*continuous curve*) as a function of the size of the polymer they belong to (see text)

average size, a quantity which was found to vary from ≈100 up to ≈ 900 in the investigated ω range (see Table 2). In the present context, we need to consider the explicit L dependent average survival time $\propto 1/(k_s L)$ for a given ω value and Fig. 15 shows the resulting functions for two such frequencies.

- To discuss the monomer mean squared displacement, it is also important to have an estimate of the longest internal relaxation (Rouse) time of the chains. Performing a separate brownian dynamics run on a monodisperse sample of 20 "dead" polymers of length $L = 50$ (corresponding to the mean polymer length of our living polymer sample) at the same semi-dilute state point ($T = 1$, $\phi = 0.15$), we get from the long time exponential bevaviour of the end-to-end vector relaxation function a value of $\tau_R = 4.24\ 10^3$. For our very flexible polymers, the scaling of the Rouse time with L in the semi-dilute conditions should scale as

$$\tau_R(L) = AL^2 \ . \tag{70}$$

On the basis of our estimated value for $L = 50$ chains, we get $A \approx 1.70$, a value which can be compared (at least its order of magnitude) to theoretical estimate [6, 26] $A \approx \tau_{blob}/(L^*)^2$ in terms of τ_{blob}, the blob relaxation time, and L^*, the number of monomers per blob. We get $A = 5.8$ for our conditions (see discussion in the section on chain con-

formational properties) where the number of monomers per blob $L^* \approx 37$ and $\tau_{blob} \approx \frac{\eta_s \chi^3}{T} \approx 8000$ using a blob size $\chi = 20$ and a solvent viscosity $\eta_s = \frac{\xi}{3\pi\sigma} = 1$ estimated from the friction coefficient using the Stokes law with an hydrodynamic radius fixed to $\sigma/2$.

On the basis of the above considerations, it is useful to define a particular chain length Λ for which the Rouse time is equal to the survival time, giving $\Lambda = (1.7 * k_s)^{-1/3}$, that is $\Lambda = 30, 20, 18, 15$ for $\omega = 0.1, 0.5, 1.0, 5.0$ respectively. Chains longer than Λ have their internal dynamics strongly altered by scissions, while chains shorter than Λ should have an internal dynamics little affected with respect to dead polymers of the same size. As Fig. 15 illustrates, when the scission attempt frequency increases, the fraction of monomers which belong to chains whose dynamics is little affected by scissions decreases.

Up to a time of 100, the MSD shown in Fig. 14 is ω independent and presents a Rouse like behaviour with a time scaling exponent $t^{0.6}$. In the reference case where ω was set to zero (we simulate polydisperse sample of dead polymers), this power law bahaviour persists at least up to time 10000, a time corresponding to the Rouse time of a chain of length 170. When scissions and recombinations are allowed, the MDS deviates from the master curve (of dead polymers) to adopt progressively a MSD linear in time. This takes place sooner and sooner for increasing values of ω. At long times ($t > 2000$), the MSD has evolved towards usual Einstein diffusion with a diffusion coefficient increasing with k_s. In Fig. 16, D is observed to (very roughly) scale like $k_s^{0.23}$ in the explored range, a result which is not so far from the 1/3 power behaviour observed by Milchev with the lattice BFM lattice [14], but which requires further analysis with better estimates of D as statistics is poor above $t = 10000$. The 1/3 power law scaling is explained by assuming that clusters of Λ monomers are responsible of the long time behaviour of living polymers, giving $D \propto \frac{R_g^2(\Lambda)}{\tau_{Rouse}(\Lambda)} \propto \Lambda^{-1} \propto k_s^{1/3}$.

5.6 Analysis of the First Recombination Times Distribution and Corresponding Diffusive Steps

In Fig. 17, we report for two different bond change attempt frequencies ω, the mean squared displacement of a chain end between its creation and its next recombination, as a function of the corresponding life time of that chain end. Interestingly, we observe a common behaviour of this dependence for both frequencies, meaning that the diffusive dynamics of chain ends is little sensitive to the scission/recombination processes of the chains as a whole.

As chain ends have very different life times for different ω's, the typical average distance travelled by a chain end between its creation and its recombination is also function of that frequency. Mean travel distances of 3.8, 4.6 and 6.8 are obtained for the average chain end life times at $\omega = 5.0$, 1.0 and 0.1 respectively, to be compared with the mean static distance between chain ends (assuming no spatial correlations) given by $h = (\langle L \rangle / 2\phi)^{\frac{1}{3}}$ which turns to be

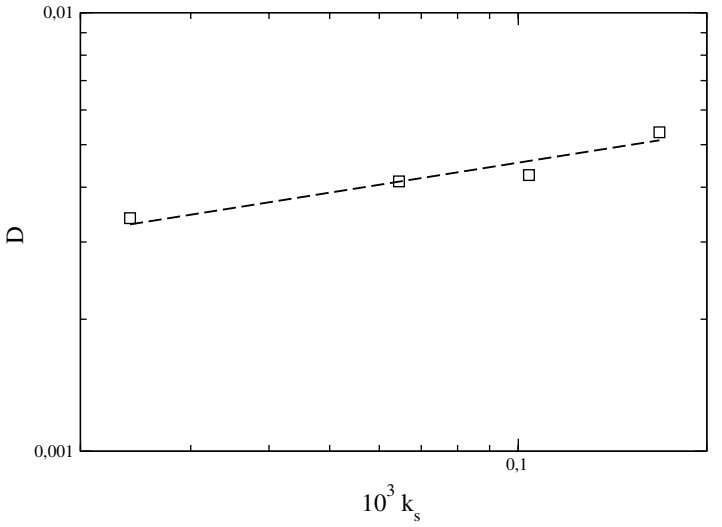


Fig. 16. Monomeric Einstein diffusion coefficient as a function of k_s for the semi-dilute case. The value of D was obtained from the long time behaviour (beyond $t = 8000$) of the monomeric mean square displacement shown in Fig. 14. The *dashed line* represent a $k_s^{0.23}$ power law

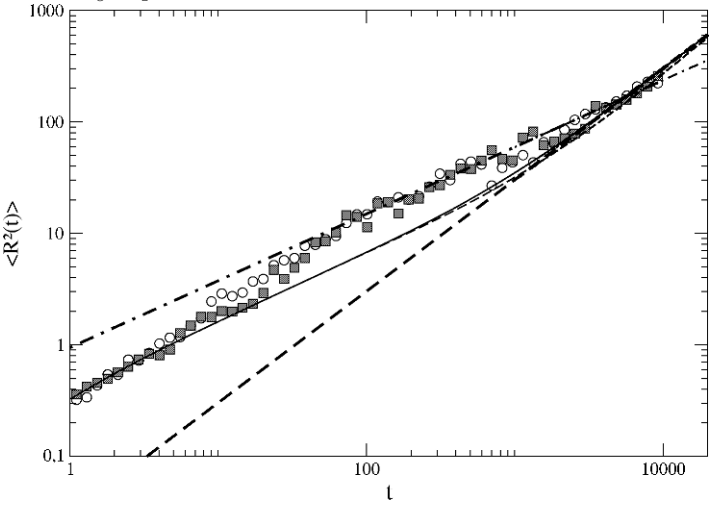


Fig. 17. Mean square displacement of end chain monomers between a time set to zero when produced by scission up to a time t corresponding to their first recombination, as a function of their total life time t. Due to poor statistics, only data for $\omega = 1$ (*filled squares*) and $\omega = 5$ (*empty circles*) are shown and are limited to $t = 10000$ maximum. Data on the mean square displacement of all monomers for corresponding ω values already shown in Fig. 14 are shown again here for comparison (see text) up to a maximum time of 20000. Scaling power laws $1.0t^{1.0}$ (*thick dashed line*) and $0.943t^{0.6}$ (*thick dot-dashed line*) shown up to $t = 20000$ are used in the text

$h = 5.5$ in our semi-dilute case. By comparison, the radius of gyration of the average chain length in the system is $R_g(< L >) \approx 10$. The dynamic picture is thus that chain ends diffuse similarly with time in an environment of other chains which continuously break and recombine. As the reacting frequency increases, the probability of a chain end to recombine with a neighbouring monomer increases and thus takes place earlier, after diffusion over a shorter distance.

6 Summary and Conclusions

In this paper, we have proposed a new version of a mesoscopic model of equilibrium polymers in solution and we have exploited this model to study the structure and the dynamics of these complex systems at a dilute and at a semi-dilute state point.

The dynamics at the semi-dilute state point has been more specifically investigated. Our system consists of chains with an average number of $L_0 \approx 50$ monomers which can be viewed as a collection of blobs of ≈ 37 monomers. Given the large flexibility of the model and the moderately semi-dilute character, our polydisperse polymer solution should be in a non-entangled dynamic regime, with a chain relaxation time growing with chain length L as L^2.

Working with a unique chain length distribution governed by the structural and thermodynamic microscopic parameters, we have analyzed the specific influence of the Monte Carlo binding/unbinding change attempt frequency on the dynamical properties. Mean field theories are found to be valid for times of the order and beyond the mean life time of a chain of average size provided the kinetic constants computed from the Monte Carlo accepted binding/unbinding changes are rescaled by a transmission coefficient which is a measure of the fraction of successful scissions. This transmission coefficient, when estimated on the basis of a time scale separation, turns out to be close to the fraction of recombinations involving the reunion of a chain end with a new partner and no longer with the monomer it was attached to previously.

The rate constants have been estimated by different techniques which give coherent estimates. The cumulative hazard function introduced by Helfand to compute isomerisation rates in chain molecules has been found to be particularly efficient and simple to implement. In paticular, we find agreement between 1) kinetic constant estimates based on the the total number of transitions rescaled by the transmission coefficient coming from a diffusion controlled mechanism and 2) estimates based directly on the time scale of the long time exponential decay of relaxation functions (mean field Poisson process). It would be interesting to perform so-called T-jump experiments [26] with our system to test whether the time evolution of the average polymer length L_0 towards its new equilibrium value at a different temperature confirms the mean field analytical prediction with the kinetic constants computed in the present work from various population relaxation functions at equilibrium.

The chain size that we chose originally was adequate to demonstrate the validity of the methodological aspects developped in this work. In order to test dynamical scaling with a wider range of the scission-recombination process frequency, longer chains will be needed which may require better optimized methods to take solvent effects into account. The extension of this study towards the rheological properties of our model system are under way.

Acknowledgments

This work has taken profit of numerous discussions with M. Baus of the Univeristé Libre de Bruxelles. C.C.H. was financially supported by the Action de Recherche Concertée N^0 00/05-257 of the Communauté Française de Belgique. The simulations have been performed thanks to ressources from the Centre de Calcul of the ULB/VUB and the French Centre CINES, which we acknowledge.

References

1. P. van der Schoot (2005) *in "Supramolecular polymers"*. Ed. A. Ciferri **77**, N.Y.
2. M. E. Cates and S. J. Candau (1990) Statics and dynamics of worm-like surfactant micelles. *J. Phys: Cond. Matt.* **2**, p. 6869
3. S. Lerouge, J. P. Decruppe, and C. Humbert (1998) Shear Banding in a Micellar Solution under Transient Flow. *Phys. Rev. Lett.* **81**, p. 5457
4. H. Hoffmann, S. Hoffmann, A. Rauscher, and J. Kalus (1991) Prog. *Colloid Polym. Sci.* **84**, p. 24
5. A. Ott, J. P. Bouchaud, D. Langevin, and W. Urbach (1990) Anomalous diffusion in "living polymers": A genuine Levy flight? *Phys. Rev. Lett.* **65**, p. 2201
6. P. G. de Gennes (1979) *Scaling concepts in polymer physics.* Cornell University Press, Ithaca
7. A. Y. Grosberg and A. R. Khokhlov (1994) *Statistical Physics of Macromolecules.* AIP Press, New-York
8. M. Doi and S.F. Edwards (1986) *The theory of polymer dynamics.* Clarendon, Oxford
9. M. E. Cates (1987) Reptation of living polymers: Dynamics of entangled polymers in the presence of reversible chain-scission reactions. *Macromolecules* **20**, p. 2289
10. N. A. Spenley, M. E. Cates and T. C. B.MacLeish (1993) Nonlinear rheology of wormlike micelles. *Phys. Rev. Lett.* **71**, p. 939
11. S. M. Fielding and P. D. Olmsted (2003) Early Stage Kinetics in a Unified Model of Shear-Induced Demixing and Mechanical Shear Banding Instabilities. *Phys. Rev. Lett.* **90**, p. 224501
12. B. O'Shaughnessy and J. Yu (1995) Rheology of Wormlike Micelles: Two Universality Classes. *Phys. Rev. Lett.* **74**, p. 4329
13. J. P. Wittmer, A. Milchev and M. E. Cates (1998) Computational confirmation of scaling predictions for equilibrium polymers. *Europhys. Lett.* **41**, p. 291; *J. Chem. Phys.* **109**, p. 834

14. A. Milchev (2000) *in "Computational methods in colloid and interface science"*. Ed. M. Borowko and M. Dekker **510**, N.Y.
15. Y. Rouault (1999) Off-lattice Brownian dynamics simulation of wormlike micelles: The dependence of the mean contour length on concentration. *J. Chem. Phys.* **111**, p. 9859
16. J. T. Padding and E. S. Boek (2004) Evidence for diffusion-controlled recombination kinetics in model wormlike micelles. *Europhys. Lett.* **66**, p. 756
17. M. Kroger and R. Makhloufi (1996) Wormlike micelles under shear flow: A microscopic model studied by nonequilibrium-molecular-dynamics computer simulations. *Phys. Rev. E* **53**, p. 2531
18. C. C. Huang, H. Xu, and J. P. Ryckaert (2006) Kinetics and dynamic properties of equilibrum polymers. *J. Chem. Phys.* (accepted)
19. D. Chandler (1987) *Introduction to Modern Statistical Mechanics*. Oxford University Press
20. K. Kremer and G. S. Grest (1990) Dynamics of entangled linear polymer melts: A molecular-dynamics simulation. *J. Chem. Phys.* **92**, p. 5057
21. A. Milchev, J. P. Wittmer, and D. P. Landau (1999) Monte Carlo Study of Equilibrium Polymers in a Shear Flow. *EPJ* B, **12**(2), pp. 241–251; cond-mat/9905336
22. A. Milchev, J. P. Wittmer, P. van der Schoot, D. Landau (2001) Osmotic pressure of solutions containing flexible polymers subject to an annealed molecular weight distribution *Europhysics Letters*, **54**, pp. 58–64; conf-mat/0008276.
23. J. P. Wittmer, M. Milchev, P. van der Schoot, J.-L. Barrat (2000) Dynamical Monte Carlo study of equilibrium polymers. II. The role of rings. *J. Chem. Phys.* **113**, 6992–7005; cond-mat/0006465.
24. E. Helfand (1978) Brownian dynamics study of transitions in a polymer chain of bistable oscillators. *J. Chem. Phys.* **69**, p. 1010
25. D. Brown and J. H. R. Clarke (1990) On the determination of rate constants from equilibrium molecular dynamics simulations. *J. Chem. Phys.* **93**, p. 4117
26. Y. Rouault and A. Milchev (1996) A Monte Carlo study of diffusion in "living polymers". *Eur. Phys. Lett.* **33**, p. 341

Equilibration and Coarse-Graining Methods for Polymers

D.N. Theodorou

School of Chemical Engineering, National Technical University of Athens,
9 Heroon Polytechniou Street, Zografou Campus, 157 80 Athens, Greece
doros@central.ntua.gr



Georgios Boulougouris, who kindly accepted to present the talk in Erice on behalf of Doros Theodorou, who had to cancel his participation at the very last minute

D.N. Theodorou: *Equilibration and Corase-Graining Methods for Polymers*, Lect. Notes Phys. **704**, 419–448 (2006)
DOI 10.1007/3-540-35284-8_16
© Springer-Verlag Berlin Heidelberg 2006

1 Introduction 421

2 Connectivity-Altering Algorithms for the Efficient Equilibration of Polymer Melts 423

3 Coarse-Graining of Polymer Models 426

3.1 Iterative Boltzmann Inversion 427
3.2 Coarse-Graining Using Pretabulated Potentials 431

4 Chain Entanglements in Polymer Melts 436

5 Topological Analysis of Atomistic Melt Configurations and Reduction to Entanglement Networks 439

6 Concluding Remarks 444

References 446

The broad spectra of length and time scales governing the physical properties of polymers call for hierarchical modelling methods, based on systematic coarse-graining of the molecular representation. This chapter examines some particularly promising elements of such a modelling hierarchy for predicting polymer properties. Starting from the atomistic level, a class of connectivity-altering Monte Carlo algorithms capable of equilibrating long-chain polymer melts at all length scales is briefly discussed. Coarse-graining strategies which reduce detailed atomistic models into models involving fewer degrees of freedom are then addressed, using the Iterative Boltzmann Inversion technique of Müller-Plathe et al. and a new, multidimensional potential pretabulation technique as examples. Finally, the Contour-Reduction Topological Analysis (CReTA) method is discussed, whereby detailed atomistic or coarse-grained polymer melt configurations are reduced to networks of entanglements, useful in the analysis and simulation of rheological properties. Examples are presented from application of these methods to vinyl polymer and polydiene melts.

1 Introduction

Understanding and predicting the relations between structure, properties, processing, and performance is a central goal of materials science and engineering, since it can form a basis for faster and more economical design of materials and processes for specific applications.

A challenge faced by materials modellers, which is especially serious in the case of polymeric materials, is that structure and dynamics are characterized by extremely broad spectra of length and time scales. Intramolecular correlations and local packing of chains in the bulk exhibit features on the length scale of bond lengths and atomic radii, i.e. 0.1 nm. The length b of a Kuhn (statistical) segment, defined such that the real chain can be considered as a random flight of such segments, is on the order of 1 nm for a typical synthetic randomly coiled polymer and can be more than an order of magnitude higher for macromolecules with stiff backbones. The radius of gyration of entire chains in the amorphous bulk scales with the chain length N as $N^{1/2}$ and is on the order of 10 nm for typical molecular weights. Domain sizes of different phases in semicrystalline polymers and in immiscible polymer blends may well be on the order of 1 μm. Even broader is the range of time scales characterizing molecular motion in polymers. While localized vibrational modes of chains have periods on the order of 10^{-14} s, conformational transitions of individual bonds over torsional energy barriers in the melt state have waiting times in excess of 10^{-11} s. Longer and longer sequences of segments along the backbone exhibit longer and longer correlation times. The longest relaxation time, required for a chain to diffuse by a length commensurate to its size and thus 'forget' its previous conformation, controls the viscoelastic response of polymer melts to flow. This time scales as N^2 for low molecular weight

melts in the Rouse regime and as $N^{3.4}$ for N exceeding a critical chain length sufficient for the development of entanglements in the reptation regime; for the molecular weights encountered in typical processing operations, it easily exceeds the millisecond range. Microphase and phase separation phenomena in molten copolymers and polymer mixtures may have characteristic times longer than seconds, whereas structural relaxation in the glassy state, 20°C below the glass temperature, has characteristic times on the order of a year (10^7 s).

This tremendous multiplicity of characteristic times and length scales, to which polymeric materials largely owe their widespread technological use, makes it clear that a single modelling or simulation approach cannot address all properties of these materials. The development of hierachical, or multiscale modelling approaches is imperative. Such approaches consist of many levels, each level addressing phenomena over a specific window of length and time scales [1, 2]; links between different levels can be established through systematic 'coarse-graining' of the model used to represent the material. At the base of this hierarchy, powerful atomistic simulation algorithms are needed, which can efficiently sample the complex potential energy hypersurfaces of polymeric systems, ensuring equilibration over the relevant time and length scales.

In this chapter we discuss some key methods that can serve as building blocks of hierarchical computational approaches for predicting polymer properties. Section 2 briefly discusses a class of connectivity-altering Monte Carlo (MC) algorithms that can fully equilibrate long-chain model polymer melts at both atomistic and coarse-grained levels. Section 3 addresses the problem of coarse-graining, i.e., deriving from a detailed atomistic model a model with fewer degrees of freedom that is capable of capturing the long length- and time scale properties of the original model in a small fraction of the computational cost. Two continuous-space coarse-graining approaches are discussed. In the first, chains are represented as sets of spherical "super-atoms;" the potentials of mean force describing interactions among such superatoms are obtained through iterative Boltzmann inversion of pair distribution functions accumulated in the course of atomistic simulations of short-chain analogues. In the second coarse-graining approach, chains are represented as sequences of inflexible moieties of arbitrary shape. Nonbonded interactions between pairs of these moieties are computed atomistically and pretabulated as functions of four or six degrees of freedom describing the relative configuration of the moieties. The final parts of the chapter address a "coarser" level of coarse-graining: one wherein dense amorphous polymer configurations are mapped onto networks of chain ends and entanglement points. Such network representations are valuable in predicting the rheological properties of polymer melts and the terminal mechanical properties of glassy polymers. Section 4 gives a brief introduction to the concepts of entanglements and reptative motion and of a simple approach that has been advanced to relate the molar mass between entanglements to chemical constitution. Section 5 introduces an algorithm for the topological analysis of atomistic (or coarse-grained) configurations and

their reduction to entanglement networks. Key distributions and parameters characterizing the structure of the entanglement network, as extracted by the algorithm, are discussed and compared to experimental evidence. Section 6 summarizes some important points.

2 Connectivity-Altering Algorithms for the Efficient Equilibration of Polymer Melts

Present-day computer hardware and dynamic simulation algorithms allow conducting atomistic molecular dynamics (MD) runs of duration several μs. This is still short in comparison to the longest relaxation times of high-molecular weight melts, which are often in the ms to s range. Thus, equilibrating the long-length scale properties of dense systems of long chains, such as the mean squared end-to-end distance $\langle R^2 \rangle$ and the mean squared radius of gyration $\langle R_g^2 \rangle$, is problematic with MD. One may argue that equilibration at the length scale of entire chains is not necessary, if one is interested in properties which depend on local packing and segmental dynamics (e.g., permeability by small molecules). It has been shown, however, that failure to equilibrate at the chain level may affect local packing in subtle ways [3, 4]. Furthermore, there is a host of properties of crucial importance in polymer synthesis, processing and end-use applications (e.g., viscosity, chain self-diffusivity, storage and loss moduli) which depend sensitively on the conformation of entire chains and therefore require full equilibration at all length scales in order to be predicted reliably.

MC offers the possibility of designing moves that can provide much more vigorous sampling of configuration space than MD. Since the long relaxation times of polymers are ultimately due to the connectivity of segments along chains, moves that modify this connectivity afford much faster convergence of long-length scale conformational properties. In the last decade, a class of MC algorithms based on this idea of "cutting the Gordian knot" and rearranging connectivity has been very useful in achieving equilibration of dense, long-chain polymer systems at all length scales. Figure 1 provides an example of a connectivity-altering MC move, termed double bridging (DB), which is appropriate for monodisperse polymer systems. In a DB move, one excises a short chain segment (here, a trimer) from each one of two neighboring chains. One subsequently builds two new short segments (here, trimers) which bridge the remaining chain sections in a different way. The result is two new chains of exactly the same length, but drastically different conformation. A simpler connectivity-altering move, end bridging (EB), involves a chain attacking another (victim) chain through one of its ends and appending a terminal section of the victim chain. EB is appropriate for polydisperse polymer systems. It has been cast in a semigrand canonical ensemble formalism, which ensures that the simulated system at equilibrium conforms to a prescribed chain length distribution.

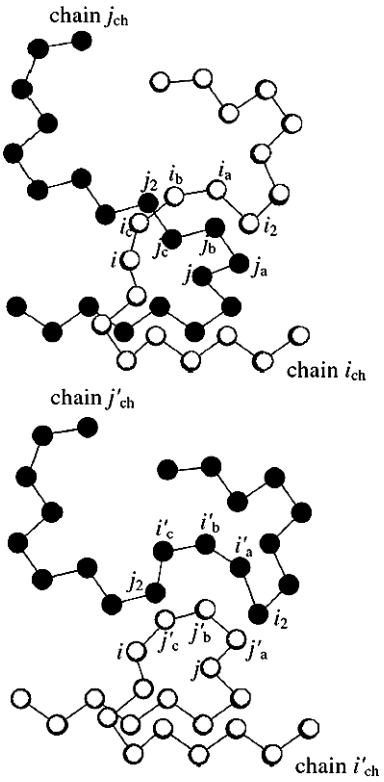


Fig. 1. Schematic of the double bridging (DB) move in a monodisperse system. *Top*: local configuration of the two chains before the move. *Bottom*: local configuration of the two chains after the move

In practice, connectivity-altering moves such as EB and DB are combined with moves which bring about local conformational rearrangement of terminal or internal segments of chains, such as end rotation, Configuration Bias (CB), flip, and Concerted Rotation (CONROT) moves. In atomistic polymer models, each bridging construction subject to prescribed bond lengths and bond angles is an interesting geometric problem, which admits multiple solutions. The multiplicity of solutions can be exploited in bias schemes designed to minimize excluded volume interactions with surrounding chains. Use of the bridging construction entails a transformation of coordinates with respect to the original Cartesian coordinates of skeletal atoms. The Jacobian of this transformation must be incorporated in the MC selection criteria, to ensure detailed balance. A detailed review of the geometric bridging construction, CONROT, EB, DB, and their variants [Directed Internal Bridging (DIB) or Self-Adapting Fixed Endpoint Configuration Bias (SAFE-CB), Directed End Bridging (DEB), scission and fusion algorithms], including acceptance criteria and measures of efficiency, has been provided in [5]. Some newer moves

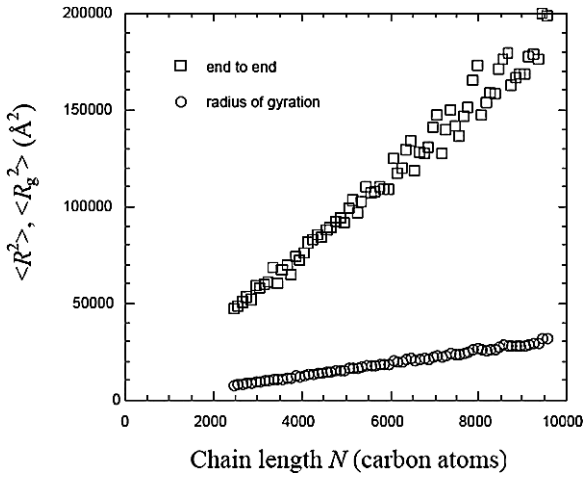


Fig. 2. Mean square end-to-end distance $\langle R^2 \rangle$ and mean square radius of gyration $\langle R_g^2 \rangle$ as functions of the chain length N from a connectivity-altering simulation of a C_{6000} PE melt, where chain lengths range from 2400 to 9600. The temperature of the simulation is 450 K and the pressure 1 bar

(Double Flip, Branch Point Flip, Branch Point Slithering) appropriate for the simulation of branched macromolecules have been described by Peristeras et al. [6]. So far, connectivity-altering MC strategies have been applied to linear, star-shaped, and H-shaped polyethylene (PE), cis-1,4 and 1,2 polybutadiene, cis-1,4 and trans-1,4 polyisoprene, and polyethylene oxide. When used in conjunction with potentials that have been well calibrated against volumetric and thermodynamic properties of low-molecular weight analogues, they provide excellent predictions of structure and thermodynamics of long-chain polymer melts.

Figure 2 provides an illustration of the ability of connectivity-altering MC algorithms to equilibrate long-chain melts at all length scales. It displays the mean squared end-to-end distance $\langle R^2 \rangle$ and the mean squared radius of gyration $\langle R_g^2 \rangle$ as functions of the chain length N (number of skeletal carbons), as obtained from MC simulation of a molten sample of 24 chains of linear PE with uniform (rectangular) distribution of chain lengths ranging from 2400 to 9600 carbon atoms at 450 K and 1 bar. The mean chain length is C_{6000} and the mean molar mass is 84000 g mol^{-1}, comparable to those encountered in plastics processing operations. The simulation [7] employed DEB, along with CONROT, reptation, flip, end rotation, and volume fluctuation moves. The dependence of both $\langle R^2 \rangle$ and $\langle R_g^2 \rangle$ on N is linear, conforming to Flory's "random coil hypothesis." This indicates that the system is well-equilibrated at the level of entire chains and that perturbations of chain dimensions due to the finite size of the simulation box are absent. The slope of the $\langle R_g^2 \rangle (N)$ curve is 1/6 the slope of the $\langle R^2 \rangle (N)$ curve, as expected for long random

coils. From this slope, one can immediately estimate the characteristic ratio $C_\infty = \lim_{N\to\infty} \frac{\langle R^2 \rangle}{N\ell^2}$ and the Kuhn length $b = \langle R^2 \rangle / L$ of the chains, with ℓ and L being the skeletal bond length and the chain contour length, respectively.

3 Coarse-Graining of Polymer Models

For systems of chains possessing bulky inflexible (e.g., aromatic) moieties along their backbones (e.g., poly(ethylene terephthalate), PET) or as side chains (e.g., polystyrene), direct atomistic MC simulation becomes expensive computationally; connectivity-altering MC moves, such as the ones discussed in the previous section, have very low acceptance rates, when implemented directly at the atomistic level. This difficulty can be overcome by resorting to coarse-grained models, which involve many fewer degrees of freedom for the representation of the material. Given a detailed atomistic model for the polymer chains, rather than trying to equilibrate dense phases of high molar mass directly at the atomistic level, it is often more efficient to (a) reduce the atomistic model into a coarse-grained one; (b) equilibrate at the coarse-grained level with MD or MC, e.g. using the algorithms discussed in Sect. 2; (c) "reverse-map" the equilibrated coarse-grained configurations into appropriate atomistic ones, if necessary. The equilibrated coarse-grained configurations may be useful in their own right, if the properties one is interested in are determined by structure and dynamics at higher than atomistic levels.

Coarse-graining is central to hierarchical, or multiscale modelling of polymers. Understandably, it has been a focus of attention since at least a decade. There is certainly no unique way to coarse-grain. A good coarse-grained model must be able to discriminate the effect of chemical constitution on the properties of interest. What level of detail is required for this, however, depends on the properties at hand. A general framework for coarse-graining the dynamical description of multibody systems is provided by the projection operator formalism of Zwanzig and Mori [8], which, however, is seldom applied to practical problems without approximations. In this discussion we will confine ourselves to coarse-grained models appropriate for capturing the structure and thermodynamic properties of polymer melts, leaving the more complex problem of dynamics aside.

Strategies for coarse-graining have been reviewed [9–11]. Coarse-grained models can be categorized broadly into lattice-based and continuous-space. In lattice-based models, the short-range interactions are realized through a potential that reproduces single unperturbed chain rotational isomeric state (RIS) bond length and bond angle distributions when mapping onto the bond fluctuation model of Binder and collaborators [12], or via an extended RIS formalism when mapping onto the second nearest neighbor diamond lattice of Mattice and collaborators [13]. In this brief discussion we will concentrate on continuous-space coarse-grained models. In those, chains are represented

as sets of multiatom moieties, or "superatoms," connected by effective bonds. Coarse-grained intramolecular potentials have to be defined for effective bond lengths, effective bond angles, and effective torsion angles; in addition, nonbonded potentials have to be defined among the superatoms in such a way that the coarse-grained model reproduces as closely as possible the structural and thermodynamic properties of the original atomistic models. In Subsect. 3.1 below we briefly discuss the "Iterative Boltzmann Inversion" method of Müller-Plathe et al., which employs spherical superatoms. To fix ideas, we will use a recently developed model of atactic polystyrene (aPS) as an example for this method. The next Subsect. 3.2 discusses a potential pretabulation approach by Zacharopoulos et al., which employs nonspherical coarse-grained moieties.

3.1 Iterative Boltzmann Inversion

Perhaps the most important step in constructing coarse-grained models is the selection of the degrees of freedom to be used in the coarse-grained representation. Superatoms must be defined in a way that enables a drastic reduction in the number of degrees of freedom with respect to the detailed atomistic representation, with the concomitant economy in CPU time, and at the same time preserves important information about the chemical structure of the polymer. An efficient coarse-grained representation for vinyl polymers has been proposed recently by Milano and Müller-Plathe [14]. Vinyl chains other than PE possess asymmetric -CHR- groups on alternate skeletal carbon atoms. As a result, they can adopt a variety of regular or irregular stereochemical configurations. It is important that the coarse-grained model preserve information on the stereochemical configuration, as this has a direct bearing on physical properties. Given a direction of the main chain, it is possible to assign an absolute configuration to each asymmetric carbon. The stereochemical configuration of the entire chain is conveniently characterized as a sequence of diad types, each diad containing two successive asymmetric carbons. If the two asymmetric carbons in a diad have the same absolute configuration, the diad is designated as *meso*, or *m*; if they have different absolute configurations, the diad is designated as *racemo*, or *r*. In the coarse-grained model of Milano and Müller-Plathe, each diad along the backbone is considered as a superatom, with its center at the methylene carbon connecting the two asymmetric carbons of the diad. Thus, there are two kinds of superatoms: m and r. Figure 3 provides a pictorial representation of the coarse-graining scheme for a short section of atactic polystyrene (PS). An atactic vinyl chain is, to a good approximation, a Bernoullian sequence of m and r diads, with a fraction of m diads usually around 0.50. Other coarse-grained models can be conceived for PS [14]. The one shown in Fig. 3, however, is simple, can form the basis of computationally efficient simulations, and at the same time preserves the all-important information about the stereochemical sequence. With this model, 16 atoms (8 carbons and 8 hydrogens) are "condensed" into a single superatom.

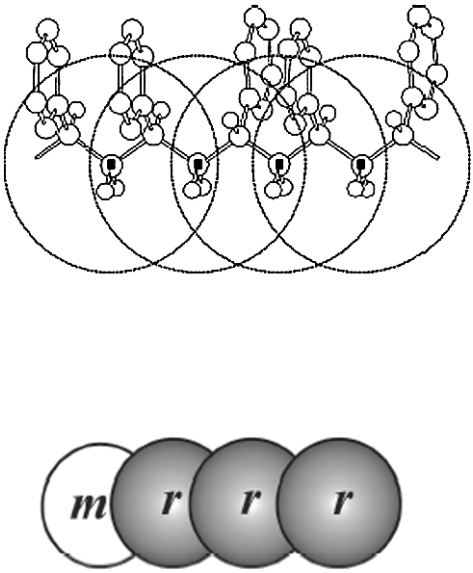


Fig. 3. Atomistic (*small spheres and bonds*) and coarse-grained (*large spheres*) representation of a short section of atactic polystyrene, with the stereochemical sequence *mrrr*, according to the coarse-graining approach of Milano and Müller-Plathe. The hydrogen atoms attached to the aromatic carbons have been omitted from the atomistic representation. (Reproduced from [14] with permission)

Having picked the coarse-grained representation, we must next define the effective potentials that go with it. In general, there are two sets of potentials we worry about: Bonded and nonbonded. It is usually a good approximation to consider the bonded force field of the coarse-grained model as consisting of additive contributions from effective bond lengths, effective bond angles, and effective torsion angles. (An effective bond connects two superatoms, and so on.) In the method of Müller-Plathe et al., coarse-grained potentials are generated by fitting certain structural features of an oligomeric analogue reference system at the temperature and pressure of interest. The reference system is subjected to both atomistic and coarse-grained simulations. Equilibration of the reference system is straightforward, due to its low molar mass. In the case of [14] the reference system was an atactic PS melt of 10-unit long chains.

There are three types of effective bonds in the coarse-grained model of Fig. 3: *m*–*m*, *m*–*r*, and *r*–*r*. A common strategy for extracting the stretching potential $U_{\ell_i}(\ell)$ for effective bonds of type i is to accumulate the normalized probability density $\rho_{\ell_i}(\ell)$ for these bonds from the atomistic simulation of the reference system, and then Boltzmann-invert it:

$$U_{\ell_i}(\ell) = -k_B T \ln\left(\rho_{\ell_i}(\ell)\right) \tag{1}$$

There are six types of effective bond angles in the coarse-grained model of Fig. 3: *m–m–m*, *m–m–r*, *m–r–m*, *m–r–r*, *r–m–r*, *r–r–r*. The associated effective bending potentials $U_{\theta_i}(\theta)$ are obtained from normalized probability densities $\rho_{\theta_i}(\theta)$ through

$$U_{\theta_i}(\theta) = -k_B T \ln\left(\rho_{\theta_i}(\theta)/sin(\theta)\right) \tag{2}$$

(The factor $\sin(\theta)$ is a Jacobian of transformation from Cartesian to generalized coordinates $\{\ell_i, \theta_i, \phi_i\}$ of the coarse-grained model.) Effective torsional potentials U_{ϕ_i} for the coarse-grained model may be accumulated in similar fashion. For the atactic PS coarse-grained model of Fig. 3, U^{ℓ_i} are Gaussians to an excellent approximation, while U_{θ_i} are representable as sums of Gaussians. U_{ϕ_i} are disregarded in [14]. Bonded effective potentials may be revised iteratively based on a comparison of probability densities from atomistic and coarse-grained simulations of the reference system, as is done for nonbonded potentials (see below). Such revision is seldom required in practice, however, due to the relatively hard nature of bonded interactions.

Deriving appropriate nonbonded potentials $U_{\mathrm{NB}}(r)$ between superatoms in the coarse-grained model is a bit more challenging. One needs one nonbonded potential $U_{\mathrm{NB}i}$ for each pair i of superatom types. In the PS example there are three such pairs: *m–m, m–r, r–r*. An initial estimate of the effective potential (potential of mean force) is obtained by accumulating the corresponding intermolecular pair distribution function $g_i^{\mathrm{target}}(r)$ from atomistic simulations of the reference system and Boltzmann inverting it:

$$U_{\mathrm{NB},i}^{(0)}(r) = -k_B T \ln\left(g_i^{\mathrm{target}}(r)\right) \tag{3}$$

If one performs a coarse-grained simulation of the reference system with bonded and nonbonded effective potentials obtained as described above, one usually observes that the intermolecular pair distribution functions $g_i^{(0)}(r)$ from the coarse-grained simulation do not quite match the atomistically computed target distribution functions $g_i^{\mathrm{target}}(r)$. To remedy this, an iterative correction procedure of the Boltzmann-inverted nonbonded effective potentials is employed:

$$U_{\mathrm{NB},i}^{(j+1)}(r) = U_{\mathrm{NB},i}^{(j)}(r) + k_B T \ln \frac{g_i^{(j)}(r)}{g_i^{\mathrm{target}}(r)} \tag{4}$$

where $j = 0, 1, \ldots$ and $g_i^{(j)}(r)$ signifies the intermolecular pair distribution function of superatom species pair i calculated by coarse-grained simulations of the reference system using the set $U_{\mathrm{NB},i}^{(j)}$ of effective nonbonded interaction potentials. This revision procedure usually converges in a few iterations. Effective nonbonded interaction potentials are usually stored as functions of distance between superatoms in tabular form. Upon convergence, they are ready for use in coarse-grained simulations of high-molecular weight systems under the conditions of interest.

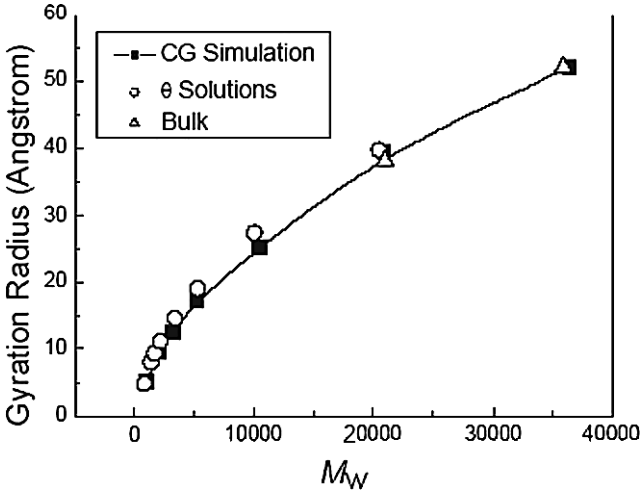


Fig. 4. Root mean squared radius of gyration $\langle R_g^2 \rangle^{1/2}$ as a function of molecular weight, as predicted from coarse-grained MD simulations of a series of monodisperse atactic PS melts at 500 K and 1 atm by Milano and Müller-Plathe. The simulation results are shown as *filled squares*. *Open triangles* are experimental data from small-angle neutron scattering (SANS) in melts. *Open circles* are experimental data from θ solutions of atactic PS. (Reproduced from [14] with permission)

One should emphasize that effective potentials of the coarse-grained model are temperature and (to a certain extent) density-dependent. Nonbonded effective potentials are designed to fit structure. They may be less than satisfactory in predicting the pressure. To bring the pressure of the coarse-grained model closer to the atomistically calculated pressure, a "ramp" potential correction which varies linearly with distance is often added a posteriori. The single parameter of the ramp is fitted to the atomistic pressure [14].

Figure 4 displays the radius of gyration $\left\langle R_g^2 \right\rangle^{1/2}$ as a function of molecular weight M_w, as determined through a series of coarse-grained MD simulations of atactic PS melts based on the model of Milano and Müller-Plathe [14]. Clearly, this important measure of overall chain conformation is predicted very well in comparison to experiment. Note that the smaller number of degrees of freedom, the smoother nature of effective potentials and the absence of fast vibrational modes afford a speedup of approximately 2000 in the coarse-grained MD relative to atomistic MD of the same melt systems [14]. This is very significant for being able to simulate long-chain systems reliably. Efficiency could be pushed even further by use of connectivity-altering MC (see Sect. 2) in conjunction with the coarse-grained model.

Figure 5 displays another interesting result from the coarse-grained MD simulations of Milano and Müller–Plathe [14]. It concerns the dependence of the chain self-diffusivity D, as extracted from the mean squared displacement

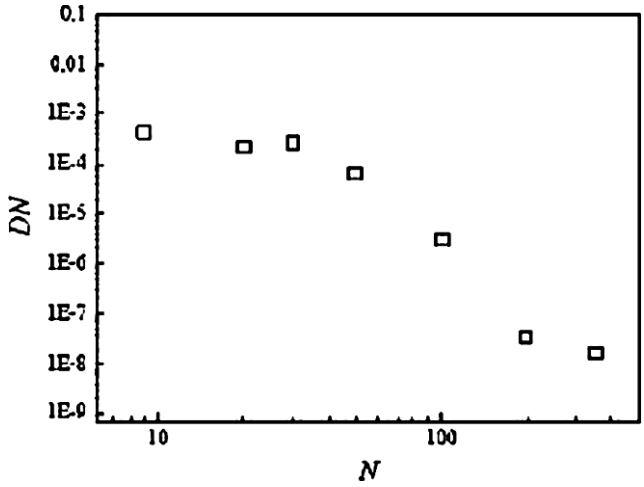


Fig. 5. Product of the chain self-diffusion coefficient D times the chain length N plotted against chain length N (in number of diads) in log-log coordinates, as obtained from the coarse-grained MD simulations of atactic PS by Milano and Müller-Plathe. (Reproduced from [14] with permission)

of chain centers of mass with elapsed time, on the chain length N (in diads). For short-chain systems, the simulations yield a dependence close to $D \propto N^{-1}$, expected from the Rouse model [15]. For longer-chain systems, a drastically different dependence close to $D \propto N^{-2}$ is seen from the coarse-grained MD, as predicted by the reptation model [15]. According to the coarse-grained MD, the crossover from Rouse to reptation behavior in the dynamics occurs around 100 diads, i.e., a molar mass of approximately 11000 g mol.$^{-1}$ This is not very far from the experimental estimate of the molar mass between entanglements for atactic PS, reported as $M_e \simeq 13000$ g mol^{-1} on the basis of the plateau modulus, [16] although the molar mass M_c where viscosity crosses over from a Rouse to a reptation-like chain length dependence is roughly double that. Thus, although absolute time scales for segmental and chain motion seem to be somewhat faster in the coarse-grained MD than in atomistic MD or in experiment, [14] it seems that the coarse-grained simulations can capture the chain constitution-dependent transition from short-chain to long-chain (entangled) terminal dynamics rather well. A further discussion of entanglements and their role in melt dynamics is provided in Sects. 4 and 5.

3.2 Coarse-Graining Using Pretabulated Potentials

Another approach to coarse-graining polymer models in continuous-space, which differs substantially from Iterative Boltzmann Inversion, has been advanced by Zacharopoulos et al. [17]. In this approach, the coarse-grained moieties which constitute a chain are not constrained to be spherical, but are

envisioned as rigid bodies of arbitrary shape. Bonded effective potentials are again calculated through Boltzmann inversion of histograms of effective bond lengths, bond angles and torsion angles accumulated via atomistic simulations of short-chain analogues or sampling of atomistic unperturbed chains in continuous space (cf. (1), (2)). Nonbonded effective potentials between pairs of coarse-grained moieties, however, are obtained via direct calculation and summation of atomistic interactions. The total interaction between two coarse-grained moieties is pretabulated as a function of the degrees of freedom specifying the relative configuration of the moieties. In the course of a coarse-grained simulation, interactions are obtained through multidimensional interpolation among the pretabulated values. In this way, there is no fitting of coarse-grained potentials to the atomistic structure. The structure and thermodynamic properties obtained from the coarse-grained model constitute direct predictions. The scheme is memory intensive and generally more complex than Iterative Boltzmann Inversion, because of the non-spherically symmetric effective nonbonded potentials it invokes. On the other hand, it is free of fitting and particularly appropriate for polymers composed of stiff, multiatom groups. Furthermore, reverse mapping into an atomistic representation is more straightforward.

Zacharopoulos et al. tested their pretabulation approach on liquid benzene at 300 K and a variety of densities. The fully flexible, atomistic COMPASS force field [18] served as a starting point for the calculations. As a first step of the coarse-graining, the actual bonded potentials are replaced by harmonic expressions where all force constants associated with the vibrational degrees of freedom go to infinity for all molecules. In the case of benzene, every molecule is ascribed a planar geometry with perfect hexagonal symmetry and bond lengths equal to their average values in an atomistic liquid simulation under the conditions of interest. Clearly, the harder the vibrational degrees of freedom, the better this approximation of making the atomistic model infinitely stiff.

In the infinitely stiff approximation, the relative configuration of two moieties depends on six degrees of freedom. As shown in Fig. 6 for two benzene molecules, these can be chosen conveniently as the distance r_{ij} between the centers of the moieties; the angles θ_i and θ_j formed between each of the vectors, $\mathbf{u}_i$, $\mathbf{u}_j$, normal to the plane of the corresponding moiety and the center-to-center vector $\mathbf{r}_{ij}$; the angles of rotation ω_i, ω_j of each of the moieties about the axes $\mathbf{u}_i$, $\mathbf{u}_j$, respectively; and the dihedral angle, ϕ_{ij}, formed by $\mathbf{u}_i$, $\mathbf{r}_{ij}$, and $\mathbf{u}_j$. Two potential pretabulation schemes were implemented for the pair of moieties: A six-dimensional one, in which the atomistic interaction energy is calculated and stored at the nodes of a six-dimensional grid in the variables r_{ij}, θ_i, θ_j, ω_i, ω_j, and ϕ_{ij}; and a four-dimensional one, in which the atomistic interaction energy is pre-averaged with respect to ω_i and ω_j and stored as a function of the remaining variables r_{ij}, θ_i, θ_j, and ϕ_{ij}. In the four-dimensional case the effective nonbonded potential is obtained as

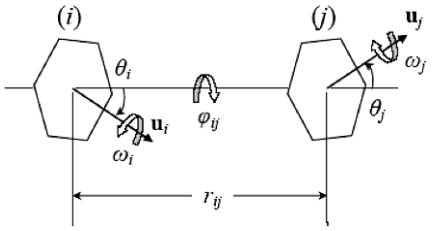


Fig. 6. Sketch of the six degrees of freedom needed to specify the relative configuration of two inflexible moieties, using two benzene molecules as an example. (Reproduced from [17], with permission)

$$U_{\mathrm{mf},ij}(r_{ij},\theta_i,\theta_j,\phi_{ij}) = -\frac{1}{\beta}\ln\left\{\frac{9}{\pi^2}\int_0^{\pi/3} d\omega_i \int_0^{\pi/3} d\omega_j \exp\left[-\beta U_{ij}^0(r_{ij},\theta_i,\theta_j,\phi_{ij},\omega_i,\omega_j)\right]\right\} \quad (5)$$

where U_{ij}^0 is the potential energy according to the atomistic (COMPASS) force field between the two moieties in their stiff, ground state geometries. Averaging over ω_i and ω_j is performed numerically, using a discrete set of nodal points along the ω_i and ω_j directions.

In the course of coarse-grained simulations, the interaction energy between each pair of coarse-grained moieties is obtained through linear interpolation in six (respectively, four) dimensions of the values in the corresponding table. The four-dimensional scheme allows higher resolution in the tabulation for fixed memory size; on the other hand, it is subject to the preaveraging approximation. Several discretizations were used in the tabulation and in the numerical integration for obtaining $U_{\mathrm{mf},ij}$, (5), to assess the effect of table resolution on the results [17].

The results from coarse-grained MC simulations of liquid benzene based on the six-dimensional and the four-dimensional pretabulation schemes, with various discretizations, were compared against fully atomistic simulations of the same systems at the same thermodynamic state points using the COMPASS force field [17]. Here we discuss some of these comparisons based on the four-dimensional pretabulation scheme.

Figure 7 shows the pair distribution function $g(r)$ between centers of mass of benzene molecules in room-temperature liquid benzene, as computed from the four-dimensional interpolation scheme and from fully atomistic MD based on the COMPASS force field. The most probable distance between centers of benzene molecules is around 5.6 Å. The coarse-grained 4D-MC simulation gives an excellent prediction of $g(r)$ under the considered conditions. Figure 8 displays another aspect of liquid structure: orientational correlations between molecules. These have been assessed through the second- order Legendre poly-

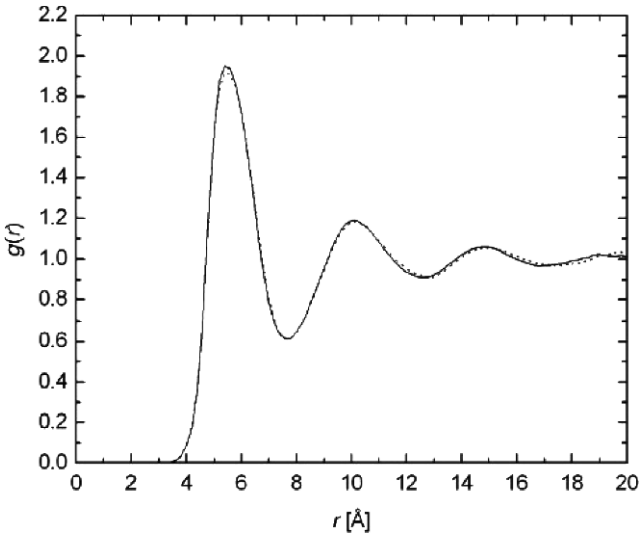


Fig. 7. Center-of-mass pair distribution functions of liquid benzene at temperature $T = 300$ K and mass density $\rho = 0.870$ g cm^{-3}, as obtained from fully atomistic MD (*solid line*) and from coarse-grained MC based on the four-dimensional pretabulation scheme of Zacharopoulos et al. (*dotted line*). The intervals used in constructing the four-dimensional table were $\Delta\phi = 10°$, $\Delta\theta_i = \Delta\theta_j = 5°$, $\Delta r = 0.2$ Å. The minimum and maximum values of r in the tabulation were 2 Å and 13 Å, respectively. Integrations for preaveraging over ω_i and ω_j were performed using $\Delta\omega_i = \Delta\omega_j = 10°$. (Reproduced from [17], with permission)

nomial $P_2(r)$ of the cosine of the angle between unit normal vectors $\mathbf{u}_i$, $\mathbf{u}_j$, averaged over all molecules at given center-to-center distance r:

$$P_2(r) = \frac{1}{2}\left[3\left\langle \cos^2(\mathbf{u}_i \cdot \mathbf{u}_j)\right\rangle_{r_{ij}=r} - 1\right] \tag{6}$$

There is clearly a trend towards orthogonal (T-shaped) configurations at the most common separation distance; this is well documented in the literature and attributed to the quadrupole moments of the molecules. The agreement between coarse-grained MC and atomistic MD is excellent. Some differences can be seen at the distances of closest approach, which, however, are not probable (see Fig. 7). MC can achieve near parallelity between molecules at these distances, because of the perfectly planar shape it postulates for the molecules.

Figure 9 assesses the ability of coarse-grained MC based on the four- dimensional pretabulation scheme to predict atomistically computed pressures. Capturing the equation-state behavior is generally a serious challenge for coarse-grained simulations. Coarse-grained results are shown for five different discretization levels. For all sets, $\Delta\omega_i = \Delta\omega_j = 10°$ was used in averaging over the in-plane degrees of freedom. Clearly, as the table becomes finer, the

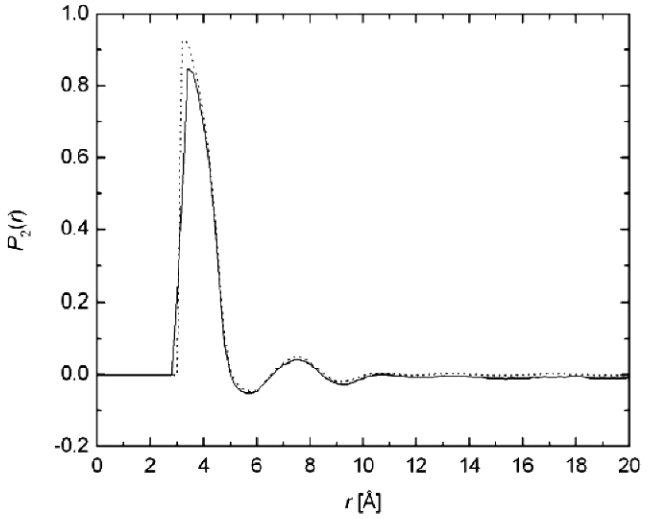


Fig. 8. Orientational correlation function $P_2(r)$ between pairs of benzene molecules in a liquid phase at $T = 300$ K and $\rho = 0.870$ g cm^{-3}, as a function of distance r between the centers of the molecules. Results from atomistic MD (*solid line*) are shown along with results from coarse-grained MC (*dotted line*) using a four-dimensional potential pretabulation scheme with $\Delta\phi = 10°$, $\Delta\theta_i = \Delta\theta_j = 5°$, $\Delta r = 0.2$ Å. (Reproduced from [17], with permission)

coarse-grained results approach those of the atomistic MD. Furthermore, the MC simulations reach a limit for each suite of discretization of the rotational degrees of freedom: finer discretization along r_{ij} has no noticeable effect on the results. At a discretization level of 10° for ϕ_{ij}, 5° for θ_i, θ_j and 0.05 Å for r_{ij}, the coarse-grained and atomistic pressures practically coincide.

Another thermodynamic property that can be compared between coarse-grained and atomistic simulations is the cohesive energy density. At the orthobaric density of 0.870 g cm^{-3} at 300 K, the four-dimensional coarse-grained scheme at finest discretization gives 7.665 ± 0.004 kcal cm,$^{-3}$ which is 6% lower than the atomistic value of 8.152 ± 0.029 kcal cm.$^{-3}$ (The experimental cohesive energy density of benzene under these conditions is 8.35 kcal cm.$^{-3}$)

Based on the comparisons presented above, one can conclude that the four-dimensional pretabulation scheme proposed by Zacharopoulos et al. [17] performs very well in capturing structural and thermodynamic properties, without adjustable parameters. One should bear in mind that the computer time of coarse-grained simulations was less than $1/15^{\text{th}}$ the time required for the fully atomistic MD.

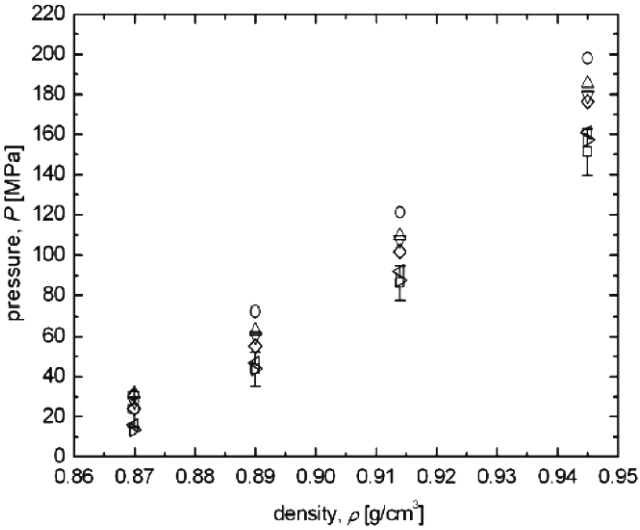


Fig. 9. Pressure as a function of density in liquid benzene at 300 K. Results from the coarse-grained simulations based on the four-dimensional pretabulation scheme are shown for $\Delta\theta_i = \Delta\theta_j = 10^\circ$ with $\Delta r = 0.2$ Å (*circles*), 0.1 Å (*triangles*), 0.05 Å (*inverted triangles*), and for $\Delta\theta_i = \Delta\theta_j = 5^\circ$ with $\Delta r = 0.2$ Å (*diamonds*), 0.1 Å (*triangles pointing to the left*), 0.05 Å (*triangles pointing to the right*). In all cases, $\Delta\phi = 10^\circ$. Atomistic MD averages are shown as squares. Standard deviations are shown for the MD. Error bars for the coarse-grained MC are within symbol size. (Reproduced with permission from [17])

4 Chain Entanglements in Polymer Melts

The rheological properties of long-chain polymer melts are both fascinating and extremely important for the design of processing operations and products made of these materials [19]. For monodisperse melts of linear flexible chains, rheological properties depend sensitively on molecular weight. For example, the zero-shear rate viscosity, η_0, scales roughly as N^1 in melts of chains which are shorter than a critical molar mass M_c, but still long enough for their end-to-end vector to follow a Gaussian distribution. For melts of molar mass substantially exceeding M_c, η_0 exhibits a much stronger dependence on chain length, $\eta_0 \propto N^{3.4}$. Likewise, the self-diffusivity D of chains scales roughly as N^{-1} for short-chain melts, but as $N^{-2.2}$ for long-chain melts (compare Fig. 5). All viscoelastic properties in the linear regime can be expressed in terms of the shear stress relaxation modulus, $G(t)$, i.e., the shear stress with which the material responds at time t to the imposition of a small step shear strain at time 0, divided by the magnitude of the step strain [20]. $\log G(t)$ plotted as a function of $\log t$ exhibits a sharp, more or less molar mass-independent drop at short times, which reflects segmental relaxation processes. For short-chain melts, $G(t)$ continues dropping steadily in the terminal relaxation region.

For long-chain melts, however, $G(t)$ exhibits a more or less chain length-independent plateau value G_N^0 over a broad range of time scales, extending out to the (strongly molar mass-dependent) longest relaxation time, beyond which the terminal drop occurs. Over the plateau region, the mechanical response of the long-chain melt is comparable to that of a crosslinked rubber.

A successful conceptual framework for understanding these phenomena is provided by the Rouse and reptation models [15]. The Rouse model, appropriate for melts of short (but still Gaussian) chains, views a chain as a string of Brownian particles connected by harmonic springs and moving in a viscous medium, which represents the rest of the polymer. The Rouse model predicts $\eta_0 \propto N^1$ and $D \propto N^{-1}$. It expresses viscoelastic properties in terms of two molecular parameters, which can be chosen as the Kuhn length b and the friction factor per Kuhn segment, ζ; the latter sets the time scale for the dynamics. In longer-chain melts, the entanglement tube model postulates that the mutual uncrossability of chains generates topological constraints, referred to as entanglements, which effectively restrict individual chain conformation in a curvilinear tube-like region enclosing each chain. Lateral chain motion is confined to the length scale of the tube diameter. Large-scale motion is promoted via reptation, an effective one-dimensional diffusion of the chain along the tube axis. The tube axis provides a coarse-grained representation that characterizes the chain topology and is called the Primitive Path (PP). An additional important molecular parameter in the reptation regime is the tube diameter, d. The primitive path is envisioned as a random walk of step size d. The picture of reptative motion of a PP of fixed length in a fixed set of obstacles leads to $\eta_0 \propto N^3$ and $D \propto N^{-2}$. More refined reptation-based approaches, which take into account PP contour length fluctuations, constraint release due to the motion of surrounding chains, and longitudinal stress relaxation along the tube [21] capture very well the scaling behavior seen in experiments.

In the plateau region of $G(t)$ the melt exhibits the response of a rubber because chains do not have the time to disengage from their entanglements and thus entanglements act as effective crosslinks. Using the Doi-Edwards theory, one can arrive at an estimate of the mean molar mass between entanglement points, M_e, based on the plateau modulus G_N^0 and the mass density ρ [15, 16, 22]. The tube diameter d is often viewed as the root mean square end-to-end distance of a chain of molar mass M_e:

$$M_e = \frac{4}{5}\frac{\rho RT}{G_N^0} \tag{7}$$

An alternative measure of M_e can be obtained by comparing the single-chain intermediate coherent dynamic structure factor obtained from Neutron Spin Echo experiments against the expressions given for this quantity by reptation theory [23]. Inevitably, estimates of M_e are colored by the approximations in the theory invoked to intepret experimental data. Also, estimates of

M_e are considerably different from M_c, M_e being roughly half M_c for most flexible polymers [20].

Not only do entanglements dominate the rheological behavior of long-chain polymers in the melt state; they are also intimately related to the strain hardening and ultimate mechanical properties exhibited by these materials in the glassy state [24].

From the above discussion it is clear that being able to relate d or M_e to chemical constitution is a highly worthwhile objective. A difficulty one faces when trying to do this is that entanglements are not directly observable experimentally. Their precise microscopic definition is somewhat elusive. Nevertheless, a number of theoretical treatments have been advanced for the estimation of M_e or d from chemical constitution. We briefly present a simple argument by Fetters et al. [16], which leads to a particularly successful correlation between polymer molecular weight, density, chain dimensions, and melt viscoelastic properties. The basis of the argument lies in the idea, advanced some time ago by Kavassalis and Noolandi, [25] that, if the number of chains visiting the volume that is "pervaded" by a given chain exceeds a certain value, then the chain is entangled with its surroundings.

Envision a randomly coiled chain of molar mass M in a melt. The volume taken up (occupied) by the chain segments is M/ρ, where ρ is the mass density of the melt. A different (much larger) volume is the volume "pervaded," or spanned, by the chain, V_{sp}. V_{sp} can be defined as the volume of the smallest sphere in space which completely contains all segments of the chain. It can be estimated from the chain mean square radius of gyration as

$$V_{sp} = A \left\langle R_g^2 \right\rangle^{3/2} \tag{8}$$

where A is a numerical constant of order unity.

$\left\langle R_g^2 \right\rangle$ is a strongly molar-mass dependent quantity. We can write

$$\left\langle R_g^2 \right\rangle \simeq \frac{1}{6} \left\langle R^2 \right\rangle = \frac{1}{6} M \frac{\left\langle R^2 \right\rangle}{M} \tag{9}$$

where the last separation is intended to emphasize that, for melt chains, $\frac{\langle R^2 \rangle}{M}$ is a chain length-independent measure of conformational stiffness (compare Fig. 2).

How many chains (including the considered one) does one find in the volume V_{sp}? A back-of-the-envelope calculation of this number, n, gives:

$$n = V_{sp}\rho/(M/N_A) = A/6^{3/2} M^{3/2} \left(\frac{\left\langle R^2 \right\rangle}{M} \right)^{3/2} \frac{\rho N_A}{M} \tag{10}$$

where N_A is Avogadro's number. Furthermore,

$$n = A/6^{3/2} \rho \left(\frac{M}{N_A} \right)^{1/2} \left(\frac{N_A \left\langle R^2 \right\rangle}{M} \right)^{3/2} \tag{11}$$

The quantity

$$p = \frac{M}{\rho N_{\mathrm{A}} \langle R^2 \rangle} = \frac{v}{b^2} \tag{12}$$

with v being the volume occupied by a Kuhn segment, is a chain length-independent quantity with dimensions of length. p is large for "fat", flexible and small for thin, stiff chains. It plays a significant role in theories of dynamics and miscibility, and is termed the "packing length."

In terms of the packing length, (11) is written:

$$n = A/6^{3/2} \left(\frac{M}{N_{\mathrm{A}} \rho} \right)^{1/2} p^{-3/2} \tag{13}$$

The idea is that, when n reaches a critical value n_{e}, M becomes M_{e}, the molar mass between entanglements. This implies

$$M_{\mathrm{e}} = 6^3/A^2 n_{\mathrm{e}}^2 N_{\mathrm{A}} \rho p^3 = C N_{\mathrm{A}} \rho p^3 \tag{14}$$

with C being a universal constant. Comparison of (14) against experimental M_{e} data from the plateau modulus of a variety of linear chain melts has led to a very successful correlation [16] with $C \simeq 370$. In the calculation of Fetters et al. [16] n_{e} is assigned a value of 2, implying that a chain of molar mass M_{e} has exactly one other chain of the same molar mass (with which it entangles) in the volume it pervades. For this value of n_{e}, $A = 1.518$.

Interpreting the tube diameter d as the root mean square end-to-end distance of a chain of molar mass M_{e},

$$\frac{d^2}{M_{\mathrm{e}}} = \frac{\langle R^2 \rangle}{M} \ , \tag{15}$$

which becomes

$$d = C^{1/2} p \tag{16}$$

leading to the rule of thumb that the entanglement tube diameter is roughly 19 times the packing length.

The analysis of simulation results from coarse-grained model melts with a wide variety of p values has yielded results which conform to (14) [26].

5 Topological Analysis of Atomistic Melt Configurations and Reduction to Entanglement Networks

Given the importance of chain entanglements for the physical properties of polymers, it is understandable that a sizeable body of simulation work has been directed towards investigations of entanglements and entangled dynamics. In this section we will focus on recent simulation work on entanglements

based on atomistic (more precisely, united atom) models of polyolefins and polydienes.

Perhaps the simplest way to detect and analyze entangled dynamical behavior in a melt is to conduct equilibrium MD simulations starting from configurations that have been thoroughly equilibrated, for example using the connectivity-altering MC schemes of Sect. 2. Simulation times up to a few μs are accessible on present-day parallel clusters using efficient multiple time step integration and domain decomposition algorithms [1]. This is sufficient for detecting all regimes of dynamical behavior in melts of up to C_{500} chains. A useful quantity for detecting the crossover from Rouse to reptation behavior upon increasing N is the mean square displacement of monomers lying sufficienty far from the chain ends, $\phi(t)$ [15,27]. The self-diffusivity D can readily be extracted from the slope of the mean square displacement of the chain center of mass with respect to time, provided the terminal linear (Einstein) regime has been reached. $G(t)$ could, in principle, be calculated from the time autocorrelation function of instantaneous shear stress. The time integral of $G(t)$ would give η_0, according to the Green-Kubo relation for viscosity.

In practice, simulation times are not adequately long for accumulating $G(t)$ with acceptable statistics. An alternative strategy is to interpret the MD simulation results in the light of a mesoscopic theoretical model of dynamics, such as the Rouse and the Doi-Edwards reptation models and their more refined variants. Harmandaris et al. [28] mapped MD trajectories of short-chain PE melts onto the Rouse model. They showed that a consistent value of ζ can be extracted from the chain self- diffusivities and the decay of the first few Rouse modes. This ζ value is independent of chain length in the interval C_{80} – C_{150} and provides very good estimates of the linear viscoelastic properties of these systems when used in conjunction with the Rouse model. Karayiannis and Mavantzas [29] accumulated mean square displacements $\phi(t)$ of internal segments as functions of time from atomistic MD simulations of PE melts up to C_{500} up to times of 4 μs. Invoking the reptation theory of Likhtman and McLeish, [21] they estimated the entanglement tube diameter from the ordinate of the point where the slope of the $\log\phi$ versus $\log t$ plot crosses over from 1/2 to 1/4, reflecting onset of confinement by the tube. This mapping onto the Likhtman-McLeish model yielded $d = 32 \pm 2$ Å, in reasonable agreement with the value 38 Å reported on the basis of plateau modulus measurements [16].

A more attractive possibility for learning about entanglements and entangled dynamics from atomistic simulations is to take the large set of uncorrelated configurations generated, e.g., through connectivity-altering Monte Carlo and analyze them topologically. This calculation circumvents the need to undertake long MD runs. Apart from giving us a detailed picture of the entanglement structure, topological analysis reduces every atomistic melt configuration into a mesoscopic entanglement network. Such networks are very useful in their own right as starting points for simulations of flow under prescribed boundary and initial conditions (see, for example, [30]) Up to now,

initial network configurations were generated in a more or less ad hoc fashion, consistent with specified moments of the distribution of interentanglement lengths, such as M_e. Now one has, for the first time, the possibility to derive these configurations from atomistics. Thus, topological analysis is a tool for generating coarse-grained configurations of polymers at a level which is considerably less detailed than the examples discussed in Sect. 3, but perfectly appropriate for the prediction of rheological properties of melts as functions of chain length distribution and architecture.

Our strategy for reducing atomistic configurations into entanglement networks [31] invokes Edwards's definition of the PP [32]. Edwards regarded the PP as the shortest path constructed by keeping the chain ends fixed while continuously tightening (shrinking) the chain contour, so that the resulting path has the same topology relative to other chains as the chain itself. The algorithm we devised, which we refer to as CReTA (Contour Reduction Topological Analysis) reduces simultaneously the contour length of all chains by shrinking them until topological constraints block further contour reduction. CReTA implements random aligning string moves and hardcore interactions. When chain contour lengths are no longer diminishing, chain thickness is reduced by reducing the hard-sphere diameter and the process starts anew. This helps tighten meshed unentangled loops, which, although temporarily blocked, do not represent true topological constraints. Upon decreasing the hard-sphere diameter, to preserve the fused sphere sequence of the chain, we place an auxiliary atom between successive skeletal atoms, that have lost contact. The whole procedure terminates when a hard sphere diameter of 0.5 Å is reached. Further reduction of the diameter to attain the infinitely thin, continuous line limit would be time consuming and superfluous. The final entanglement structure is remarkably insensitive to the order in which chains are reduced. Our use of hard sphere interactions and random moves differentiates the algorithm from another recently proposed algorithm [26], where contour reduction is achieved through minimization of an elastic energy. Slippage of chains whenever they touch each other during contour shrinkage is eliminated in our new algorithm.

Figure 10 displays an example of an atomistic long-chain PE configuration and its reduction to an entanglement network using the CReTA algorithm. In the reduced picture one can clearly see that the lengths of interentanglement strands, N_{ES}, measured in skeletal carbons, are broadly distributed.

Having a large number of equilibrated, topologically uncorrelated configurations of the atomistic model from the connectivity-altering MC, one can perform thorough statistical analyses of the reduced networks. Figure 11 shows the normalized probability densities of interentanglement strand lengths, N_{ES}, as obtained from three model polymer melts [31]: A PE of mean chain length C_{500} and polydispersity index 1.083 at 450 K and 1 atm; a monodisperse C_{1000} PE at 450 K and 1 atm; and a C_{1000} cis-1,4 polybutadiene (PB) with polydispersity index 1.053 at 413 K and 1 atm. N_{ES} has been reduced by its average value, $\bar{N}_{ES}$ in the abscissa. Remarkably, the three distributions coincide when

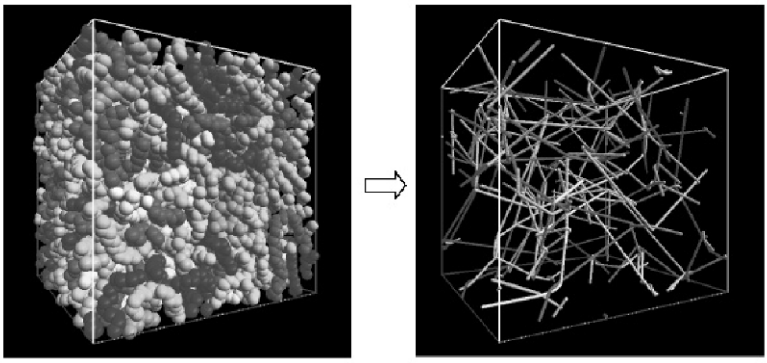


Fig. 10. Atomistic configuration of a C1000 melt at 450 K and 1 atm (*left*) and corresponding reduced network obtained by application of the CReTA algorithm (*right*). Chain ends, which remain fixed during the reduction process, are shown as white spheres on the left

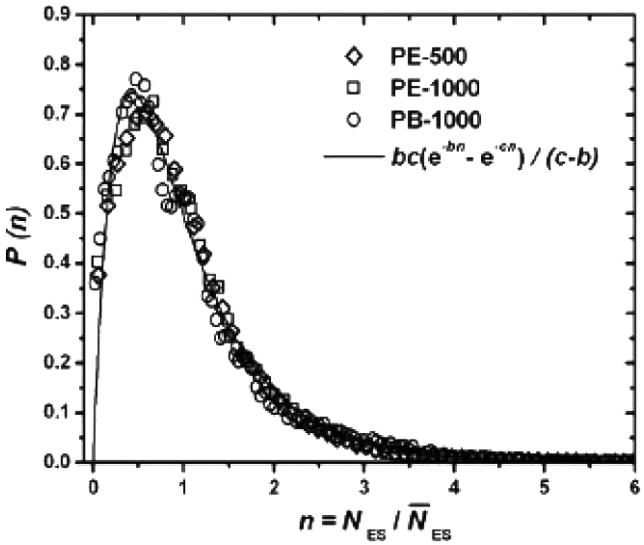


Fig. 11. Distribution of interentanglement strand lengths N_{ES} in two melts of polyethylene and one of cis-1,4 polybutadiene, as obtained by applying CReTA to large ensembles of well-equilibrated configurations. Strand lengths, measured in carbon atoms, have been reduced by their mean values in each polymer. All distributions are normalized

plotted in this way. Clearly, the reduced distribution remains invariant with molar mass and is common for two rather different chemical constitutions. The distribution possesses an exponentially decaying tail. A Poisson process of starting from one end of a chain, moving towards the other end, and laying down entanglements in uncorrelated fashion with prescribed frequency would

Table 1. Characteristics of Entanglement Networks from Reduction of Atomistic Configurations

System	N	$\bar{N}_{\mathrm{ES}}$	$\bar{d}_{\mathrm{ES}}$(Å)	N_{e}	d(Å)	p(Å)
PE	500	28.3	14.0	75.1(61.4)	38.4(38.5)	1.53(1.69)
PE	1000	29.1	14.1	74.2(61.4)	36.6(38.5)	1.65(1.69)
PB	1000	80.9	18.7	178.7(173.8)	42.3(43.0)	2.59(2.44)

yield an exponential distribution for N_{ES}. Such a distribution was predicted by Schieber [33] with simple arguments. Our atomistic simulations indicate that the distribution of $n = N_{\mathrm{ES}}/\bar{N}_{\mathrm{ES}}$ is not quite exponential. The data from both materials in Fig. 11 fall on a master curve of the form

$$P(n) = \frac{bc}{c-b}\left(e^{-bn} - e^{-cn}\right) \tag{17}$$

with $b = 1.30$ and $c = 3.78$.

A stochastic interpretation of (17) can be given in terms of a renewal process, which is the generalization of a Poisson process. $P(n)$ is the convolution of two exponential distributions, be^{-bn} and ce^{-cn}. It can be interpreted as the result of two uncorrelated alternating Poisson processes with rates b, c, evolving on the monomer sequence space. The first process, with rate c, stochastically creates an unentangled monomer sequence in front of an entanglement; in other words, it accounts for the fact that a new entanglement cannot follow immediately an existing entanglement, but there is an effective repulsion between successive entanglements. The second process, with rate b, takes over when an unentangled sequence has been created and places the next entanglement on one of the monomers following the unentangled sequence.

Table 1 shows some average values characterizing the reduced networks obtained from the three atomistic melts. $\bar{N}_{\mathrm{ES}}$ and $\bar{d}_{\mathrm{ES}}$ are the mean number of carbon atoms and the mean distance in Å, respectively, between connected entanglement points. N_{e} and d are measures of conformational stiffness of the reduced chains (PP), extracted by mapping the reduced chains to Kuhn chains. N_{e} is the number of monomer units in a Kuhn segment of a reduced chain and d is the Kuhn segment length of the reduced chain. The values in parentheses are experimental, from [16]. They show the chain lengths between entanglements and the tube diameters, as estimated from plateau moduli (7). The last column in the Table gives packing lengths p, as calculated from the simulations and as obtained experimentally. Agreement between simulated and experimental p's is excellent, reflecting the good predictions of volumetric properties and end-to-end distance afforded by the models.

Several important conclusions can be drawn on the basis of Table 1. First of all, network properties are chain length-independent for the N range considered here. Secondly, the network natural mesh size, as quantified by $\bar{N}_{\mathrm{ES}}$ and $\bar{d}_{\mathrm{ES}}$, is clearly smaller than the corresponding PP Kuhn segment quantities

N_e, d. There are directional correlations between successive strands along the same reduced chain, which die exponentially with strand separation. In other words, PPs are not random walks on the length scale of the mesh size of the network and $\bar{N}_{ES} \simeq 0.4N_e$. Interestingly, in recent [30] Brownian dynamics simulations of three-dimensional entangled networks of primitive paths, the linear viscoelastic response of PB solutions is reproduced quantitatively only when an N_e of about half the experimental value is utilized. In all simulated polymers of Table 1, the values of N_e and d extracted from the Kuhn length of the primitive path agree very well with experimental values extracted from the plateau modulus.

The structure of the "gas" of entanglement points in space was analyzed and found to correspond to an ideal dilute gas for distances exceeding $0.8\bar{d}_{ES}$ [31]. Conributions to the pair distribution function from pairs of neighboring entanglement points sharing one or two "parent chains" were analyzed separately. The statistical information accumulated from this study can be used to reconstruct mesoscopic entanglement networks consistent with the atomistic picture. As already mentioned, such networks are excellent starting points for the simulation of flow and large-scale deformation.

6 Concluding Remarks

In this chapter we discussed some recent accomplishments in the equilibration of atomistic polymer models, coarse-graining of these models into models involving fewer degrees of freedom, and prediction of physical properties through simulation.

We pointed out that the multiplicity of length and time scales calls for hierarchical approaches to polymer modeling and simulation, consisting of many levels, each level addressing phenomena over a specific window of length and time scales.

At the atomistic level, we have seen that a new class of connectivity-altering Monte Carlo algorithms enable the equilibration of long-chain polymer melts, comparable to those encountered in plastics processing operations, at all length scales. When used in conjunction with force fields that have been calibrated carefully against structural and thermodynamic properties of small-molar mass analogues, these algorithms yield excellent predictions of chain packing, conformation, volumetric and thermal properties.

Coarse-graining is valuable in addressing the properties of polymers of complex chemical constitution. We have discussed two coarse-graining strategies in continuous space. In the first, the coarse-grained model is constructed of spherical "superatoms." Effective bond length, bond angle, and torsion angle potentials are obtained through Boltzmann inversion of the corresponding distributions, accumulated in the course of atomistic simulations of oligomers. Nonbonded interactions, on the other hand, are obtained through iterative Boltzmann inversion of the correlation functions accumulated in the course

the oligomer atomistic simulations. As an example, we have seen an efficient strategy proposed by Milano and Müller-Plathe for coarse-graining vinyl polymers of various tacticities and its application to atactic polystyrene. A second coarse-graining strategy of nonbonded interactions by Zacharopoulos et al. involves the definition of inflexible multiatom moieties of arbitary shape and pretabulation of the interactions between pairs of such moieties as functions of four or six degrees of freedom specifying the relative configuration of the moieties. In the course of a coarse-grained simulation, interactions are computed by interpolation among the pretabulated values. As an example we saw a four-dimensional pretabulation/interpolation scheme for liquid benzene. When the tables are sufficiently fine, the scheme captures pair distribution functions, orientational correlation functions, pressure, and cohesive energy density in very good agreement with atomistic simulations with no adjustable parameters, at $1/15^{\text{th}}$ of the computational cost of atomistic simulation.

We briefly discussed the importance of entanglements for the rheological and terminal mechanical properties of long-chain polymers and presented a correlation between entanglement molar mass, mass density, and packing length. We saw that the crossover from Rouse to entangled dynamics can readily be detected in equilibrium MD simulations of long-chain melts and discussed strategies for extracting the segmental friction factor and the entanglement tube diameter by mapping atomistic simulation results onto state-of-the-art models of Rouse and reptation dynamics. Finally, we showed how topological analysis of atomistic configurations can be used to reduce these configurations into networks of entanglement points. This level of coarse-grained representation is valuable as a starting point for mesoscopic entanglement-based simulations of flow and deformation. Important conclusions from our topological analysis were (a) that the strand length (in monomer units) between connected entanglements follows a broad distribution that can be interpreted as a convolution of two Poisson processes: one of blocking the region along the contour immediately following an entanglement from receiving another entanglement, and a second one of laying down new entanglements in regions that have been freed of the first process; (b) that primitive paths have some stiffness, the mean number of monomers between entanglements (mesh size of the network) being only 0.4 of the Kuhn length of the primitive path; (c) that the Kuhn length of the primitive path, as extracted from the topological analysis, is remarkably close to estimates of the tube diameter from the plateau modulus, in both polymer melts examined (polyethylene and cis-1,4 polybutadiene).

Acknowledgements

I wish to thank my collaborators Prof. Vlasis Mavrantzas, Dr. Christos Tzoumanekas, Dr. Nikos Karayiannis, Dr. Vagelis Harmandaris, Dr. Nikos Zacharopoulos, Ms. Niki Vergadou, and Dr. Alfred Uhlherr (CSIRO, Aus-

tralia) for their important contributions to this work. I am grateful to Dr. Pino Milano and Prof. Florian Müller-Plathe for making their manuscript on polystyrene available to me before publication. Financial support from the Dutch Polymer Institute (project 430), from the European Commission (Brite-EuRam program MPFLOW, Growth program DEFSAM), and from the Greek Secretariat of Research and Technology (PENED projects 218-95 EΔ, 95-99 EΔ) is gratefully acknowledged.

References

1. M. J. Kotelyanskii and D. N. Theodorou, Eds. (2004) *Simulation Methods for Polymers*. Marcel Dekker, New York
2. D. N. Theodorou (2004) Understanding and predicting structure-property relations in polymeric materials through molecular simulations. *Mol. Phys.* **102**, pp. 147–166
3. L. R. Dodd and D. N. Theodorou (1994) Atomistic Monte Carlo Simulation and Continuum Mean Field Theory of the Structure and Equation of State Properties of Alkane and Polymer Melts. In *Atomistic Modeling of Physical Properties*, Eds. L. M. Monnerie and U. W. Suter, *Adv. Polym. Sci.* **116**, pp. 249–281 Springer-Verlag, Berlin
4. R. Auhl, R. Everaers, G. S. Grest, K. Kremer, and S. J. Plimpton (2003) Equilibration of long chain polymer melts in computer simulations. *J. Chem. Phys.* **119**, pp. 12718–12728
5. D. N. Theodorou (2002) Variable Connectivity Monte Carlo Algorithms for the Atomistic Simulation of Long-Chain Polymer Systems. In *Bridging Time Scales: Molecular Simulations for the Next Decade*, Eds. P. Nielaba, M. Mareschal, and G. Ciccotti, pp. 69–128 Springer-Verlag, Berlin
6. L. Peristeras, I. Economou, and D. N. Theodorou (2005) Structure and volumetric properties of linear and triarm star polyethylenes from atomistic Monte Carlo simulation using new internal rearrangement moves. *Macromolecules* **38**, pp. 386–397
7. A. Uhlherr, M. Doxastakis, V. G. Mavrantzas, D. N. Theodorou, S. J. Leak, N. E. Adam, and P. E. Nyberg (2002) Atomic structure of a high polymer melt. *Europhys. Lett.* **57**, pp. 506–511
8. M. P. Allen and D. J. Tildesley (1987) *Computer Simulation of Liquids*. Oxford University Press, Oxford
9. J. Baschnagel, K. Binder, P. Doruker, A. A. Gusev, O. Hahn, K. Kremer, W. L. Mattice, F. Müller-Plathe, M. Murat, W. Paul, S. Santos, U. W. Suter, and V. Tries (2000) Bridging the gap between atomistic and coarse-grained models of polymers: status and perspectives. *Adv. Polym. Sci.* **152**, pp. 41–156
10. F. Müller-Plathe (2002) Coarse-graining in polymer simulation: From the atomistic to the mesoscopic scale and back. *Chem. Phys. Phys. Chem.* **3**, pp. 754–769
11. F. Müller-Plathe (2003) Scale-hopping in computer simulations of polymers. *Soft Mater.* **1**, pp. 1–31
12. W. Paul, K. Binder, K. Kremer, and D. W. Heermann (1991) Stucture property correlations in polymers: A Monte Carlo approach. *Macromolecules* **24**, pp. 6332–6334

13. P. Doruker and W. L. Mattice (1997) Reverse mapping of coarse-grained polyethylene chains from the second nearest-neighbor diamond lattice to an atomistic model in continuous space. *Macromolecules* **30**, pp. 5520–5526
14. G. Milano and F. Müller-Plathe (2005) Mapping atomistic simulations to mesoscopic models: A systematic coarse-graining procedure for vinyl polymer chains. *J. Phys. Chem. B* **109**, pp. 18609–18619
15. M. Doi and S. F. Edwards (1986) *The Theory of Polymer Dynamics*. Clarendon, Oxford
16. L. J. Fetters, D. J. Lohse, D. Richter, T. A. Witten, and A. Zirkel (1994) Melt-chain polymer chain dimensions as functions of temperature. *Macromolecules* **27**, pp. 4639–4647
17. N. Zacharopoulos, N. Vergadou, and D. N. Theodorou (2005) Coarse-graining using pre-tabulated potentials: Liquid benzene. *J. Chem. Phys. B* **122**, 244111
18. H. Sun (1998) COMPASS: An ab initio force-field optimized for classical condensed-phase applications: Overview on details with alkane and benzene compounds. *J. Phys. Chem.* **102**, pp. 7338–7364
19. R. B. Bird, R. C. Armstrong, and O. Hassager (1977) *Dynamics of Polymeric Liquids*. vol. I John Wiley, New York
20. W. Graessley (1993) Viscoelasticity and Flow in Polymer Melts and Concentrated Solutions. In *Physical Properties of Polymers*, Eds. J. E. Mark, A. Eisenberg, W. W. Graessley, L. Mandelkern, E. T. Samulski, J. L. Koenig, and G. D. Wignall 2nd Edition, ACS, Washington D.C.
21. A. E. Likhtman and T. M. McLeish (2002) Quantitative theory for linear dynamics of linear entangled polymers. *Macromolecules* **35**, pp. 6332–6343
22. L. J. Fetters, D. J. Lohse, and W. W. Graessley (1999) Chain dimensions and entanglement spacings in dense macromolecular systems. *J. Polym. Sci. Part B: Polym. Phys.* **37**, pp. 1023–1033
23. A. Wischnewski, M. Monkenbusch, L. Willner, D. Richter, A. E. Likhtman, T. M. McLeish, and B. Farago (2002) Molecular observation of contour-length fluctuations limiting topological confinement in polymer melts. *Phys. Rev. Lett.* **88**, 058301
24. J. J. Benkoski, G. H. Fredrickson, and E. J. Kramer (2002) Model for the fracture energy of glassy polymer-polymer interfaces. *J. Polym. Sci. Part B: Polym. Phys.* **40**, pp. 2377–2386
25. T. A. Kavassalis and J. Noolandi (1987) New view of entanglements in dense polymer systems. *Phys. Rev. Lett.* **59**, pp. 2674–2677
26. R. Everaers, S. K. Sukumaran, G. S. Grest, C. Svedborg, A. Sivasubramanian, and K. Kremer (2004) Rheology and microscopic topology of entangled polymeric systems. *Science* **303**, pp. 823–826
27. V. Harmandaris, V. G. Mavrantzas, D. N. Theodorou, M. Kröger, J. Ramírez, H. C. Öttinger, and D. Vlassopoulos (2003) Crossover from the Rouse to the entangled polymer melt regime: Signals from long, detailed atomistic molecular dynamics simulations, supported by rheological experiments. *Macromolecules* **36**, pp. 1376–1387
28. V. Harmandaris, V. G. Mavrantzas, and D. N. Theodorou (1998) Atomistic Molecular Dynamics Simulation of Polydisperse Linear Polyethylene Melts. *Macromolecules* **31**, pp. 7934–7943
29. N. Karayiannis and V. G. Mavrantzas (2005) in preparation

30. Y. Masubuchi, G. Ianniruberto, F. Greco, and G. Marrucci (2003) Entanglement molecular weight and frequency response of sliplink networks. *J. Chem. Phys.* **119**, pp. 6925–6930
31. C. Tzoumanekas and D. N. Theodorou (2006) Topological analysis of linear polymer melts. *Macromolecules* **39**, pp. 4592–4604
32. S. F. Edwards (1977) Theory of rubber elasticity. *Br. Polym. J.* **9**, pp. 140–143
33. J. D. Schieber (2003) Fluctuations in entanglements of polymer liquids. *J. Chem. Phys.* **118**, pp. 5162–5166

Drug-Target Binding Investigated by Quantum Mechanical/Molecular Mechanical (QM/MM) Methods

U. Rothlisberger[1] and P. Carloni[2]

[1] Computational Chemistry and Biochemistry, EPFL, Ecole Polytechnique Federale de Lausanne, 1015 Ecublens, CH,
[2] International School for Advanced Studies, 34100 Trieste, Italy, and DEMOCRITOS, Modeling Center for Research in Atomistic Simulation, INFM, Italy



Ursula Rothlisberger and Paolo Carloni

U. Rothlisberger and P. Carloni: *Drug-Target Binding Investigated by Quantum Mechanical/Molecular Mechanical (QM/MM) Methods*, Lect. Notes Phys. **704**, 449–479 (2006)
DOI 10.1007/3-540-35284-8_17 © Springer-Verlag Berlin Heidelberg 2006

1 Introduction 451

2 Drug-Target Interactions 452

2.1 Non-covalent Drug-Target Interactions 452
2.2 Covalent Drug-Target Interactions 453

3 Quantum Mechanics/Molecular Mechanics (QM/MM) Methods 454

3.1 The Basic Idea 454
3.2 A Bit of History 456
3.3 Formalism of Hybrid Car-Parrinello QM/MM Simulations 456
3.4 Possible Problems and Pitfalls 464
3.5 Newest Developments and Current Limitations 468

4 Covalent Drug-Target Binding: Examples 469

5 Concluding Remarks 474

References 474

Many important drugs, also used in the clinics, exert their function by binding covalently to their targets. Understanding their action requires quantum mechanical simulations. Here, after briefly reviewing few basic concepts of thermodynamics and kinetics of drug-target binding, we summarize principles and applications of Car-Parrinello quantum mechanics/molecular mechanics (QM/MM) simulations. From this discussion, this approach emerges as a computational methodology particularly well suited to investigate covalent binding in systems of pharmacological relevance.

1 Introduction

Drugs (D) exert their beneficial effect by associating to their cellular targets (T) (both proteins and DNA) following the scheme:

$$D + T \underset{K_{eq}}{\rightarrow} D \bullet T \tag{1}$$

The change in free energy is linked to the equilibrium constant by the relation $\Delta G = G_{D\bullet T} - G_{D+T} = \Delta H - T\Delta S = -\ \mathrm{RT} \ln K_{eq}$. The process may or may not be activated, i.e. may or may not require an activation free energy $\Delta G^{\#} \gg k_B T$. ΔG and $\Delta G^{\#}$ dictate affinity, selectivity and the kinetics of binding for the target, and therefore are key parameters in drug discovery.

Predicting the free energy profile associated to the molecular recognition process (1) is a major biological challenge for theoreticians. Recognition must occur between the correct partners within a dense and complex cellular environment, and as the result of a subtle free energy balance. In addition, the process is highly sensitive to environmental changes (such as ionic strength and pH).

To begin to address this issue, let us analyze the forces involved. These can be conventionally divided in two categories: non-covalent interactions and covalent bonds. Breaking and forming non-covalent interactions requires relatively small activation energies ($\Delta G^{\#} \sim kT$ or slightly larger). Formation of such interactions may provide a stabilization up to few tens of eV. It is the weakness of non-bonded interactions that makes them so useful in biology, as a small change in chemical environment can break and form them at physiological temperature (i.e. around 300 K): a typical and famous example is the forming and breaking of H-bonds in DNA base pairing during DNA replication.

In some specific cases, covalent bonding plays a role. The activation free energy is much larger ($\Delta G^{\#} \sim 0.5{-}1$ eV), as the process involves a chemically labile transition state; likewise, the stabilization due to the formation of a covalent bond is large: $\Delta G \sim 3.6$ eV for the formation of a single C–C bond [51].

We conclude that drug/target molecular recognition involves *invariably* breaking of non-bonded interactions with the solvent and *almost always* (except maybe in some cases of covalently-bound drugs) formation of new, likewise interactions with the targets.

Here, we focus on recent progress in the theoretical study of interactions between drugs and biological systems. Excellent reviews are available for the case of non-covalently bound complexes, which are by far the majority- [25, 60, 71]. Here emphasis will be given to covalent binding.

2 Drug-Target Interactions

2.1 Non-covalent Drug-Target Interactions

Interaction Energy Between *D* and *T*

We dissect here the non-bonded energy between D and T in terms of separated contributions, as it is customary in standard biomolecular force-field based calculations such as AMBER [8] and CHARMM [42].

Electrostatics

This is the Coulombic interaction energy between nuclei and electrons of D with those of T. Assuming the nuclei as point charges, the negative electronic charge is still smeared out over space. Thus, a rigorous calculation involves an integration over the electronic charges of D and T. In biomolecular force fields, the electrons along with the nuclei are represented by point charges (q_D and q_T for drug and target, respectively) whose position and magnitude are set so as to reproduce known molecular properties:

$$E = \sum_{DT} \frac{q_D \boldsymbol{q}_T}{R_{DT}} \tag{2}$$

Here R_{DT} is the distance modulus between atoms D and T. Such a term is very often used also to describe H-bond interactions.

Polarization and Charge Transfer

Upon binding, T and D redistribute their charge, leading to an additional, negative energy term. Often, it may be that there is also a small amount of electron flow from one to the other [8]. The energies associated to such effects are smaller than electrostatic energies and they are usually neglected in biomolecular potentials, although significant progress in their theoretical treatment has been made in the last years [13].

Exchange Repulsion

The Pauli repulsion keeps electrons of D and T with the same spin spatially apart. In addition, the nuclei repel each other increasingly upon approaching each other. This repulsion energy is usually accounted for by a positive term which grows rapidly as the distance between the atoms increases:

$$E = \sum_{DT} \frac{A_{DT}}{R_{DT}^{12}} \tag{3}$$

London Dispersion

Dispersion forces arise from quantum fluctuations of the charge densities on T and D. This attractive term is described by:

$$E = \sum_{DT} \frac{B_{DT}}{R_{DT}^{6}} \tag{4}$$

Thermodynamic and Kinetic Considerations

As illustrative example, let us consider the formation of an H-bond between T and D. The formation of such a H-bond is accompanied by the disruption of H-bonds of D and T with the solvent as well as the formation of H-bonds between water molecules. Thus, $\Delta H < 0$ only if the interactions with the solvent with the uncomplexed drug will be weaker than the interactions of the solvent with the $D \bullet T$ complex. Similarly, for the entropy, $\Delta S > 0$ only if the entropy gain associated to the hydrophobic effect [31] is larger than the entropy cost due to the association of the complex. From this very simplified discussion it is clear that the affinity will be governed by an algebraic sum of competing and very subtle factors. Analogous considerations could be done for other interactions (such as electrostatic salt bridges and van der Waals interactions [16]). On the basis of these simple arguments, it should therefore not be surprising that enthalpy and entropy contributions may vary largely in magnitude and that some non-covalent binding processes are purely entropy driven, with $\Delta H \sim 0\,\text{eV}$ or even slightly positive, as shown by a set of specific complexes in Fig. 1. We further notice that the free energy of non-covalent binding is well below 1 eV (Fig. 1).

Free-energy calculations based on effective potentials have shown to be predictive for a variety of non-covalent complexes [25, 60, 71, 80]. As expected, systems which are difficult to parameterize (e.g. metal-containing drugs or proteins) may still pose challenges to the biosimulator [30].

2.2 Covalent Drug-Target Interactions

A comparison between the free energies in Fig. 1 and those reported above for chemical bonds allows us to conclude that: (i) ΔG of efficient non-covalent

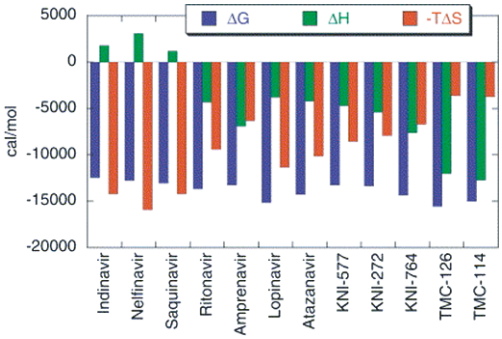


Fig. 1. Binding thermodynamics of $D \bullet T$ complexes. T is the HIV-1 protease enzyme, a major target for antiAIDS therapy. The drugs are a variety of enzyme non-covalent inhibitors. Data taken at 298 K. Dark grey $= \Delta G$, medium grey $= \Delta H$ and light grey $= -T\Delta S$. 1 eV $= 23060.5$ cal/mol. Reproduced with permission from [50]

drugs is always much smaller than that of a covalent bond. Thus, covalent binding is irreversible, not allowing the dissociation of the complex (ii) Some non-covalent interactions that the drug forms with the solvent may be lost, unlike in non-covalent complexes, where the loss of say a H-bond might be of the same order of magnitude as ΔG. In addition, the activation free energy associated to the formation of chemical bonds to the target determines often the rate-limiting step of the entire binding process.

Thus, a key issue in covalently bound drugs is to describe the formation of the chemical bond between D and T. This requires quantum chemical approaches, which are described in the next Section.

3 Quantum Mechanics/Molecular Mechanics (QM/MM) Methods

3.1 The Basic Idea

In principle, the use of a quantum mechanical method for the description of covalent drug/target interactions offers many advantages. First, the complex invariably contains non-standard groups for which force field parameterizations are not readily available and with a QM treatment the painstaking process of force field development can be avoided elegantly. Second, and most importantly, bond forming/breaking processes can be directly studied. However, most quantum mechanical electronic structure methods are restricted to the treatment of relatively small systems in the gas phase whereas complex chemical and biochemical processes usually occur in heterogeneous condensed phase environments consisting of thousands of atoms.

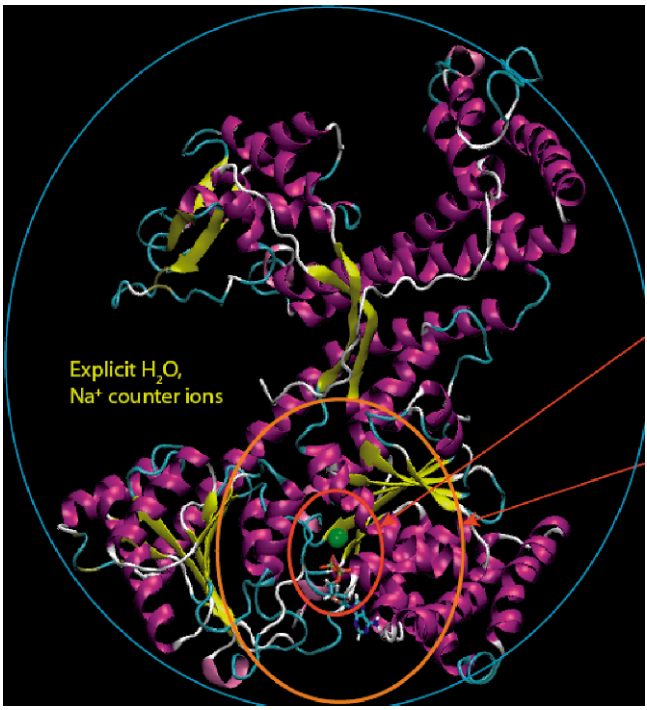


Fig. 2. Typical QM/MM partitioning of a complex system. QM: region treated with a quantum mechanical method; MM: region treated with a molecular mechanics force field; IL: optional intermediate layer between QM and MM parts. Note that in most QM/MM schemes the initial partitioning of the system is maintained throughout, i.e. atoms belonging to the QM part remain QM atoms and MM atoms are always treated on the classical level

One possible solution for the modelling of such systems is the choice of a hierarchical hybrid approach in which the whole system is partitioned into a localized chemically active region (treated with a quantum mechanical method) (QM region) and its environment (MM region) (treated with empirical potentials, Fig. 2). This is the so-called quantum mechanical/molecular mechanical[1] (QM/MM) method, in which the computational effort can be concentrated on the part of the system where it is most needed whereas the effects of the surroundings are taken into account with a more expedient model.

[1] The term "molecular mechanics" is used for historical reasons although most of the current QM/MM schemes are actually performing molecular dynamics at finite temperatures and not just 'mechanical' optimizations in the molecular mechanics sense.

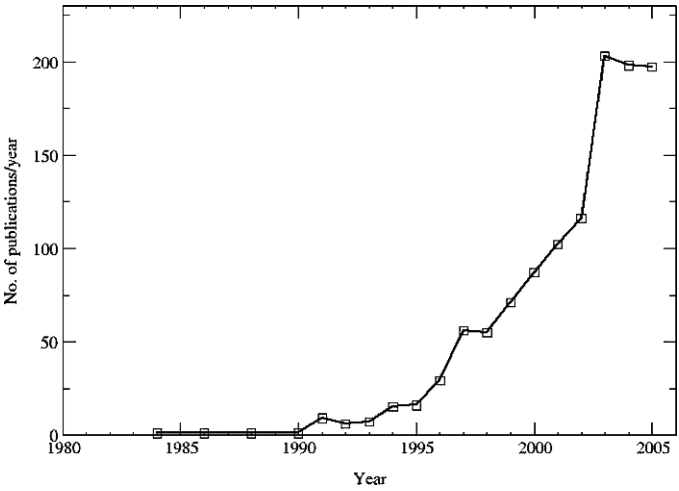


Fig. 3. Number of ISI publications per year that contain the terms 'quantum mechanical and molecular mechanical' or "QM/MM"

3.2 A Bit of History

QM/MM simulations were first introduced in 1976 by Warshel et al. [74] followed by several other groups [15, 61]. Possibly due to the rather limited computer resources available at the time, the original pioneering work found relatively little immediate impact as illustrated by a plot of the number of published articles that contain the terms "quantum mechanical and molecular mechanical" or "QM/MM" as a function of time (Fig. 3). Only in the beginning of the 90'ties a larger community started to work with this new method. In spite of the increasing popularity of the QM/MM approach there were still a number of technical problems that had to be solved in order to make the method numerically robust and reliable (also illustrated by the large oscillations in the number of publications). Around the mid 90'ties a regular QM/MM boom started. Nowadays, QM/MM simulations have become of age as a well established tool. Several commercially available packages include QM/MM options and in most implementations the technical problems have been well mastered. A number of excellent reviews are available on this topic [38, 59].

3.3 Formalism of Hybrid Car-Parrinello QM/MM Simulations

In this section, we will give a general introduction to different QM/MM methodologies, in particular view of applications of these relatively new tools for a fundamental understanding and possible rational design of complex drug/target interactions. We will focus in particular on Car-Parrinello QM/MM simulations [32] as the applications presented here use this method.

This scheme applies the density functional based Car-Parrinello method [4] as a QM core. We also dedicate one paragraph to the "possible problems and pitfalls" that hopefully provides some help for novices in this field to distinguish between differences in the various QM/MM implementations and to validate their performance and accuracy.

There are two fundamentally different ways to carry out calculations on a system that has been partitioned into a QM and a MM region. In the *subtractive* scheme, the QM calculation is performed on an isolated QM system and the environment effects (i.e. the influence of the MM system on the QM system) are estimated at the lower level by the difference between two MM calculations, one treating the entire system (QM+MM) and one the QM region only. In this approach, the total energy of the embedded system is written as

$$E = E^{QM}(QM) + E^{MM}(QM + MM) - E^{MM}(QM) \tag{5}$$

with the analogous expression

$$F_I = -\frac{\partial E}{\partial \boldsymbol{R}_I} = -\frac{\partial E^{QM}(QM)}{\partial \boldsymbol{R}_I} - \frac{\partial E^{MM}(QM + MM)}{\partial \boldsymbol{R}_I} + \frac{\partial E^{MM}(QM)}{\partial \boldsymbol{R}_I} \tag{6}$$

for the force F_{I} acting on atom I at position $\mathbf{R}_{\mathrm{I}}$. A QM/MM implementation that uses a subtractive scheme is the integrated molecular-orbital molecular mechanics (IMOMM) scheme developed by Maseras and Morokuma [47] available in Gaussian03 (`http://www.gaussian.com`). Note that in a subtractive scheme, all calculations are performed with "pure" (either fully QM or fully MM) Hamiltonians. The advantage of such an approach lies in the fact that there is no QM/MM interface that has to be dealt with. The disadvantage is that the environment influence is often described at a very simple level (MM level). The electrostatic interaction between QM and MM part for instance, is described entirely at the force field level, i.e. by Coulomb interactions between effective point charges. Such an electrostatic coupling between QM and MM part is called "mechanical coupling" to indicate that in such an approach, the electrons do not feel anything of the classical electrostatic field of the environment and all the electrostatic interactions between QM and MM part act solely on the level of the atoms. The influence of the environment as described by the lower level method is only a reasonable estimate for the environment effect at the higher level if the two descriptions are not too different. To minimize the difference in the treatment of the two regions, the original IMOMM scheme has been extended to three (respectively multiple) layers (ONIOM) [64]. A typical ONIOM calculation consists for example of a first-principles (DFT or wavefunction based method) region, adjacent to a layer treated with a semiempirical method followed by a third layer treated at the molecular mechanics level.

Most of the current QM/MM implementations however use an *additive* scheme. In an *additive* QM/MM implementation the system is described by a single hybrid Hamiltonian

$$H = H_{QM} + H_{MM} + H_{QM/MM} \tag{7}$$

where H_{QM} is the quantum Hamiltonian, H_{MM} is the molecular mechanics Hamiltonian, and $H_{QM/MM}$ is the interaction Hamiltonian between QM and MM system. The lowest eigenvalue of the Hamiltonian in (7) determines the total energy of the mixed quantum/classical system

$$E = E_{QM} + E_{MM} + E_{QM/MM} \tag{8}$$

The advantage of an additive scheme is that the QM calculation can be directly executed in the presence of the classical environment in such a way that e.g. the electron density of the QM system is optimized in (and is polarized by) the external electrostatic field of the surroundings. The prize for this is that the real system is replaced by a somewhat artificial, heterogeneous construct, in which different parts of the system are described at largely disparate levels, i.e. one part of the system is represented in electronic detail whereas all the surroundings is reduced to a purely classical (mechanical and electrostatic) description. In this way, an abrupt QM/MM border is created. One of the drastic consequences of this approach is the fact that when passing from the QM to the MM zone of the system the electrons suddenly cease to exist. Such a simplified description can necessarily only constitute a somewhat crude representation of the true uniform system.

To identify where the main approximations enter and how severe they are, let us consider the case when the entire system (QM + MM) is described uniformly at the QM level, say at the density functional level. The total (electronic plus core-core interaction) energy of such a system is given by the density functional [22]

$$E = T[\rho] + \int_{\Omega} V^{ex}(\boldsymbol{r})\rho(\boldsymbol{r})d\boldsymbol{r} + \frac{1}{2}\iint \frac{\rho(\boldsymbol{r}_1)\rho(\boldsymbol{r}_2)}{r_{12}} d\boldsymbol{r}_1 d\boldsymbol{r}_2 + E_{xc}[\rho] + \frac{1}{2}\sum_I \sum_J \frac{Z_1 Z_J}{R_{IJ}} \tag{9}$$

where T and E_{xc} are the kinetic, respectively the exchange-correlation energy density functionals, V^{ex} is the external electrostatic potential created by the positively charged nuclei. r represents the electronic coordinates, while r_{12} refers to interelectronic and R_{IJ} to internuclear distances. Z_I and Z_J represent the nuclear (or core) charge of atom I and J, respectively.

Now, we partition the system into two parts, A and B, with respective densities ρ_A and ρ_B. The total density ρ can be expressed as (see also the constrained electron density KSCED embedding formalism of Wesolowski et al. [76]) as

$$\rho(\boldsymbol{r}) = \rho_A(\boldsymbol{r}) + \rho_B(\boldsymbol{r}) \tag{10}$$

Analogous to (8) the total energy is given by

$$E = E_A + E_B + E_{A-B} \tag{11}$$

with

$$\begin{aligned} E = {} & T\lceil\rho_A\rceil + T\lceil\rho_B\rceil + T^{NL} + \int_\Omega V^{ex}(\boldsymbol{r})\rho_A(\boldsymbol{r})d\boldsymbol{r} \\ & + \int_\Omega V^{ex}(\boldsymbol{r})\rho_B(\boldsymbol{r})d\boldsymbol{r} \\ & + \frac{1}{2}\iint \frac{\rho_A(\boldsymbol{r}_1\rho_A(\boldsymbol{r}_2)}{r_{12}}d\boldsymbol{r}_1 d\boldsymbol{r}_2 + \frac{1}{2}\iint \frac{\rho_B(\boldsymbol{r}_1)\rho_B(\boldsymbol{r}_2)}{r_{12}}d\boldsymbol{r}_1 d\boldsymbol{r}_2 \\ & + \frac{1}{2}\iint \frac{\rho_A(\boldsymbol{r}_1)\rho_B(\boldsymbol{r}_2)}{r_{12}}d\boldsymbol{r}_1 d\boldsymbol{r}_2 \\ & + E_{xc}[\rho_A] + E_{xc}[\rho_B] + E_{xc}^{NL} + \frac{1}{2}\sum_I \sum_J \frac{Z_I Z_J}{R_{IJ}} \end{aligned} \tag{12}$$

The terms T^{NL} and E_{xc}^{NL} are

$$T^{NL} = T[\rho_A + \rho_B] - T[\rho_A] - T[\rho_B] \tag{13}$$

$$E_{xc}^{NL} = E_{xc}[\rho_A + \rho_B] - E_{xc}[\rho_A] - E_{xc}[\rho_B] \tag{14}$$

These terms account for the nonlinearity of the kinetic energy and the exchange-correlation density functionals. They are only zero if ρ_A and ρ_B are spatially well separated.

The term E_{xc}^{NL} in (14), arises also in the construction of *ab initio* atomic pseudopotentials when the system has to be partitioned into valence and core densities. In this case E_{xc}^{NL} is called *nonlinear core correction* [41].

For the particular case that we describe part A of the system with another approach than part B, it is useful to separate also the external potential V^{ex} into contributions from the nuclear cores of A and those of B

$$V^{ex}(\boldsymbol{r}) = V_A^{ex}(\boldsymbol{r}) + V_B^{ex}(\boldsymbol{r}) \tag{15}$$

E_A and E_B in (11) are given by the terms

$$\begin{aligned} E_Y = {} & T\lceil\rho_Y\rceil + \int_\Omega V_Y^{ex}(\boldsymbol{r})\rho_Y(\boldsymbol{r})d\boldsymbol{r} + \frac{1}{2}\iint \frac{\rho_Y(\boldsymbol{r}_1)\rho_Y(\boldsymbol{r}_2)}{r_{12}}d\boldsymbol{r}_1 d\boldsymbol{r}_2 \\ & + E_{xc}[\rho_Y] + \frac{1}{2}\sum_{I\varepsilon Y}\sum_{J\varepsilon Y}\frac{Z_I Z_J}{R_{IJ}} \end{aligned} \tag{16}$$

where $Y = A$, respectively or B. Often the nuclear charges Z_I Z_J are expanded into Gaussian shaped charge distributions with width R_c of the form

$$Z_I = \int_\Omega \rho_I^{nuclear}(\boldsymbol{r} - \boldsymbol{R}_I)dr = \int_\Omega \frac{Z_I}{R_c^3}\pi^{-2/3}\exp\left[-\frac{|\boldsymbol{r} - \boldsymbol{R}_I|^2}{R_c^2}\right] \tag{17}$$

and the three Coulombic terms can be summarized into one expression, which depends on the combined nuclear and electronic charge density $\rho^{el+nuclear} = \rho^{el} + \rho^{nuclear}$

$$\int_\Omega V_Y^{ex}(\boldsymbol{r})\rho_Y(\boldsymbol{r})dr + \frac{1}{2}\iint \frac{\rho_Y(\boldsymbol{r}_1)\rho_Y(\boldsymbol{r}_2)}{r_{12}}d\boldsymbol{r}_1 d\boldsymbol{r}_2 + \frac{1}{2}\sum_{I\varepsilon Y}\sum_{J\varepsilon Y}\frac{Z_I Z_J}{R_{IJ}}$$
$$= \frac{1}{2}\iint \frac{\rho_Y^{el+nuclear}(\boldsymbol{r}_1)\rho_Y^{el+nuclear}(\boldsymbol{r}_2)}{r_{12}}\boldsymbol{d}\boldsymbol{r}_1\boldsymbol{d}\boldsymbol{r}_2 \tag{18}$$

The interface term E_{A-B} describes the interaction between A and B and therefore contains all the remaining terms

$$E_{A-B} = T^{NL} + \int_\Omega V_B^{ex}(\boldsymbol{r})\rho_A(\boldsymbol{r})d\boldsymbol{r} + \int_\Omega V_A^{ex}(\boldsymbol{r})\rho_B(\boldsymbol{r})d\boldsymbol{r}$$
$$+\frac{1}{2}\iint \frac{\rho_A(\boldsymbol{r}_1)\rho_B(\boldsymbol{r}_2)}{r_{12}}d\boldsymbol{r}_1 d\boldsymbol{r}_2 + E_{xc}^{NL} + \frac{1}{2}\sum_{I\in A}\sum_{J\in B}\frac{Z_I Z_J}{R_{IJ}} \tag{19}$$

For the special case that part A is treated with a QM and part B with an MM method, the first two energy terms in (8) correspond to

$$E_{QM} = T\lceil\rho_{QM}\rceil + \frac{1}{2}\iint \frac{\rho_{QM}^{el+nuclear}(\boldsymbol{r}_1)\rho_{QM}^{el+nuclear}(\boldsymbol{r}_2)}{r_{12}}d\boldsymbol{r}_1 d\boldsymbol{r}_2 + E_{xc}\left[\rho_{QM}\right] \tag{20}$$

$$E_{MM} = T\lceil\rho_{MM}\rceil + \frac{1}{2}\iint \frac{\rho_{MM}^{el+nuclear}(\boldsymbol{r}_1)\rho_{MM}^{el+nuclear}(\boldsymbol{r}_2)}{r_{12}}d\boldsymbol{r}_1 d\boldsymbol{r}_2 + E_{xc}[\rho_{MM}] \tag{21}$$

E_{MM} is delegated to the classical force field. Clearly, none of the current force fields for biomolecular simulations can provide an exact match of the terms in (21). However, we will try to point out which typical analytical expressions are currently in use to mimic the physical effects described by (21).

As electrons are not considered explicitly, force fields are parameterized to single (or to the average of several) configurations with fixed electron density distributions. Therefore the kinetic energy term in (21) can be considered as an additive constant that is not taken explicitly into account. The effect of the exchange-correlation energy functional is often replaced by a pair-additive van der Waals term:

$$E_{xc} \approx E^{vdW} = \sum_{I'}\sum_{J'} 4\varepsilon_{I'J'}\left(\left(\frac{\sigma_{I'J'}}{R_{I'J'}}\right)^{12} - \left(\frac{\sigma_{I'J'}}{R_{I'J'}}\right)^{6}\right). \tag{22}$$

The electrostatic potential due to the combined electronic and ionic charge distribution is approximated via effective point charges, usually located at atomic positions

$$\frac{1}{2}\iint \frac{\rho_{MM}^{el+nuclear}(\boldsymbol{r}_1)\rho_{MM}^{el+nuclear}(\boldsymbol{r}_2)}{r_{12}} d\boldsymbol{r}_1 d\boldsymbol{r}_2 \approx \frac{1}{2}\sum_{I'}\sum_{J'}\frac{q_{I'}q'_{J'}}{R_{I'J'}} \qquad (23)$$

The set of effective (often empirical) point charges commonly used in biomolecular force fields cannot be expected to faithfully reproduce the left hand side of (23), i.e. to be fully consistent with the electronic structure method used for the QM part. However, due to the extremely cumbersome work involved in the development of a general and transferable force field for complex biological systems, people usually prefer to employ existing parameterizations instead of constructing a fully "ab initio" derived force field. In addition, it turns out, that although the magnitude of effective point charges used in different force fields can vary largely, the average electrostatic potentials seem to be in surprisingly good agreement with each other as well as with e.g. DFT descriptions [58].

If we compare the electrostatic field of a QM/MM Car-Parrinello molecular dynamics (MD) run of a QM Gly-Ala dipeptide in its zwitterionic form in aqueous solution (MM) with the equivalent quantities resulting from purely classical MD runs with the AMBER/parm95 [10] and the GROMOS96 [57] force fields, it turns out that these standard non-polarizable force field models are able to reproduce the electrostatic field all along the MD trajectory very well. The standard deviations for an MD trajectory exploring only one conformer is between 6 and 13% for the AMBER charge set and only slightly higher, between 9 and 16%, for the GROMOS charge set, which is a united atom model and hence includes less degrees of freedom to reproduce the field [58].

In spite of this somewhat reassuring caveat, the fact remains that real electronic charge distributions are far from mere assemblies of point charges. The point charge approximation breaks completely down in the description of covalent chemical bonds that are characterized by highly inhomogeneous, highly directional distributions of the electron density. Clearly, simple electrostatic/van der Waals descriptions such as those in (22) and (23) cannot reproduce the intricacies of chemical bonding. In most force fields, the interaction between nearest, second nearest and third nearest neighbor atoms linked by chemical bonding are therefore mimicked by mechanical bond, angle and torsional angle terms of the typical form given in (24)

$$\begin{aligned} E_{MM}^{bonded} = &\sum_{b}\frac{1}{2}k_b(R_{I'J'}-b_0)^2 + \sum_{\theta}\frac{1}{2}k_\theta(\theta_{I'J'K'}-\theta_0)^2 \\ &+ \sum_{\varphi}\sum_{n}k_n\left[1+\cos(n\varphi_{I'J'K'L'}-\varphi_0)\right] \end{aligned} \qquad (24)$$

where the first term runs over all bonds b with harmonic force constant k_b and equilibrium bond length b_0, the second term runs over all bonding angles θ with harmonic force constant k_θ and equilibrium bonding angle θ_0, whereas the last term is a sum over all dihedral angle interactions φ with multiplicity n and corresponding force constants k_n and phases φ_0. For the atoms connected via bonded terms, the nonbonded (electrostatic and van der Waals) interactions are either omitted or scaled down (so called "exclusion rules").

Using (22)–(24) the interaction energy in (19) becomes

$$\begin{aligned} E_{QM/MM} = & \sum_{I'} \int_{\Omega} \frac{q_{I'}}{|\boldsymbol{R}_{I'} - \boldsymbol{r}|} \rho_{QM}^{el+nuclear}(\boldsymbol{r}) d\boldsymbol{r} \\ & + \frac{1}{2} \sum_{I'} \sum_{I} 4\varepsilon_{I'I} \left(\left(\frac{\sigma_{I'I}}{R_{I'I}} \right)^{12} - \left(\frac{\sigma_{I'I}}{R_{I'I}} \right)^{6} \right) \\ & + \sum_{b} \frac{1}{2} k_b (R_{I'I} - b_0)^2 + \sum_{\theta} \frac{1}{2} k_\theta (\theta_{I''J''K''} - \theta_0)^2 \\ & + \sum_{\varphi} \sum_{n} k_n \left[1 + \cos(n\varphi_{I''J''K''L''} - \varphi_0) \right] \end{aligned} \tag{25}$$

where I runs over QM and I′ over MM atoms and at least one atom of the triple (I″J″K″) and quadruples (I″J″K″L″) of bonded atoms is a QM atom. In this formulation, the effective classical point charges act as an external field to the QM calculation, i.e. the electron density of the QM part is polarized by the classical environment (in contrast to the *subtractive* approach described earlier). Note however, that both the van der Waals term and the bonded terms are acting on atomic positions only i.e. are not part of the total electronic potential and thus are not directly felt by the electrons. If we want to achieve a close facsimile of a full QM description, the deviations caused by the actual MM representation have to be compensated by a correction term ΔV in the total potential that the electrons of the QM part experience

$$V_{tot} = V_{QM} + V_{QM/MM} + \Delta V \tag{26}$$

$$\Delta V = \Delta V^{NL} + \Delta V_{ele} \tag{27}$$

Where ΔV_{ele} accounts for the error in the electrostatic terms (deviation of the classical electrostatic potential from the QM reference and reduction of the electronic density distribution to a point charge representation) whereas the nonlinear correction term ΔV^{NL} results from the nonlinearity corrections in (13) and (14). Thus this term is a mere artifact of the density partitioning and is not present in a system treated at the uniform level. To keep this term minimal, the somewhat trivial but important condition has to be fulfilled that the QM part has to be chosen in such a way that the electronic wavefunctions are localized to this region. If this condition cannot be fulfilled, the correction term ΔV^{NL} gains in importance (see paragraph problems and pitfalls).

How can we assess the importance of the correction term ΔV in practice? Our main goal is to perform a QM calculation in a realistic environment. Ideally, we would like that the electron density in the QM region, ρ_{QM}, matches as close as possible the electron density in the same region produced by a full QM representation of the system (ρ_{true}). According to the Hohenberg–Kohn theorem [22], if the two densities are identical, all the properties we calculate for the QM region are identical to those of the real system. In other words, if we determine the correction potential ΔV in such a way that the total electronic potential in a QM/MM simulation V_{tot} minimizes the density difference

$$\int\limits_{\Omega'} \left(\rho_{true}(\boldsymbol{r}) - \rho_{QM}\left(\boldsymbol{r}\right)\right)^2 d\boldsymbol{r} \tag{28}$$

where Ω' is a suitably chosen volume of the QM region, our QM/MM simulation approaches the full QM reference results in an optimal way (see also paragraph on possible problems and pitfalls).

For a concrete example of a QM/MM implementation, let us consider the mixed QM/MM Car-Parrinello approach used in the following applications. The Car-Parrinello method [4] can be extended into a QM/MM scheme using a mixed Lagrangian of the form [32]

$$\begin{aligned} L = \frac{1}{2}\mu \sum_i \int d\boldsymbol{r} \dot{\psi}_i^* \left(\boldsymbol{r}\right) \dot{\psi}_i \left(\boldsymbol{r}\right) + \frac{1}{2} \sum_I M_I \dot{\boldsymbol{R}}_I^2 - E_{MM} - E_{QM/MM} - E_{QM} \\ + \sum_{i,j} \Lambda_{i,j} \left(\int d\boldsymbol{r} \psi_i^* \left(\boldsymbol{r}\right) \psi_j \left(\boldsymbol{r}\right) - \delta_{i,j} \right) \end{aligned} \tag{29}$$

Where μ is the fictitious mass associated with the electronic degrees of freedom, ψ_i are the Kohn-Sham one particle orbitals, M_I is the mass of atom I and $\Lambda_{i,j}$ are Lagrange multipliers that enforce orthonormality of the Kohn-Sham orbitals. The energy of the QM system E_{QM} is given by the Kohn-Sham energy density functional [82]

$$\begin{aligned} E_{QM} = E_{KS}\left[\psi_i, \boldsymbol{R}_I\right] = -\frac{1}{2} \sum_i \int d\boldsymbol{r} \psi_i^* \left(\boldsymbol{r}\right) \nabla^2 \psi_i \left(\boldsymbol{r}\right) \\ + \int dr V^{ex}(\boldsymbol{r}) \, \rho_{QM}\left(\boldsymbol{r}\right) \\ + \frac{1}{2} \int d\boldsymbol{r} d\boldsymbol{r}' \rho_{QM}\left(\boldsymbol{r}\right) \frac{1}{\left|\boldsymbol{r} - \boldsymbol{r}'\right|} \rho_{QM}\left(\boldsymbol{r}'\right) + E_{xc}\left[\rho_{QM}\left(\boldsymbol{r}\right)\right] \end{aligned} \tag{30}$$

where for the spin unpolarized case, the electron density $\rho_{QM}(\boldsymbol{r})$ is given by the sum of the densities of the doubly occupied one-particle states

$$\rho_{QM}\left(\boldsymbol{r}\right) = 2 \sum_i \psi_i^* \left(\boldsymbol{r}\right) \psi_i \left(\boldsymbol{r}\right) \tag{31}$$

The purely classical part E_{MM} is described by a standard biomolecular force field

$$E_{MM} = E_{MM}^{bonded} + E_{MM}^{non-bonded} \tag{32}$$

as given by (22)–(24). The interaction between QM and MM parts, $E_{QM/MM}$, is included in the form of (25) with the only exception of the harmonic bond interactions between QM and MM atoms which are omitted from the classical description and treated at the QM level.

Usually, standard implementations of the Car-Parrinello methods use plane wave basis sets. In this case, due to the high intrinsic flexibility of a plane wave basis set (in contrast to e.g. the minimal basis sets used in semiempirical QM/MM calculations), special care has to be taken that the QM/MM interface is described in an accurate and consistent way. In our case, the quantum/classical correction term ΔV consists of specifically designed monovalent pseudo potentials to represent bonds between QM and MM parts of the system [69] and of modified screened Coulomb potentials for the interaction of the quantum electron density with closeby classical point charges [33].

In the context of a plane wave based Car-Parrinello scheme, a direct evaluation of the first term of (25) is prohibitive as it involves of the order of $\mathrm{N}_r \times N_{MM}$ operations, where N_r is the number of real space grid points (typically ca. 100^3) and N_{MM} is the number of classical atoms (usually of the order of 10,000 or more in systems of biochemical relevance). Therefore, the interaction between the QM system and the more distant MM atoms is included via a Hamiltonian term explicitly coupling the multipole moments of the quantum charge distribution with the classical point charges. This two level electrostatic coupling scheme can also be refined to an intermediate third layer that makes efficient use of variational D-RESP charges [33, 34]. Recently, highly efficient schemes based on a dual grid approach [79] or a multigrid approach with Gaussian expansion [37] have also been proposed in this context.

The QM/MM Car-Parrinello implementation that was used for the applications in this article, establishes an interface between the Car-Parrinello code CPMD [11] and the classical force fields GROMOS96 [57] and AMBER [10] in combination with a particle-particle-particle mesh (P3M) treatment of the long-range electrostatic interactions [23]. This QM/MM interface is freely distributed with the CPMD package (www.cpmd.org).

With this implementation, efficient and consistent QM/MM Car-Parrinello simulations of complex extended systems of several 10'000–100'000 atoms can be performed in which the steric and electrostatic effects of the surroundings are taken explicitly into account.

3.4 Possible Problems and Pitfalls

The Link Atom Problem

In the previous section, we have seen that the main intrinsic approximations of a QM/MM approach lie in the reduction of the real electron density distri-

bution of the MM part to a mere point charge representation and the neglect of the kinetic energy and exchange-correlation corrections (26) and (27) on the electronic level. All three of these terms are particularly severe in the neighbourhood of a covalent chemical bond, where the electron density distribution is far from isotropic and the densities of QM and MM part are strongly overlapping. In force field descriptions, these deficiencies in the description of chemical bonding are remedied by including the special bonding terms given in (24). However, these terms are a function of atomic coordinates only and do not influence the electronic potential in a direct way. One of the most current problems in QM/MM simulations thus occurs when the border between QM and MM parts has to run across a chemical bond (*link atom problem*). For QM/MM simulations of covalent drug/target interactions, this is essentially always the case. A typical QM/MM partitioning for such a problem is to include the entire drug (or parts of it) as well as the region of the target that forms the covalent link into the QM region. This means that the biological macromolecule that constitutes the target has to be cut into a QM and a MM region. As electrons cease to exist when passing from the QM to the MM region, the QM system contains unsaturated valencies and has to be made chemically inert.

This can be done in the spirit of (26) by introducing an explicit correction term in the total electronic potential felt by the QM electrons. For the case of a QM/MM bond cut, we represent the correction potential in (27) by a monovalent pseudopotential situated at the position of the first MM atom (Fig. 4). This pseudopotential is constructed in such a way that the electrons of the QM region are scattered correctly by the classical environment. In the Car-Parrinello based QM/MM scheme used here we employ analytic, nonlocal pseudopotentials of the Goedecker type [21]

$$
\begin{aligned}
V^{eff}(\boldsymbol{r},\boldsymbol{r}') &= V^{loc}(\boldsymbol{r})\,\delta(\boldsymbol{r}-\boldsymbol{r}') + V^{nl}(\boldsymbol{r},\boldsymbol{r}') \\
V^{loc}(\boldsymbol{r}) &= \frac{-Z_{ion}}{r} erf\left[\frac{\boldsymbol{r}}{r_{loc}\sqrt{2}}\right] + \exp\left[-\frac{r^2}{2r_{loc}^2}\right] \\
&\quad \left(C_1 + C_2\left(\frac{r}{r_{loc}}\right)^2 + C_3\left(\frac{r}{r_{loc}}\right)^4 + C_4\left(\frac{r}{r_{loc}}\right)^6\right) \\
V_l^{nl}(\boldsymbol{r},\boldsymbol{r}') &= \sum_{m=-l}^{+l} Y_{lm}(\boldsymbol{r}) \sum_{j,h=1}^{3} p_{lh}(\boldsymbol{r})\, h_{lhj} p_{lj}(\boldsymbol{r}')\, Y_{lm}^*(\boldsymbol{r}) \\
&\quad p_{lh}(r) \propto r^{l+2(h-1)} \exp\left[-r^2/2r_l^2\right]
\end{aligned}
\tag{33}
$$

to represent the MM atoms involved in QM/MM bond cuts. The adjustable parameters r_{loc}, r_l and C_1 to C_4 are determined in analogy to (28) by minimizing the density penalty

$$
F\left[\rho_{QM}(\boldsymbol{r}), \{\sigma_i\}\right] = \int_{\Omega'} d\boldsymbol{r}\, \left|\rho_{ref}(\boldsymbol{r}) - \rho_{QM}(\boldsymbol{r}, \{\sigma_i\})\right|^2 \tag{34}
$$

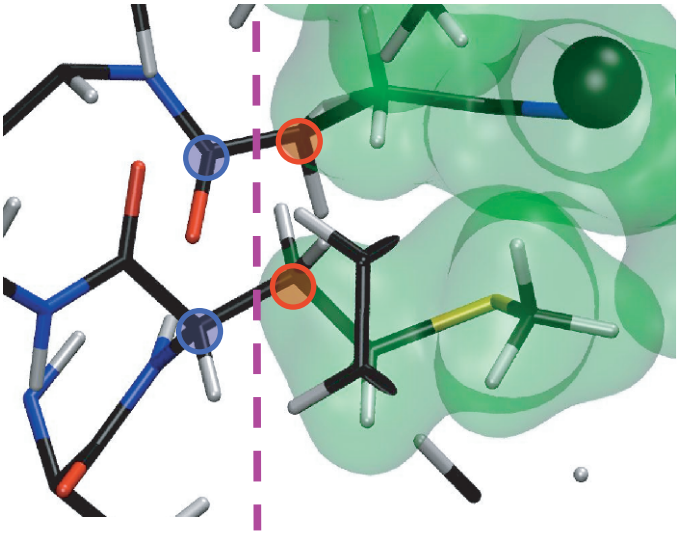


Fig. 4. Zoom in to the QM/MM interface region. In the case shown here the QM/MM boundary (indicated by the grey dashed line) runs across covalent bonds. A contour plot of the electron density distribution of the QM part is shown in grey. Atomic positions of QM and MM part are represented by sticks. The last QM atoms at the boundary are indicated by a red circle the first MM atoms with dark grey circle

where the σ_i's are the set of adjustable parameters and ρ_{ref} is a reference density that approximates ρ_{true} in (28). ρ_{ref} is usually determined from a QM/MM calculation of the system with extended QM part [72].

There are many other (ad hoc) procedures in use to cure the link atom problem. Commonly used strategies are to add capping atoms (hydrogen or fluorine) or to represent the last QM atom with frozen frontier orbitals [1]. However, hydrogen (or less commonly used fluorine) capping introduces new atoms into the QM system that are not present in the real system. As a consequence, the QM portion is chemically not identical with the real system (e.g. the "true" system may contain C–C bonds at the boundary that are now described with C–H bonds that clearly have different electronic and chemical properties). Furthermore, additional degrees of freedom have been introduced and interactions of these "nonexisting" ghost atoms with the classical environment have to be carefully removed. Some of these drawbacks are remedied by the use of frozen frontier orbitals for the boundary atoms. In this way, no additional physical interactions and degrees of freedom are introduced and the QM part retains its original composition. However, frozen orbitals have to be determined via calculations on small model systems and, as the name says, they remain frozen when transferred into the real environment. Specially parameterized pseudopotentials such as the ones described above, on the other hand have the additional flexibility to adjust to changes in the environment.

Electron Spill Out

Another possible artifact in QM/MM simulations, in particular in combination with extended and highly flexible basis sets (such as e.g. plane waves) is the *electron spill out* problem. As shown in (25), the exchange interactions between QM and MM part are taken into account on the level of atomic pair interactions only. Once again, these terms do not directly affect the electrons of the QM part. For a proper description of the electronic structure of the QM region an electronic correction term ΔV_{xc}^{NL} has to be included. As we have seen, this term is especially important for regions with overlapping or nearly overlapping densities between QM and MM parts, which is particularly the case for the nearby atoms surroundings the QM region. Due to the fact that the MM part contains no explicit electrons, the electrons of the QM part are no longer repelled by the closed-shell cores of the MM region. As a result of this missing Pauli repulsion, the electrons of the QM part can artificially localize on nearby positively charged classical point charges. This phenomenon is called *electron spill out*. This effect can be avoided by using Gaussian smeared (screened) classical charges (attention! drastic artifacts are possible by choosing to large widths for the Gaussian broadening) or by replacing the classical point charge potential by suitably constructed ionic pseudopotentials with screened electrostatic interactions [32].

$$E_{QM/MM}^{ele} = \sum_{I \in MM} q_I \int d\boldsymbol{r} \rho(\boldsymbol{r}) \nu_I(|\boldsymbol{r} - \boldsymbol{R}_I|) \tag{35}$$

where q_I is the classical point charge located at $\boldsymbol{R}_I$ and

$$\nu_I(|\boldsymbol{r} - \boldsymbol{R}_I|) = \frac{r_c^4 - r^4}{r_c^5 - r^5} \tag{36}$$

(with r_c chosen as the covalent radius of atom I) is a Coulombic interaction potential modified at short-range in such a way as to avoid spill-out of the electron density to nearby positively charged classical point charges.

Other potential sources of problems are possible incompatibilities between the QM and MM descriptions, such as imbalances in the electrostatic interactions that can lead to artificial preferences of e.g. substrate-QM, respectively substrate-MM interactions. Another problem is the consistent application of the classical exclusion rules for nonbonded interactions. In most force field definitions, nonbonded interactions (such as van der Waals and electrostatics) are not taken into account for nearby bonded neighbours. Such a selective neglect of particular pair interactions is not easily transferable to a many body QM description. A consistent approach is however possible via mapping of the many body electronic Hamiltonian to a pair additive point charge representation [34].

In summary, although most QM/MM implementations have now reached a mature state in which they have outgrown most of their children diseases, it

is always a good practice to familiarize oneself with a new QM/MM code by first running some simple test calculations to verify that the accuracy meets your expectations.

3.5 Newest Developments and Current Limitations

Nowadays, QM/MM simulations are a rather established tool for the investigation of adiabatic ground state reactions in complex environments [6]. Recently, QM/MM approaches have also been extended to the description of electronically excited states [19, 48, 55, 65] which enables the investigation of photochemical reactions, e.g. in photoactive proteins, and photochemically linked substrate target interactions. In these schemes, the excited states are either described via multiconfigurational-wavefunction-based quantum chemical methods [19] or excited state extensions of density functional theory [56]. Whereas the former are still limited to fairly small systems and relatively small basis sets, which can compromise accuracy, the latter are also extendable to fairly large systems (of the order of a few hundred atoms). Among the DFT based approaches, there are restricted open shell Kohn Sham (ROKS) implementations for the treatment of the first excited singlet states [17] and full time-dependent density functional theory (TDDFT) treatments for a general investigation of excited states [24]. TDDFT is most of the time used in the linear response (LR) approximation. The basic quantity in LR-TDDFT is the density-density response function

$$\chi(\mathbf{r}, t, \mathbf{r}', t') = \left. \frac{\delta n(\mathbf{r}, t)}{\delta v_{ext}(\mathbf{r}', t')} \right|_{v_0} \tag{37}$$

which relates the first order density response $n_1(\mathbf{r}, t)$ to the applied perturbation $v_1(\mathbf{r}, t)$

$$n_1(\mathbf{r}, t) = \int d^3\mathbf{r}' dt' \, \chi(\mathbf{r}, t, \mathbf{r}', t') \, v_1(\mathbf{r}', t') \tag{38}$$

where $v_0(\mathbf{r})$ is the ground state KS potential and $v_{ext}(\mathbf{r}) = v_0(\mathbf{r}) + v_1(\mathbf{r})$.

The response function for the physical system of interacting electrons, $\chi(\mathbf{r}, t, \mathbf{r}', t')$, can be related to the computationally more advantageous Kohn-Sham (KS) response, $\chi_s(\mathbf{r}, t, \mathbf{r}', t')$, and the problem of finding excitation energies of the interacting system reduces to a search for the poles of the response function.

In addition to the computation of excited state properties of a molecular system (excitation energies, densities and related properties), LR-TDDFT also enables the calculation of nuclear forces in the excited state. In the framework of the Lagrangian method [24], these forces are computed as derivatives of the total energy of the excited state with respect to the nuclear positions. The implementation of excited state forces within an MD package allows therefore the efficient calculation of trajectories on an excited state surface at a modest

computational cost. Such TDDFT/MM simulations have been successfully applied in many cases [48] but also failures due to current limitations of TDDFT for e.g. charge-transfer excitations and double excitations are known [65]. Currently, intense development efforts are ongoing for improvements of (TD)DFT, such as e.g. the inclusion of self interaction corrections (SIC) [67, 70].

A further recent QM/MM extension is the possibility to treat nonadiabatic effects within e.g. a surface hopping algorithm, based on Landau-Zener theory [66]. In this approach, the probability of changing adiabatic state is calculated from the energy gap between the two adiabatic potential energy surfaces (PESs) and the derivatives of the PES with respect to the reaction coordinate.

Latest QM/MM developments also allow access to many molecular properties such as e.g. optical spectra [63] (as mentioned above), and NMR chemical shifts [58] etc., that are useful to make direct contact with experimental data and facilitate verification of the simulations results for complex systems.

The most stringent current limitation is the relatively short time scale accessible via QM/MM simulations of the order of tens of picoseconds (for first principles based QM/MM) which severely restricts the accuracy of time-averaged properties, such as binding free energies. Possible remedies for this problem are the use of semiempirical methods that allow sampling for hundreds of picoseconds [53] and enhanced sampling approaches such as multiple time step sampling of QM and MM parts [77], or the use of classical (atom based) [35, 68]) or electronic [20, 69] bias potentials or by exploiting a linear response approximation with respect to a reference potential [75].

4 Covalent Drug-Target Binding: Examples

In this section we present some illustrative Car-Parrinello QM/MM examples in which covalent binding plays a role for DNA- and protein-drug recognition.

Drugs Targeting DNA

The anticancer drugs duocarmycins target adenine nucleobases of DNA (Fig. 5). We have carried out QM/MM calculations to investigate the alkylation reaction, which involves the cyclopropyl moiety of the drug (C13 in Fig. 5) and N3 of adenine **A** in the d(pGpApCpTpApApTpTpGpApC)- d(pGpTpCpApApTpTp**A**pGpTpC) oligonucleotide, for which structural information is available [14]. To the best of our knowledge, this has been the first study of this class of systems including the entire biomolecular frame, the solvent and temperature effects. Our calculations show that (i) DNA polarizes the drugs significantly and renders them more reactive with respect to the reaction in water (ii) the nature of the condensed rings in the drugs affects largely its intrinsic reactivity [62].

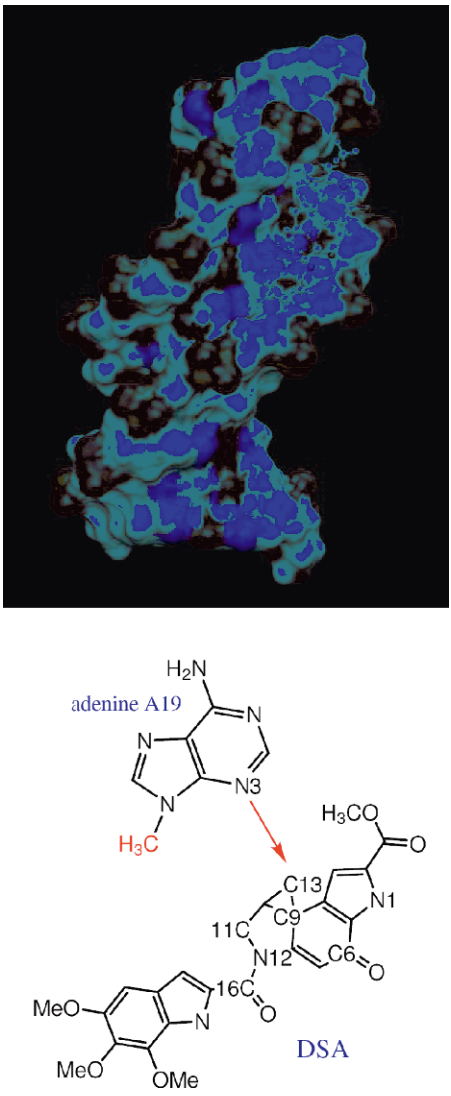


Fig. 5. DNA/duocarmycin complex. *Top:* Solvent accessible surface of the d(pGpApCpTpApApTpTpGpApC)-d(pGpTpCpApApTpTp**A**pGpTpC) double helix with electrostatic color code (light grey, negative potential, dark grey, positive potential). **Adenine 19** forms a covalent bond to a duocarmycin derivative (DSA), which is accommodated into the minor groove of DNA. The structure of the adduct has been experimentally determined at high resolution [14]. *Bottom*: Chemical structure of DSA and its target adenine [62]. DSA binding to DNA involves the formation of a covalent bond between the carbon atom C13 and adenine atom N3. Both reactive moieties have been included in the QM part of the QM/MM calculations. The methyl group colored in light grey is connected to the rest of the biomolecular frame, which is instead treated classically

Several anticancer drugs targeting DNA contain transition metal ions. The most widely used is a platinum derivative, cisplatin, (cis-diamminedichloroplatinum(II), Fig. 6), which treats a variety of cancer diseases in the clinics. Its beneficial effects arise from its binding to DNA nucleobases, preferentially to two N7 atoms belonging to two adjacent guanines. The platinated lesion induces an overall curvature of the DNA double helix towards the major groove associated to an opening of the minor groove. The platinated DNA moieties bind to a specific protein domain, the High Mobility Group (HMG) domains of proteins, impeding replication and cell repair processes, leading eventually to cell death [29, 39, 54, 73, 78].

As the stereochemical features of the platinated lesion are dictated by the electronic structure at the metal center [5], force-field based simulations of Pt adducts are far from trivial, making a QM/MM treatment crucial to correctly describe their properties. Our QM/MM simulations [81] have shown that the structure of the platinated DNA dodecamer rearranges significantly towards structural determinants of the solution structure as obtained by NMR spectroscopy [18]. The calculated ^{195}Pt chemical shifts of the QM/MM structure relative to cisplatin in water are in qualitative agreement with the experimental data [2]. The QM/MM structure of the platinated/DNA HMG complex, on the other hand, remains rather similar to the X-ray structure [49] consistent with its relatively low flexibility. Docking of $[Pt(NH_3)_2]^{2+}$ onto DNA in its canonical B-conformation causes a large axis bend and a rearrangement of DNA as experimentally observed by NMR in the platinated adducts [45].

More recently, we have extended our methodology to new and promising Pt-based drugs such as the dinuclear Pt-containing compounds [26–28] shown in Fig. 6. These have been suggested to be able to overcome the intrinsic and acquired cell resistance of cisplatin [45]. QM/MM simulations of their DNA adducts have shed light on the structural distortions that the drugs induce to the DNA duplex [45]. Our calculations show that the drugs do not provoke any kink upon binding to the double-stranded DNA, suggesting that they may act with a mechanism different than that of cisplatin. The accuracy of our calculations is established by a comparison with the NMR data for the corresponding complex.

Enzyme Inhibitors

HIV-1 Aspartyl Protease (AP) is a major target for anti-AIDS therapy (see, e.g. [7, 9, 50]). HIV-1 AP Food and Drug Administration (FDA)-approved drugs are non-covalently bound inhibitors that rely on their chemical similarity with the enzymatic transition state, namely, the presence of a hydroxyl group hydrogen-bonding the Asp dyad. Their affinity is caused mostly by the entropy associated with the transfer of partially ordered water in the active site of the enzyme to the bulk solvent (Fig. 1) [50]. The effectiveness of these drugs is continuously challenged by the emergence of protein variants that cause resistance and by severe side effects. Thus, a large effort is devoted to

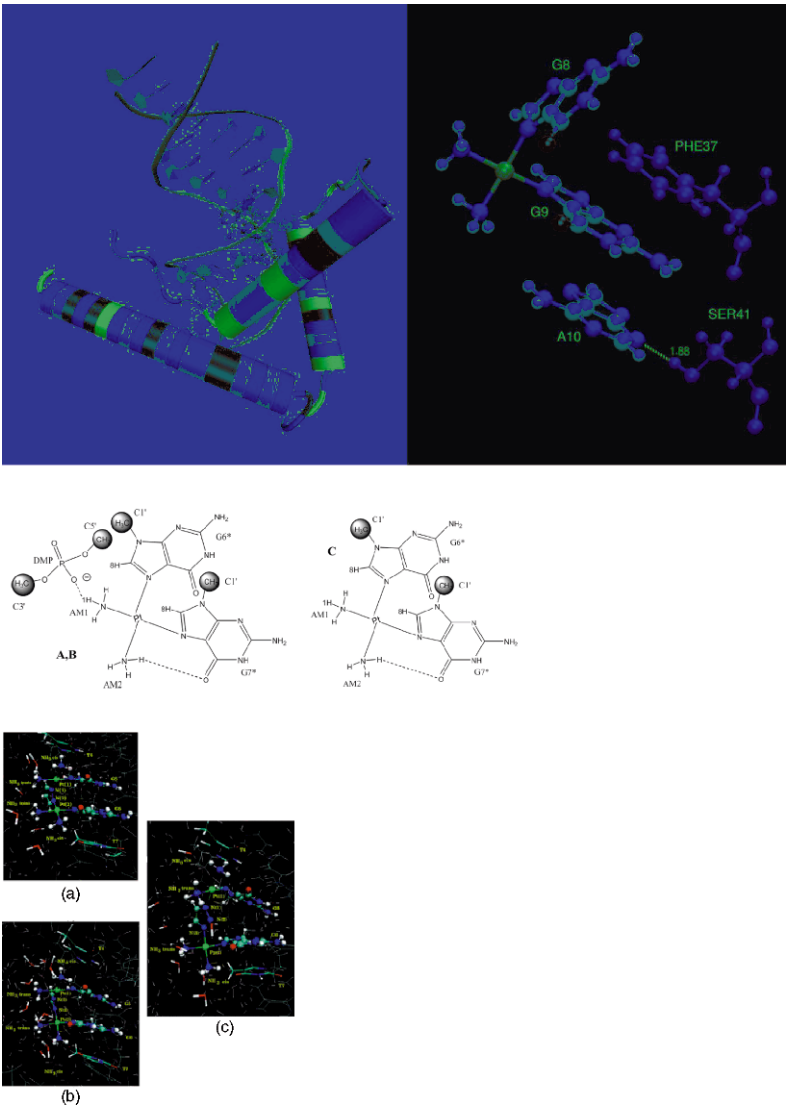


Fig. 6. Cisplatin binding to DNA. *Top*: Molecular recognition between a platinated DNA fragment and the HMG A protein. The cisplatin complex induces a distortion between two guanine residues (G8 and G9) which in turn interact with Phe37 and Ser41 in the protein. Notice the square planar geometry of cisplatin, which is dictated by electronic effects [5]. *Middle:* QM regions used in the calculations. Hydrogen bonds are indicated with dotted lines. Capped atoms are indicated by gray spheres.) *Bottom*: Covalent binding of new-generation Pt-based drugs to d(CpTpCpTp**G**p**G**pTpCpTpCp), where **G** are the guanines bound to the Pt atom. The drugs are: [{*cis*-Pt(NH_3)$_2$}$_2$ (μ-OH) (μ-pz)] (NO_3)$_2$ (with pz = pyrazolate)) and [{*cis*-Pt(NH_3)$_2$}$_2$ (μ-OH) (μ-1,2,3-ta-N(1),N(2))](NO_3)$_2$ (with ta = 1,2,3-triazolate)

identify inhibitors that feature not only positive entropy of binding but also form highly stabilizing interactions. One class of these inhibitors is that of the phosphinate and phosphonate derivatives (Fig. 7). In these compounds, the PO2 group is a noncleavable transition state analogue that binds at the AP active site. These inhibitors form hydrogen bonds in which the distance between the hydrogen-bond donor and acceptor is as small as 2.5 Å. Thus, the high affinity for the enzyme might be caused by the presence of low barrier hydrogen bonds between the ligand and AP. LBHB's can clearly not be studied with standard force fields. We have used QM/MM methods to investigate the hydrogen-bonding pattern at the binding site of the complexes of human immunodeficiency virus type-1 AP and the eukaryotic endothiapepsin and penicillopepsin. Our calculations are in fair agreement with the NMR data available for endothiapepsin [9] and show that the most stable active site configuration is the diprotonated, negatively charged form. Because of different interactions present at the active site, in the viral complex both protons are located at the catalytic Asp dyad, while in the eukaryotic complexes the proton

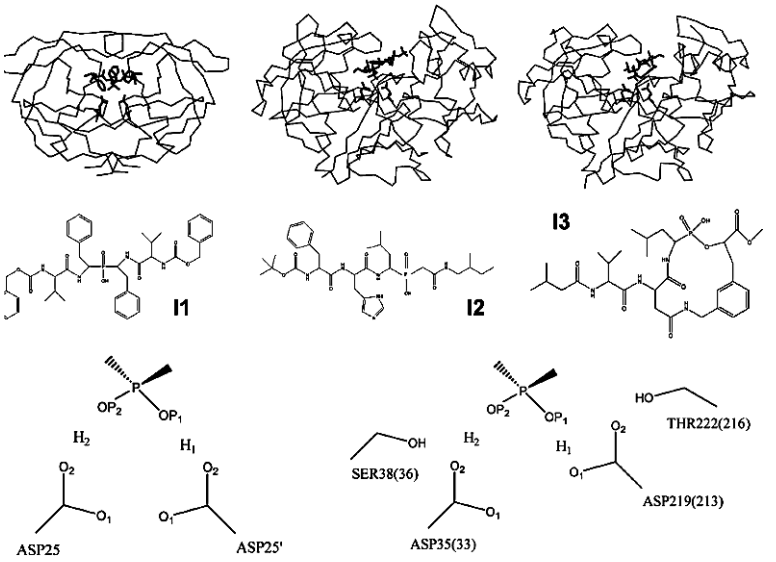


Fig. 7. *Top*: Structures of the HIV-1 AP (*left*), endothiapepsin (*middle*), and penicillopepsin (*right*) in complex with phosphorous containing inhibitors (PDB entries 1GVW, 1BXO, 1HOS). The viral and eukaryotic proteins feature a different fold: HIV-1 AP is a homodimer with the catalytic Asp dyad located on two loops at the monomer-monomer interface. Endothiapepsin and penicillopepsin are monomers composed of two asymmetric lobes, with the catalytic Asp dyad located at the lobe interface, the so-called "pepsin-like" protease. Thin line, protein C trace; thick line, inhibitor and catalytic Asp dyad. *Middle:* HIV-1 AP (*left*), endothiapepsin (*middle*), and penicillopepsin (*right*) inhibitors. *Bottom*: Close up on the cleavage site of the viral and penicillopepsin complexes

shared by the closest oxygen atoms is located at the phosphinic/phosphonic group.

As a final remark, notice that we do not discuss here investigations of enzymatic reactions, which from one hand clearly also require quantum mechanics, and from the other can also provide practical pharmaceutical applications for transition state analogs. In fact, Car-Parrinello QM/MM simulations of enzymatic systems have been reviewed recently [6] and have not been mentioned here.

5 Concluding Remarks

Extremely important drugs such as cisplatin bind covalently to their target. In that case, both the kinetics and the thermodynamics depend dramatically on the formation of the covalent bond and therefore the binding process cannot be described with standard force fields, as commonly practiced for non-covalent complexes. In addition, the investigation of enzymatic reactions, which requires quantum mechanics, can also provide practical guidance for the design of transition state analogs. How such processes are affected by the cellular environment is still mostly unknown [40] and will require much more sophisticated models than those used so far. Therefore, there is no doubt that scientists oriented towards electronic structure calculations will be more and more called upon to address these key topics, which are not investigated in traditional computational pharmacology.

Acknowledgments

The authors wish to thank all their coauthors of the papers referred to in this review, Michael L. Klein and Michele Parrinello for their continued support over the last fifteen years. Paolo Ruggerone and Ivano Tavernelli are gratefully acknowledged for a reading of the manuscript and many inspiring discussions. Last but not least, we would like to express our gratitude to Giovanni Ciccotti for his never ceasing effort to force us to write this review.

References

1. X. Assfeld and J. L. Rivail (1996) Quantum chemical computations on parts of large molecules: The ab initio local self consistent field method. *Chem. Phys. Lett.* **263**, p. 100
2. P. B. Daniel, A. L. Christopher, J. L. Stephen, D. P. Bancroft, C. A. Lepre, and S. J. Lippard (1990) Platinum-195 NMR kinetic and mechanistic studies of cis- and trans-diamminedichloroplatinum(II) binding to DNA. *J. Am. Chem. Soc.* **112**, p. 6860

3. R. Bauernschmitt and R. Ahlrichs (1996) Treatment of electronic excitations within the adiabatic approximation of time dependent density functional theory. *Chem. Phys. Lett.* **256**, p. 454
4. R. Car and M. Parrinello (1985) Unified approach for molecular dynamics and density-functional theory. *Phys. Rev. Lett.* **55**, p. 2471
5. C. Paolo, P. Carloni, A. Wanda, A. Andreoni, H. Jurg, C. Hutter, C. Alessandro, A. Curioni, G. Paolo, P. Giannozzi, P. Michele and M. Parrinello (1995) Structure and Bonding in Cisplatin and other Pt(II) Complexes. *Chem. Phys. Lett.* **234**, pp. 50–56
6. P. Carloni, U. Rothlisberger, and M. Parrinello (2002) The role and perspective of ab initio molecular dynamics in the study of biological systems. *Acc. Chem. Res.* **35(6)**, pp. 455–64
7. A. Carr (2003) Toxicity of antiretroviral therapy and implications for drug development. *Nat. Rev. Drug Discovery* **2**, pp. 624–634
8. D. A. Case, T. E. Cheatham 3rd., T. Darden, H. Gohlke, R. Luo, K. M. Merz, Jr., A. Onufriev, C. Simmerling, B. Wang, and R. J. Woods (2005) The Amber biomolecular simulation programs. *J. Comput. Chem.* **26(16)**, pp. 1668–88
9. L. Coates, P. T. Erskine, M. P. Crump, S. P. Wood, and J. B. Cooper (2002) Five atomic resolution structures of endothiapepsin inhibitor complexes: implications for the aspartic proteinase mechanism. *J. Mol. Biol.* **318**, pp. 1405–1415
10. W. D. Cornell, P. Cieplak, C. I. Bayly, K. M. Gould, K. M. Merz, D. M. Ferguson, G. L. Seibel, U. C. Singh, P. K. Weiner, and P. A. Kollmann (1995) A 2nd generation force-field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.* **117**, p. 5179
11. CPMD, IBM Corp. (1990–2001) Copyright MPI fuer Festkoerperforschung Stuttgart (1997–2001) Available online at: `http://www.cpmd.org`
12. P. M. Dal, S. Raugei, P. Carloni, and M. L. Klein (2005) Solute-solvent charge transfer in aqueous solution. *Chem. Phys. Chem.* **6(9)**, pp. 1715–8
13. A. G. Donchev, V. D. Ozrin, M. V. Subbotin, O. V. Tarasov, and V. I. Tarasov (2005) A quantum mechanical polarizable force field for biomolecular interactions. *Proc. Natl. Acad. Sci. USA* **102**(22): pp. 7829–34
14. P. S. Eis, J. A. Smith, J. M. Rydzewski, D. A. Case, D. L. Boger, and W. J. Chazin (1997) High Resolution Solution Structure of a DNA Duplex Alkylated by the Antitumor Agent Duocarmycin SA. *J. Mol. Biol.* **272**, pp. 237–252
15. M. J. Field, P. A. Bash, and M. Karplus (1990) A combined quantum-mechanical and molecular mechanical potential for molecular-dynamics simulations. *J. Comp. Chem.* **11**, pp. 700–733
16. F. Alan (1998) Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding. W. H. Freeman 3rd Revised edition
17. I. Frank, J. Hutter, D. Marx, and M. Parrinello (1998) Molecular dynamics in low-spin excited states. *J. Chem. Phys.* **108**, p. 4060
18. A. Gelasco and S.J. Lippard (1998) NMR solution structure of a DNA dodecamer duplex containing a cis-diammineplatinum(II) d(GpG) intrastrand cross-link, the major adduct of the anticancer drug cisplatin. *Biochemistry* **37**, pp. 9230–9239
19. G. Groenhof, M. Bouxin-Cademartory, B. Hess, S. P. De Visser, H. J. C. Berendsen, M. Olivucci, A. E. Mark, and M. A. Robb (2004) Photoactivation of the photoactive yellow protein: Why photon absorption triggers a trans-to-cis isomerization of the chromophore in the protein. *J. Am. Chem. Soc.* **126,** p. 4228

20. L. Guidoni and U. J. Rothlisberger (2005) Scanning reactive pathways with orbital biased molecular dynamics. *J. Chem. Theo. Comp.* **1**, p. 554
21. C. Hartwigsen, S. Goedecker, and J. Hutter (1998) Relativistic separable dual-space Gaussian pseudopotentials from H to Rn. *Phys. Rev. B* **58**, p. 3641
22. P. Hohenberg and W. Kohn (1964) Inhomogeneous electron gas. *Phys. Rev. B* **136**, p. 864
23. P. Hünenberger (2000) Optimal charge-shaping functions for the particle-particle-particle-mesh (P3M) method for computing electrostatic interactions in molecular simulations. *J. Chem. Phys.* **113**, p. 10464
24. J. Hutter (2003) Excited state nuclear forces from the Tamm-Dancoff approximation to time-dependent density functional theory within the plane wave basis set framework. *J. Chem. Phys.* **118**, p. 3928
25. P. A. Kollman and D. A. Case (2003) Burger's Medicinal Chemistry and Drug Discovery, D. J. Abraham, Ed. **1**, pp. 169–185
26. S. Komeda, S. Bombard, S. Perrier, J. Reedijk, and J. Kozelka (2003) Kinetic study of azole-bridged dinuclear platinum(II) complexes reacting with a hairpin-stabilized double-stranded oligonucleotide. *J. Inorg. Biochem.* **96**, pp. 357–366
27. S. Komeda, H. Yamane, M. Chikuma, and J. Reedijk (2004) A kinetic study on the reactions of azolato-bridged dinuclear platinum(II) complexes with guanosine 5'-monophosphate. *Eur. J. Inorg. Chem.* **24**, pp. 4828–4835
28. K. Seiji, L. Martin, L. Anthony, Y. Yasuyuki Spek, S. Takaji, C. Masahiko, and R. Jan (2002) A Novel Isomerization on Interaction of Antitumor-Active Azole-Bridged Dinuclear Platinum(II) Complexes with 9-Ethylguanine. Platinum(II) Atom Migration from N2 to N3 on 1,2,3-Triazole. *J. Am. Chem. Soc.* **124(17)**, pp. 4738–4746
29. K. R. Barnes, S. J. Lippard (2004) Cisplatin and related anticancer drugs: recent advances and insights. *Met. Ions Biol. Syst.* **42**, 143–177
30. M. Kontoyianni, L. M. McClellan, and G. S. Sokol (2004) Evaluation of docking performance: comparative data on docking algorithms. *J. Med. Chem.*, **47**, pp. 558–65
31. I. D. Kuntz and Jr. W. Kauzmann (1974) Hydration of proteins and polypeptides. *Adv. Protein Chem.* **28**, pp. 239–345
32. A. Laio, J. VandeVondele, and U. Rothlisberger (2002) A Hamiltonian electrostatic coupling scheme for hybrid Car-Parrinello molecular dynamics simulations. *J. Chem. Phys.* **116**, p. 6941
33. A. Laio, J. VandeVondele, and U. Rothlisberger (2002) D-RESP: Dynamically generated electrostatic potential derived charges from QM/MM simulations. *J. Phys. Chem. B* **106**, p. 7300
34. A. Laio, F. Gervasio, M. Sulpizi, and U. Rothlisberger (2004) A variational definition of electrostatic potential derived charges. *J. Phys. Chem.* **108**, p. 7983
35. A. Laio and M. Parrinello (2003) Escaping free-energy minima. *Proc. Nat. Ac. Sc.* **99**, 12562
36. L. D. Landau (1932) *Phys. Z. Sowjetunion* **2**, p. 46; C. Zener (1932) Non-adiabatic crossing of energy levels. *Proc. R. Soc.* **A137**, p. 696
37. T. Laino, F. Mohamed, A. Laio, and M. Parrinello (2005) An efficient real space multigrid QM/MM electrostatic coupling. *J. Comp. Theor. Chem.* **1**, 1176
38. H. Lin and D. G. Truhlar, Proceedings of the 10th Electronic Computational Chemistry Conference. http://comp.chem.umn.edu/Truhlar/hilight/QMMMreview05.htm

39. B. C. Lippert (1999) Chemistry and Biochemistry of a leading anticancer drug. *Viley-VCH*, ed., **1**
40. A. Lesk (2004) Introduction to Protein Science. *Architecture, Function, and Genomics*. Oxford University Press
41. S. G. Louie, S. Froyen, and M. L. Cohen (1982) Non-linear ionic pseudopotentials in spin-density-functional calculations. *Phys. Rev. B* **26**, p. 1738
42. A. D. MacKerell, Jr. J. Wiorkiewicz-Kuczera, and M. Karplus (1995) An all-atom empirical energy function for the simulation of nucleic acids. *J. Am. Chem. Soc.* **117**, pp. 11946–11975
43. A. D. MacKerell Jr. and N. Banavali (2000) All-Atom Empirical Force Field for Nucleic Acids: 2) Application to Molecular Dynamics Simulations of DNA and RNA in Solution. *J. Comput. Chem.* **21**, pp. 105–120
44. A. D. MacKerell Jr., D. Bashford, M. Bellott, R. L. Dunbrack Jr., J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, III, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorkiewicz-Kuczera, D. Yin, and M. Karplus (1998) All-atom empirical potential for molecular modeling and dynamics Studies of proteins. *J. Phys. Chem. B* **102**, pp. 3586–3616
45. A. Magistrato, P. Ruggerone, K. Spiegel, P. Carloni, and J. Reedijk (2006) Binding of Novel Azole-Bridged Dinuclear Platinum(II) Anticancer Drugs to DNA: Insights from Hybrid QM/MM Molecular Dynamics Simulations. *J. Phys. Chem. B.* **110**, pp. 3604–3613
46. D. Marx and J. Hutter (2000) Modern Methods and Algorithms of Quantum Chemistry. Ed. J. Grotendorst. John von Neumann Institute for Computing, Jülich, NIC Series **1**, p. 301
47. F. Maseras and K. Morokuma (1995) IMOMM - A new integrated ab initio plus molecular mechanics geometry optimization scheme of equilibrium structures and transition states. *J. Comput. Chem.* **16**, p. 1170. Implementation of an adaptive umbrella sampling
48. M. E. Moret, E. Tapavizca, L. Guidoni, U. F. Röhrig, M. Sulpizi, I. Tavernelli, and U. Rothlisberger (2005) Quantum mechanical / molecular mechanical (QM/MM) Car-Parrinello simulations in excited states. *CHIMIA* **59**, p. 493
49. U. M. Ohndorf, M. A. Rould, Q. He, C. O. Pabo, and S. J. Lippard (1999) Basis for recognition of cisplatin-modified DNA by high-mobility-group proteins. *Nature*, **399**, pp. 708–712
50. H. Ohtaka and E. Freire (2005) Adaptive inhibitors of the HIV-1 protease. *Prog. Biophys. Mol. Biol.* **88**, pp. 193–208
51. L. Pauling (1988) *General Chemistry*, Dover Publications, NY.
52. M. Petersilka, U. J. Gossmann, E. K. U. Gross (1996) Excitation energies from time-dependent density-functional theory. *Phys. Rev. Lett.* **76**, p. 1212
53. R. Rajamani, K. J. Naidoo, and J. L. Gao (2003) Implementation of an adaptive umbrella sampling method for the calculation of multidimensional potential of mean force of chemical reactions in solution. *J. Comp. Chem.* **24**, p. 1775
54. J. Reedijk, (2003) New clues for platinum antitumor chemistry: Kinetically controlled metal binding to DNA. *Proc. Natl. Acad. Sci. USA* **100**, pp. 3611–3616
55. U. F. Röhrig, I. Frank, J. Hutter, and U. Rothlisberger (2003) QM/MM Car-Parrinello molecular dynamics study of the solvent effects on the ground state and on the first excited singlet state of acetone in water. *Chem. Phys. Chem.* **4**, p. 1177

56. E. Runge and E. K. U. Gross (1984) Density-functional theory for time-dependent systems. *Phys. Rev. Lett.* **53**, p. 997
57. W. R. P. Scott, P. H. Hünenberger, I. G. Tironi, A. E. Mark, S. R. Billeter, J. Fennen, A. E. Torda, T. Huber, P. Krüger, and W. F. van Gunsteren (1999) The GROMOS biomolecular simulation program package. *J. Phys. Chem. A* **103**, p. 3596
58. D. Sebastiani and U. Rothlisberger (2003) Advances in Density Functional Based Modelling Techniques: Recent Extensions of the Car-Parrinello Approach. In P. Carloni, F. Alber 'Medicinal Quantum Chemistry', Series: Methods and Principles in Medicinal Chemistry. Series Editors: R. Mannhold, H. Kubiny, G. Folkers, Wiley-VCH, Weinheim
59. P. Sherwood (2000) Modern Methods, and Algorithms of Quantum Chemistry. John von Neumann Institute for Computing **1** p. 257
60. T. Simonson, G. Archontis, and M. Karplus (2002) Free energy simulations come of age: protein-ligand recognition. *Acc. Chem. Res.*, pp. 430–7, Review
61. U. C. Singh and P. A. Kollman (1986) A combined ab initio quantum mechanics and molecular mechanics method for carrying out simulations on complex molecular systems - applications to the CH3Cl $+Cl-$ exchange reaction and gas phase protonation of polyethers. *J. Comp. Chem.* **7**, p. 718
62. K. Spiegel, U. Rothlisberger, and P. Carloni (2006) Duocarmycins Binding to DNA Investigated by Molecular Simulation. *J. Phys. Chem. B* **110**, pp. 3647–3660
63. S. Sulpizi, U.F. Röhrig, J. Hutter, and U. Rothlisberger (2005) Optical properties of molecules in solution via hybrid TDDFT/MM simulations. *Intl. J. Quant. Chem.* **101**, p. 671
64. M. Svensson, S. Humbel, R. D. J. Froese, T. Matsubara, S. Sieber, and K. Morokuma (1996) ONIOM: A multilayered integrated MO+MM method for geometry optimizations and single point energy predictions. A test for Diels-Alder reactions and Pt(P(t-Bu)3)2+ + H2 oxidative addition. *J. Phys. Chem.* **100**, p. 19357
65. I. Tavernelli, U. F. Röhrig, U. Rothlisberger (2005) Molecular dynamics in electronically excited states using time-dependent density functional theory. *Mol. Phys.* **103**, p. 963
66. E. Tapavizca, I. Tavernelli, and U. Rothlisberger Ab initio molecular dynamics with quantum transitions using time-dependent density functional theory. (submitted)
67. I. Tavernelli Intramolecular electron transfer in bis(methylene)adamantyl radical cation: a finite temperature study using SIC corrected density functional theory. (submitted)
68. J. VandeVondele and U. Rothlisberger (2000) Efficient multidimensional free energy calculations for ab initio molecular dynamics using classical bias potentials. *J. Chem. Phys.* **113**, p. 4863
69. J. VandeVondele and U. Rothlisberger (2002) Accelerating rare reactive events by means of a finite electronic temperature. *J. Am. Chem. Soc.* **124**, p. 8163
70. J. VandeVondele and M. Sprik (2005) A molecular dynamics study of the hydroxyl radical in solution applying self interaction - corrected density functional methods. *Phys. Chem. Chem. Phys.* **7**, p. 1363
71. W. F. van Gunsteren, P. M. King, and A. E. Mark (1994) Fundamentals of drug design from a biophysical viewpoint. *Q Rev. Biophys. Dec.* **27(4)**, pp. 435–81

72. O. A. von Lilienfeld, I. Tavernelli, U. Rothlisberger, and D. Sebastiani (2005) Variational optimization of effective atom-centered potentials for molecular systems. *J. Chem. Phys.* **122**, 014113
73. D. Wang and S. J. Lippard (2005) Cellular processing of platinum anticancer drugs. *Nat. Rev. Drug. Discov.* **4**, pp. 307–320
74. A. Warshel and M. Levitt (1976) Theoretical studies of enzymic reactions - dielectric, electrostatic and steric stabilization of carbonium - ion in reaction of lysozyme. *J. Mol. Biol.* **103**, p. 227
75. A. Warshel (2006) Towards accurate ab initio QM/MM calculations of free-energy profiles of enzymatic reactions. *J. Phys. Chem. B* **110**, p. 2934
76. T. A. Wesolowski (1993) Frozen density-functional approach for ab initio calculations of solvated molecules. *J. Phys. Chem.* **97**, p. 8050
77. T. K. Woo, P. Margl, P. E. Bloechl, and T. Ziegler (2002) Sampling phase space by a combined QM/MM ab initio Car-Parrinello molecular dynamics method with different (multiple) time steps in the quantum mechanical (QM) and the molecular mechanical (MM) domains. *J. Phys. Chem. A* **106**, p. 1173
78. A. D. Yang, S. S. G. E. van Boom, J. Reedijk, J. H. van Boom, and A. H.-J. Wang (1995) Structure and Isomerization of an Intrastrand Cisplatin-Cross-Linked Octamer DNA Duplex by NMR Analysis. *Biochem.* **34**, pp. 12912–12920
79. D. Yarne, M. E Tuckerman, and G. J. Martyna (2001) A dual length scale method for plane-wave-based simulations studies of chemical systems modeled using mixed ab initio/empirical force field descriptions. *J. Chem. Phys.* **115**, p. 3531
80. Z. Zhu, D. I. Schuster, and M. E. Tuckerman (2003) Molecular dynamics study of the connection between flap closing and binding of fullerene-based inhibitors of the HIV-1 protease. *Biochemistry* **42**, pp. 1326–33
81. K. Spiegel, U. Rothlisberger, and P. Carloni (2004) Cisplatin binding to DNA oligomers from hybrid Car-Parrinello/molecular dynamics simulations. *J. Phys. Chem. B* **108**(8): pp. 2699–2707
82. W. Kohn, and L. J. Sham (1965) Self-consistent equations including exchange and correlation. *Phys. Rev.* **140**, p. 1133

Redox Free Energies from Vertical Energy Gaps: Ab Initio Molecular Dynamics Implementation

J. Blumberger[1] and M. Sprik[2]

[1] Center for Molecular Modeling and Department of Chemistry, University of Pennsylvania, 231 S. 34th Street Philadelphia, PA 19104-6323
[2] Department of Chemistry, University of Cambridge, Cambridge CB2 1EW, United Kingdom



Michiel Sprik

J. Blumberger and M. Sprik: *Redox Free Energies from Vertical Energy Gaps: Ab Initio Molecular Dynamics Implementation*, Lect. Notes Phys. **704**, 481–506 (2006)
DOI 10.1007/3-540-35284-8_18 © Springer-Verlag Berlin Heidelberg 2006

1 Introduction 483

2 Simulation of Electron Transfer 484

2.1 The Diabatic Energy Gap as Reaction Coordinate 484
2.2 Reaction Free Energies 487
2.3 Free Energy Perturbation Method 489
2.4 Relation to Marcus Theory 491

3 Redox Half Reactions 493

3.1 Ab Initio MD Considerations 493
3.2 Parallel to Electrode Reactions 495

4 Applications 496

4.1 Overview 496
4.2 Two Examples: The Ru and Ag Aqua Cations 497
4.3 Conclusion 503

References 504

1 Introduction

The development of methods for the exploration of reaction paths in condensed molecular systems (solutions and biopolymers) and the computation of the corresponding reaction free energies and kinetic parameters remains at the center of research in computational chemistry. Much has happened in recent years. It is the subject of a good number of the chapters in this book, which give an up to date overview of the enormous progress that has been made. We mention the development of the transition path sampling method by the group of Chandler at Berkeley (see the chapters by Dellago, Bolhuis and Geissler, where also the references to the original literature can be found). An alternative approach with a somewhat different purpose and scope is the metadynamics method developed by the Parrinello group (see the chapter by Laio and Parrinello). Transition path sampling and metadynamics studies to date have focused mostly on dynamical processes which never leave the adiabatic ground state potential energy surface (PES). However barriers for chemical reactions often coincide with an avoided crossing, or, alternatively, can be seen as the result of the coupling between two intersecting diabatic surfaces (see Fig. 1). The diabatic perspective offers certain advantages. This applies in particular to activation energies with a strong solvent contribution. An instructive example of such a reaction is electron transfer (ET). For outer sphere transfer the barrier is almost 100 percent due to rearrangement of the solvent polarization. This observation is a key idea in the Marcus theory of electron transfer [1–4]. In the original formulation of the theory [1] the polarization was described by the linear response of a dielectric continuum. How to quantify solvent polarization by a microscopic order parameter? Polarization is a highly collective quantity with a configurational component (the orientation of molecules) and electronic component (induced polarization).

While it should be possible to find a ground state observable incorporating all relevant aspects of the solvent response to ET, the diabatic picture at the core of Marcus theory suggests that the vertical energy gap between the two

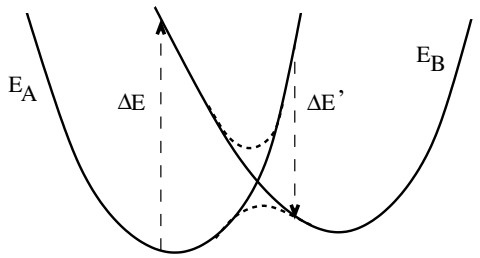


Fig. 1. Two intersecting diabatic potential energy surfaces E_{A} and E_{B} (*solid curves*). The *dashed curves* are the corresponding adiabatic surfaces. The *dashed arrows* indicate two examples of the diabatic vertical energy gap defined in (3), the gap on the left (ΔE) is positive and the one on the right ($\Delta E'$) negative

diabatic surfaces is a more suitable quantity for this purpose (see e.g. [2]). The first to use this idea in a simulation of ET reactions was Warshel [5]. A key step in his approach is the observation that the energy gap is a particularly convenient reaction coordinate for the application of free energy perturbation methods [5–11]. This enabled Warshel and coworkers to calculate not only the reaction and activation free energy from their molecular dynamics (MD) trajectories but also the reorganization free energy of Marcus theory (see below). The latter quantity, usually written as λ, is the free energy cost of deforming an equilibrium atomic configuration in the reactant state to an equilibrium configuration of the product state while staying on the diabatic PES of the reactant (so without making an electronic transition). Reorganization free energy is therefore a new quantity unique to the diabatic (two surface) picture. As the many successful applications of Marcus theory illustrate, it is a most powerful concept for the analysis and understanding of chemical reactions.

Marcus theory has inspired a rich production of simulation studies aiming to validate its assumptions or focusing on applications to specific model systems. References [6,7] and [12–27] are a selection of the many papers that have appeared since Warshel's pioneering 82 paper [5]. These studies are based on classical and semi-classical models. In a series of recent *ab initio* MD studies of redox half reactions involving transition metal coordination complexes [28–33] and organic molecules [34–36] we have attempted to implement similar methods in a density functional theory based MD (Car-Parrinello) [37] framework. A number of technical difficulties had (and some remain) to be resolved. The aim of this chapter is to discuss some of the background which our approach shares with earlier work.

2 Simulation of Electron Transfer

2.1 The Diabatic Energy Gap as Reaction Coordinate

For non-adiabatic ET treated in a semi-classical MD approximation the diabatic energy gap is the ideal reaction coordinate [5]. The electronic system in this approach is modeled by a two level system defined by a 2×2 Hamiltonian matrix

$$\mathbf{H} = \begin{pmatrix} E_{\mathrm{A}}\left(\mathbf{R}^N\right) & \gamma\left(\mathbf{R}^N\right) \\ \gamma\left(\mathbf{R}^N\right) & E_{\mathrm{B}}\left(\mathbf{R}^N\right) \end{pmatrix} \tag{1}$$

The two diagonal elements represent the diabatic PES of reactant A and product B where $\mathbf{R}^N$ are the atomic positions of (in principle) all N atoms in the system. For example in case of a charge separation reaction

$$\mathrm{D} + \mathrm{A} \rightarrow \mathrm{D}^+ + \mathrm{A}^- \tag{2}$$

the reactant state A corresponds to the donor-acceptor pair $\mathrm{D} + \mathrm{A}$ before the transfer and B to the product $\mathrm{D}^+ + \mathrm{A}^-$ after the transfer. The diabatic states

are coupled by the off-diagonal matrix element γ. This quantity depends on the overlap between the electron wavefunctions in reactant and product state and is often modelled by an exponent of the distance r_{DA} between donor and acceptor, $\gamma = \gamma_0 \exp(-\alpha r_{\mathrm{DA}})$, where α is a positive parameter with the dimension of inverse length. The vertical diabatic energy gap of the two level system of (1) is simply the difference of the diagonal matrix elements.

$$\Delta E(\mathbf{R}^N) = E_{\mathrm{B}}(\mathbf{R}^N) - E_{\mathrm{A}}(\mathbf{R}^N) \tag{3}$$

If the off-diagonal matrix element γ is small (weak coupling), for example because the distance between donor and acceptor in reaction (2) is large, the (absolute) diabatic energy gap ΔE of (3) can be identified with the adiabatic vertical excitation energy

$$\Delta E_{1\leftarrow 0}\left(\mathbf{R}^N\right) = \left[\left(E_{\mathrm{A}} - E_{\mathrm{B}}\right)^2 + 4\gamma^2\right]^{1/2} \tag{4}$$

where the index 0 denotes the ground state and 1 the (only) excited state and we have suppressed the dependence on $\mathbf{R}^N$ of the quantities on the rhs. Note, however, that ΔE of (3) can be both positive and negative, changing sign at the surface crossing. This is where, in the weak coupling limit, the radiationless transition between state A and B takes place. The vertical diabatic energy gap can therefore be used as a reaction coordinate assigning each atom configuration $\mathbf{R}^N$ either to reactant or product. Using first order time dependent perturbation theory arguments, Marcus separated the ET rate k_{ET} in a quantum transition probability proportional to the squared coupling parameter γ^2 and the Boltzmann exponent of the free energy G at the surface crossing

$$k_{\mathrm{ET}} = \kappa\gamma^2 \exp\left[-\Delta G\left(\Delta E = 0\right)/k_{\mathrm{B}}T\right] \tag{5}$$

The non-adiabatic "transition state" is unambiguously identified by the zero gap ($\Delta E = 0$). Equation (5), including the expression for the prefactor κ, can be derived using the Golden rule. The derivation is far from straightforward, requiring careful consideration of the classical limit of the atomic system (see for example [38]). An easier route is to start from a classical atomic system using a Landau Zener approach [39].

The factorization of the ET rate achieved in (5) allows us to compute the activation free energy $\Delta G\left(\Delta E = 0\right)$ treating the atoms as classical particles and make separate assumptions for the estimation of γ (or even not worry about it at all). This is the justification of the many classical force field model based simulations of non-adiabatic ET, which at first might seem somewhat of a contradiction. The vertical energy gap in the fully classical point charge model can be related to the electrostatic potentials at the site of the donor and acceptor ions. For the purpose of comparison of our Car-Parrinello results to classical model studies it is instructive to take a more detailed look at the vertical gap in the simple point charge approximation. The total energy in such a model can be written (in atomic units) as

$$U_{\mathrm{A}}\left(\mathbf{R}^N\right) = \frac{q_{\mathrm{D}}q_{\mathrm{A}}}{r_{\mathrm{DA}}} + \sum_{i}^{solvent}\left(\frac{q_{\mathrm{D}}q_i}{r_{\mathrm{D}i}} + \frac{q_{\mathrm{A}}q_i}{r_{\mathrm{A}i}}\right) + \sum_{i>j}^{solvent}\frac{q_iq_j}{r_{ij}} + \sum_{i>j}^{N} v_{ij}\left(r_{ij}\right) \quad (6)$$

where the charges and positions of the donor and acceptor ion are labeled by the subscript D respectively A and i counts solvent particles. v_{ij} is some atom-atom potential describing short range interactions between particles i and j. Summation for this pair interaction term runs over all particles N in the model (solutes plus solvent) where the donor particle D is now identified with the particle with index $i = 1$ and the acceptor with the particle indicated by index $i = 2$. If we assume that q_{D} and q_{A} are the charges of donor and acceptor in the reactant state we can interpret the energy U_{A} of (6) as the total energy before the transfer, hence the subscript A. The charge of donor and acceptor after the transfer will then be $q_{\mathrm{D}}+1$ respectively $q_{\mathrm{A}}-1$ and the total energy U_{B} of the product becomes

$$U_{\mathrm{B}}\left(\mathbf{R}^N\right) = \frac{\left(q_{\mathrm{D}}+1\right)\left(q_{\mathrm{A}}-1\right)}{r_{\mathrm{DA}}} + \sum_{i}^{solvent}\left(\frac{q_{\mathrm{D}}+1}{r_{\mathrm{D}i}} + \frac{q_{\mathrm{A}}-1}{r_{\mathrm{A}i}}\right)q_i + \sum_{i>j}^{solvent}\frac{q_iq_j}{r_{ij}} + \sum_{i>j}^{N} v_{ij}\left(r_{ij}\right) \quad (7)$$

Subtracting gives the vertical energy gap.

$$\Delta U = U_{\mathrm{B}} - U_{\mathrm{A}} = \frac{\left(q_{\mathrm{A}} - q_{\mathrm{D}} - 1\right)}{r_{\mathrm{DA}}} + \sum_{i}^{solvent}\left(\frac{1}{r_{\mathrm{D}i}} - \frac{1}{r_{\mathrm{A}i}}\right)q_i \quad (8)$$

$$= -\frac{1}{r_{\mathrm{DA}}} + \sum_{i\neq 1}^{N}\frac{q_i}{r_{\mathrm{D}i}} - \sum_{i\neq 2}^{N}\frac{q_i}{r_{\mathrm{A}i}} \quad (9)$$

All short range interaction cancel since the position of the particles are kept fixed during a vertical transfer. Recalling that the electrostatic potential at the site of a particle i can be expressed as

$$\Phi_i\left(\mathbf{R}^N\right) = \sum_{j\neq i}^{N}\frac{q_j}{r_{ij}} \quad (10)$$

we recognize in the last two terms of (9) the difference of the electrostatic potentials acting on the donor ($i = 1$) and acceptor ($i = 2$) and we can write the point charge gap as

$$\Delta U\left(\mathbf{R}^N\right) = -\frac{1}{r_{\mathrm{DA}}} + \Phi_{\mathrm{D}}\left(\mathbf{R}^N\right) - \Phi_{\mathrm{A}}\left(\mathbf{R}^N\right) = -\frac{1}{r_{\mathrm{DA}}} + IP_{\mathrm{D}} - EA_{\mathrm{A}} \quad (11)$$

Equation (11) is indeed equal to the vertical electron transfer excitation energy (4) in the classical point charge approximation (plus weak coupling limit). This is made explicit in the second identity, where IP_{D} is the ionization potential of the donor and EA_{A} the electron affinity of the acceptor. The first term is the direct "particle-hole" Coulomb contribution to the excitation energy.

2.2 Reaction Free Energies

The crucial step, as in any study of chemical reactivity based on reaction coordinates, is the computation of the free energy profile or potential of mean force (PMF). A formalism with separately defined potential energy surfaces for reactant A (see Sect. 2.1) and product B also yields two separate diabatic free energy profiles. So, to begin with the total free energy we define

$$A_M = -k_\mathrm{B}T \ln \Lambda^{-3N} \int \mathrm{d}\mathbf{R}^N \exp[-\beta E_M\left(\mathbf{R}^N\right)] \tag{12}$$

where $M = \mathrm{A}, \mathrm{B}$. E_A is the diabatic PES of state A, and E_B the PES of state B (3). Λ is the average thermal wavelength of the atoms defined as $\Lambda^{-3N} = \prod_j \lambda_j^{-3N_j}/N_j!$ with $\lambda_j = h/\sqrt{2\pi m_j k_\mathrm{B} T}$ the thermal wavelength of the nuclear species j. As usual $\beta^{-1} = k_\mathrm{B}T$ with k_B the Boltzmann constant and T the temperature. The reaction free energy change can then be related to a ratio of partition functions

$$\Delta A = A_\mathrm{B} - A_\mathrm{A} = -k_\mathrm{B}T \ln \frac{\int \mathrm{d}\mathbf{R}^N \exp(-\beta\, E_\mathrm{B}\left(\mathbf{R}^N\right))}{\int \mathrm{d}\mathbf{R}^N \exp(-\beta\, E_\mathrm{A}\left(\mathbf{R}^N\right))} \tag{13}$$

Note that a definition of order parameters distinguishing between reactant and product is not needed. Integration in (12) and (13) extends over the full configuration space.

Next the free energy profile. Unlike the total free energy (12), this quantity is not unique. It does depend on the specification of an order parameter $X\left(\mathbf{R}^N\right)$. This can be a geometric characteristic such as bond length or angle, or, what we are going to use in the end, the vertical energy gap $\Delta E\left(\mathbf{R}^N\right)$. The definition of the diabatic free energy profiles $A_M(x), M = \mathrm{A}, \mathrm{B}$ is similar to (12) but now the integral is restricted to atomic configurations having a given value x of the function $X\left(\mathbf{R}^N\right)$. The restriction is formally imposed by inserting a Dirac delta function.

$$A_M(x) = -k_\mathrm{B}T \ln \Lambda^{-3N} \int \mathrm{d}\mathbf{R}^N \exp[-\beta E_M\left(\mathbf{R}^N\right)]\delta(X\left(\mathbf{R}^N\right) - x) \tag{14}$$

The probability distribution for x can be defined by a similar integral normalizing by the full unrestricted partition function

$$p_M(x) = \frac{\int \mathrm{d}\mathbf{R}^N \exp[-\beta E_M\left(\mathbf{R}^N\right)]\delta(X(\mathbf{R}^N) - x)}{\int \mathrm{d}\mathbf{R}^N \exp[-\beta E_M\left(\mathbf{R}^N\right)]} = \langle \delta\left(X - x\right)\rangle_M \tag{15}$$

In the second identity the probability distribution is formally written as an expectation value of a Dirac delta function (recall the difference in status of X and x in our notation, X is a function of configuration $\mathbf{R}^N$, which has been suppressed here, while x is a constant). Comparing (14) and (15) we see that

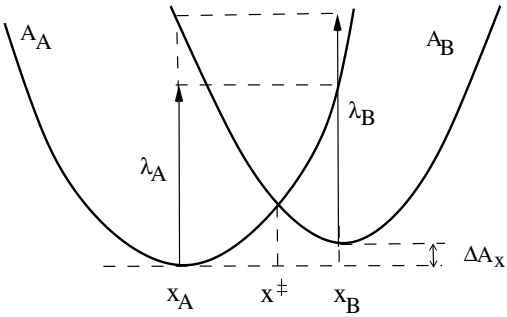


Fig. 2. Two intersecting diabatic free energy surfaces A_A and A_B (*thick solid curves*). *Thin solid arrows* represent the reorganization free energies λ_A and λ_B as defined in (17) and (18). Also indicated is the reaction free energy change ΔA_x of (19). Note the important difference with Fig. 1 which shows potential energy curves E_M and vertical energy gaps ΔE. The aim of the theory in this review (Marcus theory) is to establish a relation between energy gaps and free energies

$$A_M(x) = -k_\mathrm{B}T \ln p_M(x) + A_M \tag{16}$$

where A_M is the diabatic free energy of (12). The free energy profiles (14) are the basis for the definition of reorganization and activation free energies, using (16) for their computation. These free energies are differences relative to the values of the stable minimum of $A_M(x)$ in the reactant and product state. We will indicate the location of these minima by x_A and x_B (see Fig. 2). Reorganization free energies are then defined as

$$\lambda_\mathrm{A} = A_\mathrm{A}\left(x_\mathrm{B}\right) - A_\mathrm{A}\left(x_\mathrm{A}\right) \tag{17}$$

$$\lambda_\mathrm{B} = A_\mathrm{B}\left(x_\mathrm{A}\right) - A_\mathrm{B}\left(x_\mathrm{B}\right) \tag{18}$$

While reorganization free energies for reactant and product state are in general different, $\lambda_\mathrm{A} = \lambda_\mathrm{B}$ in Marcus theory (see Sect. 2.4). We should also be careful to distinguish between the reaction free energy change as determined from the stable values of $A_M(x)$ (see Fig. 2)

$$\Delta A_x = A_\mathrm{B}\left(x_\mathrm{B}\right) - A_\mathrm{A}\left(x_\mathrm{A}\right) \tag{19}$$

and the reaction free energy ΔA of (13). ΔA_x of (19) is, in principle, dependent on the choice of reaction coordinate, which is why we have appended the subscript x, while ΔA of (13) is not. Finally the free energy of activation for the forward ($M = \mathrm{A}$) and reverse ($M = \mathrm{B}$) reaction can be defined as

$$\Delta A^\dagger_M = A_M\left(x^\dagger\right) - A_\mathrm{M}\left(x_\mathrm{M}\right) \tag{20}$$

where $x^\dagger$ is the value of the reaction coordinate where the diabatic curves intersect.

2.3 Free Energy Perturbation Method

The gap is a difference of total energies. The free energy required to change the gap can therefore be obtained using free energy perturbation methods developed in classical force field based simulation to calculate the free energy cost of a modification of the Hamiltonian (see [40, 41] and the text book of Frenkel and Smit [42]). This special feature of the gap has a number of consequences, both from the point of view of computational method [5, 7] and interpretation [10]. We start by noting that the expression (13) can be rewritten as an average over the exponential of the vertical energy gap ΔE of (3). The result has the form of a free energy perturbation (FEP) expression [40–42]

$$\Delta A = -\frac{1}{\beta} \ln\langle e^{-\beta \Delta E} \rangle_{\mathrm{A}} \tag{21}$$

$$= \frac{1}{\beta} \ln\langle e^{\beta \Delta E} \rangle_{\mathrm{B}} \tag{22}$$

where $\langle \cdots \rangle_M$ denotes the canonical average over the PES of state $M = \mathrm{A,B}$. As profound as this expression is, it is also impossible to use in practical calculations, except in very special cases. The reason is that for most systems of interest there is no overlap between regions in configuration space accessible by thermal fluctuations of the equilibrium reactant and product. So, taking (21) as example, the configurations visited by a trajectory in state A will give large positive values for ΔE and therefore a vanishing exponential weight (see Fig. 1). On the other hand ΔE is negative for equilibrium configurations of state B resulting in a huge exponent. Unfortunately the dynamics controlled by E_{A} never reaches this part of configuration space. A sure signature of these sampling problems is a discrepancy between the averages of (21) and (22).

The reader should be familiar with the problem. It is the central problem in free energy computation which every method is trying to solve. The method favored by Warshel in his ET calculations (for a review see [7]) is umbrella sampling (US) using as biasing potential a linear superposition of the two diabatic PES:

$$E_\eta\left(\mathbf{R}^N\right) = \eta E_{\mathrm{B}}\left(\mathbf{R}^N\right) + (1-\eta) E_{\mathrm{A}}\left(\mathbf{R}^N\right) \tag{23}$$

Varying the value of the coefficient η from 0 to 1 creates a series of potentials gradually transforming E_{A} into E_{B}. The discrete increments $\Delta\eta$ must be chosen sufficiently small for the exponential sampling to be accurate so we can use (21) to estimate the corresponding increase in free energy

$$A(\eta + \Delta\eta) - A(\eta) = -k_{\mathrm{B}} T \ln\langle \exp\left[-\beta\left(E_{\eta+\Delta\eta} - E_\eta\right)\right]\rangle_\eta \tag{24}$$

where $A(\eta) = -k_{\mathrm{B}} T \ln \Lambda^{-3N} \int \mathrm{d}\mathbf{R}^N \exp(-\beta E_\eta)$ is the free energy generated by the surface E_η (conf. (12)) and the subscripted brackets denote a thermal average over this surface. The reaction free energy ΔA is obtained by

adding all intermediate free energy changes. Alternatively we can exploit the linear dependence on η which goes back to the coupling parameter method of Kirkwood [43] and use thermodynamic integration

$$\Delta A = A(1) - A(0) = \int_0^1 \mathrm{d}\eta \, \frac{\mathrm{d}A(\eta)}{\mathrm{d}\eta} = \int_0^1 \mathrm{d}\eta \, \langle \Delta E \rangle_\eta \tag{25}$$

Similar to (24) $\langle \Delta E \rangle_\eta$ is computed for a discrete set of values of η (windows) between 0 and 1 and the integral is approximated by a finite sum.

The extension of the FEP/US scheme for the computation of the order parameter probability function $p_M(x)$ (15) is an example of the histogram method. $p_M(x)$ can be obtained in "principle" by determining a histogram of order parameter fluctuations in state M. However, what we are interested in, is the value of $p_M(x)$ at the crossing of the diabatic free energy surfaces $A_\mathrm{A}(x)$ and $A_\mathrm{B}(x)$ because that is where the reaction (electron transfer) takes place. This will give us our estimate of the activation free energy (see Fig. 2 and below). But, of course, we encounter the same problem as for reaction free energy change calculation, this region will be out of reach of equilibrium fluctuations and is sampled very poorly, if at all. This is no different from the adiabatic picture. However, the diabatic picture (an approximation justified in the weak coupling limit) allows for a neat formulation of this problem. Following a derivation similar to the one which led to (21), we can relate the order parameter probability distributions of one diabatic state to the distribution of the other.

$$\frac{p_\mathrm{B}(x)}{p_\mathrm{A}(x)} = e^{\beta \Delta A} \frac{\langle e^{-\beta \Delta E} \delta(X - x) \rangle_\mathrm{A}}{\langle \delta(X - x) \rangle_\mathrm{A}} \tag{26}$$

As in (21) we have to include exponents of the vertical gap as weights in the averaging. This seriously deteriorates the accuracy except when the energy differences are small, such as is the case in the crossing region where the weights are approximately unity ($\Delta E \approx 0$). So the idea is again to bridge distributions p_A and p_B by a series of overlapping distributions (umbrellas) generated by the bias potential (23). The fluctuations of $X\left(\mathbf{R}^N\right)$ for the MD trajectories for a representative set of discrete values of η are binned in histograms. These histograms are merged into a single diabatic probability distribution $p_\mathrm{A}(x)$ which now extends all the way to values of x where $p_\mathrm{B}(x)$ has its maximum (see for example [44] and the Frenkel & Smit textbook [42]).

Equation (26) also shows why the vertical energy gap ΔE is different from other reaction coordinates. Taking logarithms and substituting (16) we find

$$A_\mathrm{B}(x) - A_\mathrm{A}(x) = -k_\mathrm{B} T \ln \left[\frac{\langle e^{-\beta \Delta E} \delta(X - x) \rangle_\mathrm{A}}{\langle \delta(X - x) \rangle_\mathrm{A}} \right] \tag{27}$$

Setting $\Delta E = X$ we see that ΔE appears both in the argument of the delta function on the rhs and the exponent. As a result the exponent can be taken outside the configurational integral giving

$$A_{\mathrm{B}}(\epsilon) - A_{\mathrm{A}}(\epsilon) = \epsilon \tag{28}$$

where ϵ stands for a given value of the gap ΔE. This relation, commonly referred to as the Zwanzig relation, states that the logarithm of the probability distributions of the diabatic states are linearly dependent. The free energy curve of the product can be obtained adding ϵ to the free energy curve for the reactant. This equation, which is rigorous and completely general has a number of interesting and useful implications for computations which use the energy gap as reaction coordinate. For example, substituting in (17) and (18) with $X = \Delta E$ we find for the reorganization energies

$$\lambda_{\mathrm{A}} = +\Delta A_{\epsilon} - \epsilon_{\mathrm{B}} \tag{29}$$

$$\lambda_{\mathrm{B}} = -\Delta A_{\epsilon} + \epsilon_{\mathrm{A}} \tag{30}$$

where ΔA_{ϵ} is the reaction free energy change as determined from the minima in the $A_M(\epsilon)$ curves at ϵ_{A} and ϵ_{B} (see (19)). Equations (29) and (30) tell us that, once we have an estimate of the reaction free energy change, we can find the reorganization free energies simply from the equilibrium values of the diabatic gap, which for all practical purposes can be equated with the average gap in the reactant and product state. This is particularly helpful when the reaction free energy change vanishes ($\Delta A_{\epsilon} = 0$), as for example for self exchange reactions. Note also that, since reorganization free energies are by definition positive, (29) and (30) imply that

$$\epsilon_{\mathrm{B}} < \Delta A_{\epsilon} < \epsilon_{\mathrm{A}} \tag{31}$$

The equilibrium energy gaps set a lower and upper bound to the reaction free energy change. Finally, since in the diabatic approximation the transfer is assumed to take place at the curve crossing ($\epsilon^{\dagger} = 0$), the activation free energy (20) becomes directly equal to the free energy cost of closing the energy gap,

$$\Delta A_M^{\dagger} = A_M\left(\epsilon = 0\right) - A_M\left(\epsilon_M\right) \tag{32}$$

2.4 Relation to Marcus Theory

We are now ready to return to Marcus theory, which was at the origin of all this formalism. Marcus assumed that the solvent responds linearly to a change in charge, which is equivalent to approximating distribution (15) for the energy gap by a Gaussian

$$p_M(\epsilon) = \frac{1}{\sqrt{2\pi}\sigma_M} \exp\left[-(\epsilon - \epsilon_M)^2/(2\sigma_M^2)\right], \tag{33}$$

where $M = \mathrm{A}, \mathrm{B}$. Since we are dealing with Gaussians

$$\epsilon_M = \langle \Delta E \rangle_M \tag{34}$$

Similarly the squared widths σ_M^2 in (33) are equal to the variance of the gap fluctuations

$$\sigma_M^2 = \left\langle \left(\Delta E - \langle \Delta E \rangle_M \right)^2 \right\rangle_M \tag{35}$$

The corresponding free energy profiles (16) are parabolic

$$A_M(\epsilon) = A_M^\epsilon + \frac{k_\mathrm{B}T}{2\sigma_M^2} \left(\epsilon - \epsilon_M \right)^2 \tag{36}$$

As pointed out in Sect. 2.2 the minimum free energy A_M^ϵ and the total free energy of A_M of (12) are not the same. For Gaussian distributions it is easy to find what is missing

$$A_M^\epsilon = A_M + \frac{k_\mathrm{B}T}{2} \ln \left[2\pi\sigma_M^2 \right] \tag{37}$$

Now comes the step where we use (28) which is special to energy gaps. According to this equation, free energy functions for vertical gaps can differ only by a linear term. This means that quadratic terms must cancel which can only be satisfied if the variance of the fluctuations is independent of the chemical state of the system

$$\sigma_\mathrm{A}^2 = \sigma_\mathrm{B}^2 \equiv 2k_\mathrm{B}T\lambda' \tag{38}$$

The second identity defines the linear response reorganization free energy λ'. Equation (38) may at first seem a rather strong claim, in particular for half reactions, but is, according to (28), a rigorous consequence of the Gaussian approximation. Any violation of this relation must be due to non-linearities as was pointed out by Tachiya in [45]. Similarly, λ' in (38) is in general not identical to the reorganization free energies of (29) and (30) (which is why we have added the prime). This is another point emphasized by Tachiya [39]. For Gaussian distributions, however, these quantities are indeed the same. This follows from the constraints imposed by (28) on the linear coefficients in (36) which require

$$\lambda' = \frac{1}{2} \left(\epsilon_\mathrm{A} - \epsilon_\mathrm{B} \right) \tag{39}$$

Substituting with (36) in expression (29) for λ_A we obtain

$$\lambda_\mathrm{A} = \frac{1}{4\lambda'} \left(\epsilon_\mathrm{A} - \epsilon_\mathrm{B} \right)^2 = \lambda' \equiv \lambda \tag{40}$$

The same result is found for λ_B of (30). For parabolic free energy curves of energy gaps we can therefore forget about the various kinds of reorganization energy, there is only one, which we call λ. Similarly, because of (38) and (37) we can conclude that in the Gaussian approximation the discrepancy between the reaction free energies of ΔA (13) and ΔA_ϵ of (19) vanishes. Rearrangement of (29) and (30) then yields two very useful relations between equilibrium Gaussian energy gaps and the reaction and reorganization free energies

$$\Delta A = \frac{1}{2}\left(\epsilon_{\mathrm{A}} + \epsilon_{\mathrm{B}}\right) \tag{41}$$

$$\lambda = \frac{1}{2}\left(\epsilon_{\mathrm{A}} - \epsilon_{\mathrm{B}}\right) \tag{42}$$

Note that (41) can be directly obtained from the two point approximation to the coupling parameter integral (25). Yet another route to (41) and (42) departs from the cumulant expansion of (21) and (22) (see e.g. [31]).

$$\Delta A = \langle \Delta E \rangle_{\mathrm{A}} - \frac{1}{2k_{\mathrm{B}}T}\sigma_{\mathrm{A}}^2 = \langle \Delta E \rangle_{\mathrm{B}} + \frac{1}{2k_{\mathrm{B}}T}\sigma_{\mathrm{B}}^2 \,, \tag{43}$$

For Gaussian statistics truncation after the second term is exact. Inserting (34) and (38), we can recover (41) by adding and (42) by subtracting. Equations (43) and (38) are the ultimate of economy in free energy computation. While (41) and (42) require two equilibrium runs (one for reactant and one for the product), (43) and (38) claim that both ΔA and λ can be obtained from the mean and variance of the gap fluctuations of a single trajectory, either reactant or product state. The trajectory must of course be of sufficient length to converge the variance σ_M^2. This quantity has been eliminated in (41) and (42) which is a significant advantage in *ab initio* MD calculations, where runs are short. Finally, substitution in (32) yields the famous Marcus gap law for the activation energy (written here for the forward reaction)

$$\Delta A_{\mathrm{A}}^{\dagger} = \frac{(\lambda + \Delta A)^2}{4\lambda} \tag{44}$$

3 Redox Half Reactions

3.1 Ab Initio MD Considerations

Equations (29), (30) and (32), or their Gaussian (Marcus) approximations of Sect. 2.4, have been the basis for much of the simulation work on electron transfer, using either semiclassical methods or fully classical point charge models. However, the use of the vertical diabatic energy gap as a reaction coordinate is not restricted to ET in the weak coupling limit even though for systems with strong off-diagonal interactions (1) the diabatic energy surfaces are not acceptable approximations to the adiabatic PES. The discrepancies, particularly in the crossing region, are too large. However, while ΔE can no longer be interpreted as the vertical (optical) ET excitation energy (4), it retains its usefulness as an order parameter for ET. Moreover, the diabatic PES can be used as a reference potential. The deviation from the true adiabatic ground state energy surface can be accounted for by means of FEP techniques similar to the methods discussed in Sect. 2.3. The adiabatic ground state energy can be obtained from diagonalization of the diabatic Hamiltonian matrix ((1) in case of a two level model) or from a completely independent calculation

using ab initio electronic structure methods such as MP2 or DFT [8,9]. The question then is how to define the diabatic PES. Warshel and coworkers opt for an empirical valence bond (EVB) scheme. In a series of papers they have applied this approach, not only to ET [5–7] but also to proton transfer and enzymatic reactions [8–11].

How to implement the diabatic energy gap approach in a DFT based ab initio MD ("Car-Parrinello") simulation of a condensed molecular system, when there is no EVB model available? In the case of ET it should be in principle feasible to reconstruct the diabatic surfaces from the adiabatic electron transfer excitation energies. Unfortunately DFT, at the level it is usually applied in ab initio MD simulation, is notoriously unreliable for treating charge transfers. The way we have avoided this problem in our work on redox reactions is to separate full redox reactions (2) in half reactions.

$$\mathrm{R} \rightarrow \mathrm{O} + \mathrm{e}^- \tag{45}$$

Model systems now consist only of a single redox active solute rather than two. Instead the number of electrons in the system can vary between n, say, for the reduced state R and $n-1$ for the oxidized state O. Implementation in a Car-Parrinello simulation is fairly straightforward. The method normally produces a finite temperature trajectory on the adiabatic ground state PES of R or O. What our scheme in essence does is adding a second calculation recomputing for the same ionic configuration the ground state energy of the system with one electron less, (O) or one electron more (R), which gives the vertical energy gap at that configuration.

The non-adiabatic picture as expressed in (5) is thus taken to the extreme in our half reaction scheme. The coupling parameter γ is not only small, it is zero, as it would be the case for very large separations between donor and acceptor. This is in fact the proper limit for the reaction free energy ΔA if we want to compare the result of our calculations to experimental *standard* redox potentials which formally correspond to reaction free energies at infinite dilution. This is the main objective of our approach. However, also the activation and reorganization free energies remain meaningful, if not quantitatively, then at least qualitatively as a way to understand the redox reaction kinetics.

There is, however, a serious objection that comes immediately to mind. The interpretation of the vertical gap ΔE or free energy ΔA as the corresponding quantities in a homogeneous solution clearly cannot be correct in model systems under periodic boundary conditions as used in our simulation. Either the R or O state, or both, end up with a net charge. In the Ewald summation methods used in our codes, net charge of the model is automatically compensated by a homogeneous back ground charge, effectively playing the role of a counter ion. Energies of half reactions have therefore no experimental meaning. Still, free energies of full reactions, obtained as differences of half reactions, can be compared to experiment, but only if charges of reactant and product species are the same. Under these (rather restrictive) conditions long

range interactions cancel (see Sect. 4.1). System size effects are not the subject of this contribution which focuses on the statistical mechanics. The practical justification we can offer here is the success of our scheme in reproducing a (limited) number of experimental redox potentials for full reactions of model aqueous transition metal coordination complexes and organic molecules (see further Sect. 4.1).

3.2 Parallel to Electrode Reactions

There is a second way to view the half reaction scheme. In electrochemistry a half reaction (45) is regarded as a zero order approximation to heterogeneous electron transfer between an ion in solution and a metal electrode E (see for example [46]). In the notation used in previous sections where reactant and product state were indicated by A respectively B, we now have A = R + E representing the state with the electron held by the reduced ion and B = O + E^- the state with the active electron transferred to the electrode E. This is a very useful parallel. It suggests that the thermodynamic driving force in the simulation can be controlled similar to the way electrochemists control a redox reaction by applying a voltage to the electrodes of the cell. In the minimal implementation of this scheme we have employed in our simulations, the electrode is replaced by a fictitious electron reservoir at electronic chemical potential μ, which exchanges electrons with the solution but has no further interactions with the solution. The PESs A and B remain the same in this approximation except that we must add a shift μ to the PES of B

$$E_{\mathrm{A}}\left(\mathbf{R}^N\right) = E_{\mathrm{R}}\left(\mathbf{R}^N\right) \tag{46}$$

$$E_{\mathrm{B}}\left(\mathbf{R}^N\right) = E_{\mathrm{O}}\left(\mathbf{R}^N\right) + \mu \tag{47}$$

The vertical energy gap (3) for the transfer to the fictitious electrode is offset by the same bias μ. We formally write for this electrochemically controlled gap

$$\Delta E_{\mu}\left(\mathbf{R}^N\right) = E_{\mathrm{B}}\left(\mathbf{R}^N\right) - E_{\mathrm{A}}\left(\mathbf{R}^N\right) = \Delta E\left(\mathbf{R}^N\right) + \mu \tag{48}$$

where ΔE is the ET energy of (3). The same linear relation holds for the free energy of oxidation

$$\Delta A_{\mu} = A_{\mathrm{B}} - A_{\mathrm{A}} = \Delta A + \mu \tag{49}$$

where ΔA is the free energy difference (13). This will be the interpretation of (45) we adopt in the remainder of the discussion. All equations of Sects. 2.2, 2.3 and 2.4 can be carried over if we make the following simple replacements

$$\begin{aligned} \mathrm{A} &\rightarrow \mathrm{R} & \mathrm{B} &\rightarrow \mathrm{O} \\ \Delta E &\rightarrow \Delta E_{\mu} & \Delta A &\rightarrow \Delta A_{\mu} \end{aligned} \tag{50}$$

In practice the electrochemical potential μ plays the role of an external parameter used to align the free energy profiles of reactant and product state.

Indeed by choosing $\mu = -\Delta A$ the effective reactive energy change ΔA_μ (49) can be made to vanish. ΔA can be estimated using the FEP methods outlined in this chapter or alternatively, we can use the "numerical titration scheme" proposed in [34] and applied in [28] and [31]. The freedom of the control over μ is directly exploited in this scheme by stepwise variation of μ until the oxidation reaction is observed (for a critical evaluation of this approach and a comparison to the FEP method see [31]).

4 Applications

4.1 Overview

The methodology outlined in the previous sections has been applied to a number of redox active solutes. The redox reactions we have investigated and the result for the reaction free energy change are listed in Table 1 together with the experimental values of the reaction free enthalpy. The density functional in all calculations was BLYP. The simulation for the aqueous transition metal coordination complexes (reactions **1** and **2**) were carried using the CPMD package [47]. Standard norm conserving pseudo potentials (Trouillier-Martins [48]) were used and conventional plane wave cutoffs ($70 - 80$Ry). Technical details can be found in the original publications given in Table 1. For the molecular systems (reactions **3** and **4)** we employed a new mixed Gaussian-plane wave *ab initio* MD method [49] and the Quickstep code [50]. The one electron orbitals are expanded in Gaussian basis sets. Core electrons are represented using separable norm conserving pseudo potentials according to the Goedecker-Hutter recipe [51]. The plane waves are used as an auxiliary basis set to describe the density [49] which enables us to compute the long range electrostatic energies using fast Fourier transform methods similar to CPMD.

Table 1. Reaction free energies in units of eV of four model redox reactions compared to experiment. The first two reactions involve transition metal aqua ions. TH (thianthrene) and TTF (tetrathiafulvalene) are two organosulfur compounds which can be oxidized to stable radical cations. BQ (benzoquinone) and DQ (duroquinone) are small quinones forming radical anions. The last column gives the reference to the original papers

	Redox Reaction	Solvent	ΔA(calc.)	ΔG(exp.)	Ref.
(**1**)	$Cu^{1+} + Ag^{2+} \rightarrow Cu^{2+} + Ag^{1+}$	water	-1.7	-1.83	28
(**2**)	$RuO_4^{2-} + MnO_4^{1-} \rightarrow RuO_4^{1-} + MnO_4^{2-}$	water	-0.3	$+0.03$	31
(**3**)	$TH^{\bullet+} + TTF \rightarrow TH + TTF^{\bullet+}$	acetonitrile	-0.9	-0.93	35
(**4**)	$DQ^{\bullet-} + BQ \rightarrow DQ + BQ^{\bullet-}$	methanol	-0.43	-0.46	36

The redox free energies ΔA of reactions **2, 3** and **4** were estimated from the average vertical energy gaps using the Marcus approximation ((41) for reactions **2** and **3** and (43) for reaction **4**). ΔA of reaction **1** was obtained using the scheme of [34]. This method for the determination of redox potentials does not have to rely on the Gaussian approximation and can be applied to systems outside the Marcus regime. The Ag ion [33] and most likely also the Cu aqua ion are examples of such non-Marcus ions (see further Sect. 4.2).

The agreement with experiment is good. Discrepancies are in the 100 meV range except for reaction **2**, where the error is 300 meV. The simulation parameters most critical to the assessment of the accuracy of these results are the duration of the MD runs and the model system size. The typical run length is 10 ps (some runs are shorter, 5 ps, others longer, 20 ps). The number of solvent molecules is between 30 and 50, with cubic box dimensions in the order of 10-15 Å. All model systems contain only a single redox active ion without counterion. In systems of such small dimensions the interaction with periodic images and the Ewald back ground charge distribution is very large (in the eV's). There are a number of considerations why these large size effects are apparently not affecting the accuracy of the redox free energies in Table 1 to the same extent. Most important is a compensation of errors. The reactions in Table 1 are all of the type

$$\mathrm{X}^m + \mathrm{Y}^{m+1} \rightarrow \mathrm{X}^{m+1} + \mathrm{Y}^m \tag{51}$$

The species in reactant and product have the same charges and approximately the same spatial dimension. From a distance they will look rather similar to the solvent and the long range errors cancel.

4.2 Two Examples: The Ru and Ag Aqua Cations

As an illustration of the methods presented in this chapter, and to underline the importance of vertical energy gaps, we will discuss two half reactions in more detail. Both involve transition metal aqua cations. The first is the $\mathrm{Ag}^{1+} \rightarrow \mathrm{Ag}^{2+} + e^-$ oxidation, i.e. half of the reaction **1** of Table 1. The second is $\mathrm{Ru}^{2+} \rightarrow \mathrm{Ru}^{3+} + e^-$. The reason that this reaction is not included in Table 1 is that we have not yet studied another metal ion with reduced and oxidized states of the same charge so as to satisfy (51). The subject of the discussion in this section is however not the redox free energy itself but the statistical mechanics of the vertical energy gap. We have used therefore the freedom of the control of the redox free energy (49), by choosing a value of μ such that $\Delta A_\mu = 0$. This effectively aligns the free energy minima of the diabatic surfaces in Fig. 2. These values of μ, which are different for each reaction, are given in the caption of Fig. 3 (recall again that similar to ΔA of half reactions, values of μ have no direct experimental meaning).

The probability distributions (15) of the (shifted) vertical energy gap of (48) are shown in Fig. 3. These distributions have been determined by sampling the time evolution of the vertical gap ΔE_μ in equilibrium runs of reduced

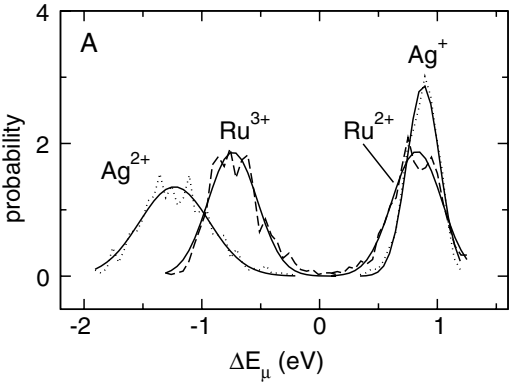


Fig. 3. Vertical energy gap probability distributions p_M (15) as determined from equilibrium trajectories for two oxidation states of the Ru and Ag aqua-cations. *Dashed* and *dotted lines* are the original histograms, *solid curves* are Gaussian fits. The oxidation free energy ΔA has been set to zero by choosing the appropriate value for the chemical potential, $\mu = 0.58$ eV for Ru^{2+}/Ru^{3+} and $\mu = -1.16$ eV for Ag^{1+}/Ag^{2+}. Note the asymmetry for the Ag system (see also Table 2)

and oxidized state. We immediately notice a difference. The distributions for Ru are to a very good approximation Gaussian. The distribution for reduced state (p_R) and oxidized state (p_O) are placed in symmetrical position relative to $\Delta E_\mu = 0$ and have the same variance (see Table 2). This is in agreement with (38) and (41), which requires that $\epsilon_O = -\epsilon_R$ for $\Delta A = 0$ (converting the notation according to (50)). As explained in Sect. 2.4 this symmetry is a necessary condition a Gaussian system must fulfill. Indeed the ET chemistry of

Table 2. Properties characterizing the diabatic free energy profiles A_M shown in Fig. 4 for the Ru^{2+}/Ru^{3+} couple and Fig. 5 for Ag^{1+}/Ag^{2+}. The parameter ϵ_M gives the location of the position of the minimum of the parabolic fit to A_M and λ_M the reorganization free energy computed from these fits according to (17) and (18). $\Delta A^\dagger$ is the activation free energy determined from the intersection point of the parabolic curves. The order parameter is the quantum energy gap of (48) except for the data marked as Ru^{2+}/Ru^{3+}(cl.) which refer to the classical point charge gap of (53). The widths σ_M are the root mean square second moments (35) of the corresponding gap fluctuations used in the criterion of (38) to decide whether Marcus theory applies. All energies are in eV. Data are taken from [32] and [33]

	ϵ_R	ϵ_O	σ_R	σ_O	λ_R	λ_O	$\Delta A^\dagger$	Fig.
Ru^{2+}/Ru^{3+}	0.78	-0.78	0.23	0.23	0.78	0.78	0.20	4A
Ru^{2+}/Ru^{3+}(cl.)	1.00	-0.82	0.39	0.33	0.24	0.42	0.08	4B
Ag^{1+}/Ag^{2+}	1.17	-1.20	0.14	0.30	1.19	1.15	0.3	5

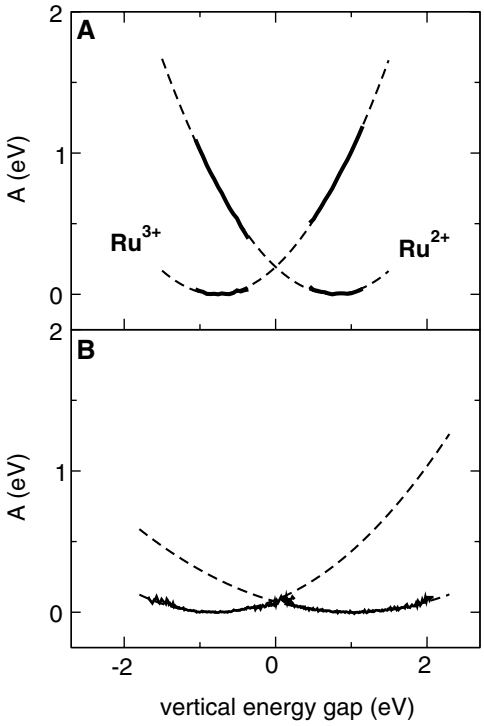


Fig. 4. A: Free energy profiles for Ru^{2+} and Ru^{3+} corresponding to the probability distributions of Fig. 3 using the shifted vertical energy gap of (48) as order parameter. The low energy set of points have been obtained according to (16). The upper parts have been obtained from the lower parts by adding or subtracting the energy gap using (28). Dashed curves are parabolic fits. Panel **B** shows the free energy profiles obtained using the fluctuations of the classical point charge gap of (53) (data taken from [32])

Ru cations is a text book example of an outer sphere process to which Marcus theory applies. Ru(II) forms an octahedral coordination complex with water, $Ru(II)(H_2O)_6$, which is, apart from a small contraction of the metal oxygen distance, virtually preserved in oxidation state III (radial distribution functions and other geometry data can be found in [32]). In contrast Ag^{1+}/Ag^{2+} shows a significant asymmetry. This is consistent with the substantial changes in the hydration structure induced by oxidation. As we found in [28] the coordination number of Ag^{1+} is on average 4 (fluctuations are however large) while the coordination number of Ag^{2+} is 5 (see also [29] and [33]). This coordination change increases the importance of non-linear effects which would violate the assumptions for the validity of the Gaussian approximation (Sect. 2.4).

The free energy profiles corresponding to the distributions of Fig. 3 are shown in Fig. 4 (upper panel) for Ru^{2+}/Ru^{3+} and for Ag^{1+}/Ag^{2+} in Fig. 5.

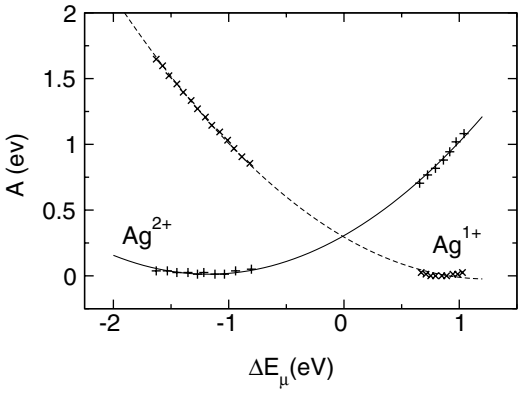


Fig. 5. Free energy profiles for Ag^{1+} ($\times$ symbols) and Ag^{2+} (+ symbols) corresponding to the probability distributions of Fig. 3. The low energy set of points have been obtained according to (16). The upper parts have been obtained from the lower parts using (28). The solid and dashed curves are parabolic fits (data taken from [33])

The curve for each oxidation state consists of two sets of data points connected by a parabolic fit (thin lines). The data points have been obtained as follows. The lower energy part of $A_{\mathrm{R}}(\epsilon)$ has been computed from $p_{\mathrm{R}}(\epsilon)$ by taking the logarithm according to (16) and setting the constant A_{R} to zero. The low energy part of $A_{\mathrm{O}}(\epsilon)$ has been computed applying the same procedure to $p_{\mathrm{O}}(\epsilon)$. In contrast the high energy part of $A_{\mathrm{R}}(\epsilon)$ (at negative gap values) has been obtained from the equilibrium trajectory of the O state. Using (28) we have simply subtracted the value of the vertical gap from $A_{\mathrm{O}}(\epsilon)$. The part of $A_{\mathrm{O}}(\epsilon)$ at positive gap values was generated from $A_{\mathrm{R}}(\epsilon)$ in a similar way.

As expected the free energy profiles of Ru^{2+}/Ru^{3+} form a mirror pair of intersecting diabatic curves. The positions of the minima as obtained from the parabolic fits are listed in Table 2 (first row) which also gives the estimates of the reorganization free energies λ_M computed from the parabolic fits according to (17) and (18) as well as the activation free energy using (32). These results are in excellent agreement with the predictions of the Marcus model ((41), (42) and (44)). However, even for Ag^{1+}/Ag^{2+}, despite of the asymmetries (Fig. 3) parabolic fit functions approximate the simulation data reasonably well. Deviations are largest in the equilibrium region of Ag^{+}.

We emphasize again that (28) used for the construction of the nonequilibrium parts of the curves is valid for every system, Gaussian or not. For Ag we have verified this relation, the most fundamental of gap laws, by comparing to the full diabatic free energy profiles computed using the FEP umbrella sampling method outlined in Sect. 2.3. Three intermediate windows were generated by selecting three values for the coupling parameter η in the bias potential of (23). The values we used were $\eta = 0.25$, 0.5, and 0.75 bridging the gap between the terminal $\eta = 0$ and $\eta = 1$ windows given by the equilibrium

results of Fig. 3. The results are reported in [33]. The FEP/US $A_M(\epsilon)$ profiles closely follow the parabolic fits in Fig. 5. The largest deviations are observed for gap values in the interval $\Delta E_\mu = 0.5 - 1$ eV. This is the region where most of the coordination number fluctuations are found to occur, confirming that they might be responsible for the non-linear behaviour [33].

The results for Ru^{2+}/Ru^{3+} and Ag^{1+}/Ag^{2+} confirm that the diabatic energy gap can be used as reaction coordinate for redox reactions, both in the Marcus regime and outside. Are there alternatives? The question is of interest because total energy gaps come with a number of technical complications for *ab initio* MD applications. One of the difficulties discussed in Sect. 3 is the computation of excited electronic states. Another drawback is that energy gaps prohibit the use of constraint methods. In this still rather popular method the free energy profiles are obtained as integrals over the mean force acting on the reaction coordinate (hence the name potential of mean force) [52]. The reaction coordinate is fixed at a series of values and the mean force is estimated from the force of constraint. So in the notation of Sect. 2.2

$$A(x) - A(x_0) = \int_{x_0}^{x} dx' \left\langle \frac{\partial H}{\partial X} \right\rangle_{X=x'} = -\int_{x_0}^{x} dx' \lambda_X(x') \tag{52}$$

where $\lambda_X(x')$ is the time average of the Lagrange parameter keeping the value of $X\left(\mathbf{R}^N\right)$ fixed at x' during the MD run (the last identity is approximate, for the corrections see [52]). This approach in practice requires that the reaction coordinate is available as an explicit function of configuration. Total energy is an implicit function so cannot be subjected to mechanical constraints.

Constraint methods, if applicable, are easy to use in *ab initio* MD which motivated us to search for configurational reaction coordinates to describe the Ru^{2+}/Ru^{3+} and Ag^{1+}/Ag^{2+} half reactions, preferably a structural (geometric) parameter. The coordination change we observed for Ag^{1+}/Ag^{2+} suggests that coordination number (n_c) might be used for this purpose in this system, which is what we tried in the work reported in [29]. The PMF we obtained is reproduced in Fig. 6. The contrast with Fig. 5 is striking. While there is a maximum in the PMF around $n_c = 4.6$, its value is one order of magnitude smaller than the activation free energy at the curve crossing in Fig. 5 (see also Table 2). The apparent barrier in Fig. 6 of 17 mev (=195 K) is well in the thermal range, implying that spontaneous oxidation might occur on the MD time scale. The reason that these events are not observed is that n_c is evidently inadequate as a reaction coordinate (in fact as the analysis in [33] shows n_c is not even a good order parameter for distinguishing unambiguously between oxidation states). The transition dynamics is dominated by solvent rearrangements which are not represented by the PMF for n_c. The consequence is underestimation of the free energy of activation. This phenomenon is now well understood thanks of the work of the Chandler group (see for example [53, 54] and the chapter on transition path sampling in this book). The special point for us here is that the solvent reorganization is much better accounted for by the free energy of the diabatic energy gap. This feature

of diabatic energy gaps has been repeatedly emphasized by Warshel (see in particular [8, 10, 11]).

The discussion above raises the question whether perhaps the classical electrostatic point charge energy gap (11), while a (probably poor) approximation to the full electronic energy gap, could non the less play a similar role as reaction coordinate capturing the solvent reorganization. The idea would be to simply assign SPC charges to the H and O atoms of the solvent and compute the electrostatic potential at the site of the ion by standard Ewald methods. What is gained is that the mean force for this quantity can be determined using mechanical constraint methods or included among the metadynamics variables. The vertical SPC gap ΔU for half reactions is given directly by the electrostatic potential (10). The derivation is similar to the one given for the ET gap in Sect. 2. As we already have established that the $Ru^{+2} \rightarrow Ru^{3+} + e^-$ reaction adheres closely to the Marcus rules, this system seems a good candidate to subject this idea to a test. For the purpose of comparison to the fully consistent DFT results of Fig. 4 it is convenient to "calibrate" ΔU by the ionization energy in vacuum, denoted by ΔE_0^{v}, leading to the following expression for the classical vertical gap

$$\Delta U_\mu(\mathbf{R}^N) = \Phi(\mathbf{R}^N) + \frac{\xi_{\mathrm{EW}}}{2L}(q_{\mathrm{O}}^2 - q_{\mathrm{R}}^2) + \Delta E_0^{\mathrm{v}} + \mu \tag{53}$$

We have also included the self interaction energy of the ion with its periodic images and background charge in a finite periodic cell [55]. The Ewald constant $\xi_{\mathrm{EW}} = -2.837297$ and L is the box length. q_M is the classical point charge of the ion in state M. Finally, to be consistent with the bias applied to the quantum gap ΔE_μ (see (48)) we have also offset ΔU by an adjusted electronic chemical potential μ ensuring that reduced and oxidized state are strictly thermodynamically equivalent ($\Delta A_\mu = 0$).

The PMF for the SPC gap of Ru^{2+}/Ru^{3+} is compared to the full quantum gap free energy profile in Fig. 4. Only the lower energy (adiabatic) part has

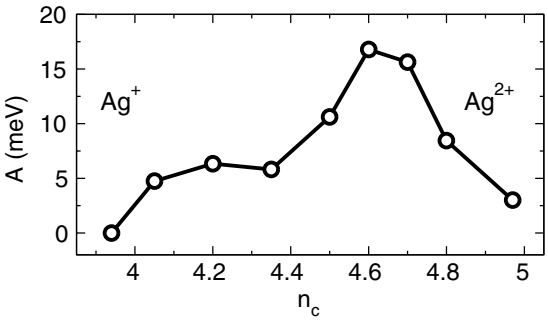


Fig. 6. Potential of mean force for the oxygen coordination number n_c of the Ag aqua cation. $n_c = 4$ for Ag^{1+} and $n_c = 5$ for Ag^{2+}. Oxidation takes place at $n_c = 4.6$ (data taken from [29]). Note the difference in energy scale with Fig. 5

been computed. It is clear, however, that the free energy profiles along the SPC gap are again harmonic but the curvature of the Ru^{3+} curve is almost twice larger than for Ru^{2+}. The same factor two difference is found for the variance of the equilibrium fluctuations (Table 2). In contrast to the quantum gap, the Marcus condition (38) is not satisfied for the classical gap. There is no conflict because (38) is only valid for the true energy gap of a Gaussian system. However, consistent with the larger width, also the activation free energy of $\Delta A_c^\dagger = 0.08$ eV predicted by the maximum in the PMF of the classical gap is lower than the estimate $\Delta A_q^\dagger = 0.20$ eV obtained for the quantum gap. Because the Ru cation is such a well behaved Marcus ion and ΔE_μ rather than ΔU_μ satisfies all the Marcus rules, we can be confident that $\Delta A_q^\dagger$ is the more reliable number. The classical electrostatic potential generated by fixed charges on the solvent atoms can be assumed to account for most of the reorganization of the orientational (inertial) solvent polarization. This can be concluded from all the work on classical models of aqueous Fe^{2+}/Fe^{3+} charge transfer, which showed that when ΔU is the consistent vertical gap the PMF is again symmetric. [7,12,14,16,21,22,24,25]. When the gap is only an approximation to the true total energy gap we can expect deviations from linear response. What is surprising, is that this effect for the Ru^{2+}/Ru^{3+} is as large as we found it to be.

4.3 Conclusion

The examples discussed in this section were meant to illustrate the use of the vertical diabatic energy gap as reaction coordinate for the study of redox reactions. First of all the energy gap proves to be an appropriate microscopic degree of freedom to represent the solvent polarization in Marcus theory. Because of its special status in free energy perturbation methods it also leads to a set of convenient and efficient expressions for the computation of the reaction free energy change of redox reactions. The most intriguing aspect of the energy gap is perhaps its potential for the computation of activation free energies. As the connection to Marcus theory already suggests, it seems to be a most suitable reaction coordinate to describe the for redox reactions all important solvent reorganization. Because of limitations in current DFT implementations, we focused on half reactions. The challenge is now to extend these methods to full electron transfer and redox reactions.

Acknowledgments

The following people have been collaborating with us in this ongoing project on redox reactions: Ivano Tavernelli, Yoshitaka Tateyama, Joost VandeVondele, Marialore Sulpizi, Regla Ayala Espinar and Rodolphe Vuilleumier. We thank them for their many and crucial contributions.

References

1. R. A. Marcus (1956) Theory of oxidation-reduction reactions involving electron transfer.1. *J. Chem. Phys.* **24**, p. 966
2. R. A. Marcus (1960) Theory of oxidation-reduction reactions involving electron transfer.4. A statistical-mechanical basis for treating contributions from solvent, ligands, and inert salt. *Discuss. Faraday Soc.* **29**, p. 21
3. R. A. Marcus (1965) On theory of electron-transfer reactions.6. Unified treatment for homogeneous and electrode reactions. *J. Chem. Phys.* **43**, p. 679
4. R. A. Marcus (1993) Electron-transfer reactions in chemistry – theory and experiment. *Rev. Mod. Phys.* **65**, p. 599
5. A. Warshel (1982) Dynamics of reactions in polar-solvents – semi-classical trajectory studies of electron-transfer and proton-transfer reactions. *J. Phys. Chem.* **86**, p. 2218
6. J. K. Hwang and A. Warshel (1987) Microscopic examination of free-energy relationships for electron-transfer in polar-solvents. *J. Am. Chem. Soc.* **109**, p. 715
7. G. King and A. Warshel (1990) Investigation of the free-energy functions for electron-transfer reactions. *J. Chem. Phys.* **93**, p. 8682
8. R. P. Muller and A. Warshel (1995) ab-initio calculations of free-energy barriers for chemical-reactions in solution. *J. Phys. Chem.* **99**, p. 17516
9. J. Bentzien, R. P. Muller, J. Florian, and A. Warshel (1998) Hybrid ab initio quantum mechanics molecular mechanics calculations of free energy surfaces for enzymatic reactions: The nucleophilic attack in subtilisin. *J. Phys. Chem. B* **102**, p. 2293
10. J. Villa and A. Warshel (2001) Energetics and dynamics of enzymatic reactions. *J. Phys. Chem. B* **105**, p. 7887
11. M. Strajbl, G. Hong, and A. Warshel (2002) Ab initio QM/MM simulation with proper sampling: "First principle" calculations of the free energy of the autodissociation of water in aqueous solution. *J. Phys. Chem. B* **106**, p. 13333
12. R. A. Kuharski, J. S. Bader, D. Chandler, M. Sprik, M. L. Klein, and R. W. Impey (1988) Molecular-model for aqueous ferrous ferric electron-transfer. *J. Chem. Phys.* **89**, p. 3248
13. E. A. Carter and J. T. Hynes (1989) Solute-dependent solvent force-constants for ion-pairs and neutral pairs in a polar-solvent. *J. Phys. Chem.* **93**, p. 2184
14. R. B. Yelle and Y. Ichiye (1997) Solvation free energy reaction curves for electron transfer in aqueous solution: Theory and simulation. *J. Phys. Chem. B* **101**, p. 4127
15. K. Ando (1997) Quantum energy gap law of outer-sphere electron transfer reactions: A molecular dynamics study on aqueous solution. *J. Chem. Phys.* **106**, p. 116
16. K. Ando (2001) Solvent nuclear quantum effects in electron transfer reactions. II. Molecular dynamics study on methanol solution. *J. Chem. Phys.* **114**, p. 9040
17. K. Ando (2001) Solvent nuclear quantum effects in electron transfer reactions. III. Metal ions in water. Solute size and ligand effects. *J. Chem. Phys.* **114**, p. 9470
18. K. Ando (2001) A stable fluctuating-charge polarizable model for molecular dynamics simulations: Application to aqueous electron transfers. *J. Chem. Phys.* **115**, p. 5228

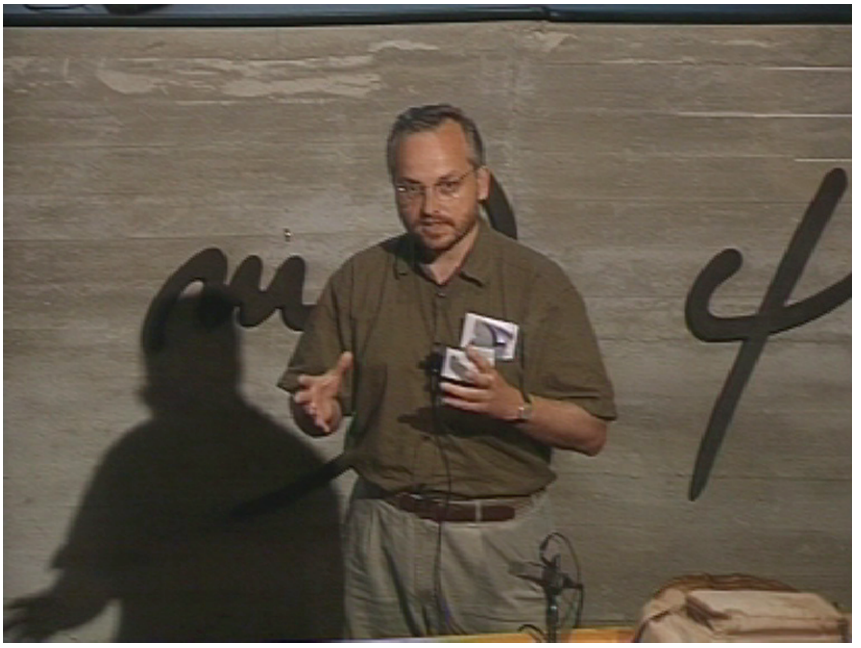
19. D. W. Small, D. V. Matyushov, and G. A. Voth (2003) The theory of electron transfer reactions: What may be missing? *J. Am. Chem. Soc.* **125**, p. 7470
20. T. Ishida (2005) Polarizable solute in polarizable and flexible solvents: Simulation study of electron transfer reaction systems. *J. Phys. Chem. B* **109**, p. 18558
21. D. A. Rose and I. Benjamin (1994) Molecular-dynamics of adiabatic and nonadiabatic electron-transfer at the metal-water interface. *J. Chem. Phys.* **100**, p. 3545
22. D. A. Rose and I. Benjamin (1995) Solvent-free energies for electron-transfer at a solution metal interface – effect of ion charge and external electric-field. *Chem. Phys. Lett.* **234**, p. 209
23. J. B. Straus and G. A. Voth (1993) A computer-simulation study of free-energy curves in heterogeneous electron-transfer. *J. Phys. Chem.* **97**, p. 7388
24. J. B. Straus, A. Calhoun, and G. A. Voth (1995) Calculation of solvent-free energies for heterogeneous electron-transfer at the water-metal interface – classical versus quantum behavior. *J. Chem. Phys.* **102**, p. 529
25. A. Calhoun and G. A. Voth (1998) Isotope effects in electron transfer across the electrode-electrolyte interface: A measure of solvent mode quantization. *J. Phys. Chem. B* **102**, p. 8563
26. C. Hartnig and M. T. M. Koper (2001) Molecular dynamics simulations of solvent reorganization in electron-transfer reactions. *J. Chem. Phys.* **115**, p. 8540
27. C. Hartnig and T. M. Koper (2004) Molecular dynamics simulation of solvent reorganization in ion transfer reactions near a smooth and corrugated surface. *J. Phys. Chem. B* **108**, p. 3824
28. J. Blumberger, L. Bernasconi, I. Tavernelli, R. Vuilleumier, and M. Sprik (2004) Electronic structure and solvation of copper and silver ions: A theoretical picture of a model aqueous redox reaction. *J. Am. Chem. Soc.* **126**, p. 3928
29. J. Blumberger and M. Sprik (2004) Free energy of oxidation of metal aqua ions by an enforced change of coordination. *J. Phys. Chem. B* **108**(21), p. 6529
30. J. Blumberger and M. Sprik (2005) Ab initio molecular dynamics simulation of the aqueous Ru_{2+}/Ru_{3+} redox reaction: The Marcus perspective. *J. Phys. Chem. B* **109**, p. 6793
31. Y. Tateyama, J. Blumberger, M. Sprik, and I. Tavernelli (2005) Density-functional molecular-dynamics study of the redox reactions of two anionic, aqueous transition-metal complexes. *J. Chem. Phys.* **122**, p. 234505
32. J. Blumberger and M. Sprik (2006) Quantum versus classical electron transfer energy as reaction coordinate for the aqueous Ru_{2+}/Ru_{3+} redox reaction. *Theor. Chem. Acc.* **115**, p. 113
33. J. Blumberger, I. Tavernelli, M. L. Klein, and M. Sprik (2006) Diabatic free energy curves and coordination fluctuations for the aqueous Ag^{+}/Ag^{2+} redox couple: A biased Born-Oppenheimer molecular dynamics investigation. *J. Chem. Phys.* **124**, p. 064507
34. I. Tavernelli, R. Vuilleumier, and M. Sprik (2002) Ab initio molecular dynamics for molecules with variable numbers of electrons. *Phys. Rev. Lett.* **88**, p. 213002
35. J. VandeVondele, R. Lynden-Bell, E. J. Meijer, and M. Sprik (2006) Density functional theory study of tetrathiafulvalene and thianthrene in acetonitrile: Structure, dynamics, and redox properties. *J. Phys. Chem. B* **110**, p. 3614
36. J. VandeVondele, M. Sulpizi, and M. Sprik (2006) From solvent fluctuations to quantitative redox properties of quinones in methanol and acetonitrile. *Angew. Chem. Intl. Ed.* **45**, p. 1936

37. R. Car and M. Parrinello (1985) Unified approach for molecular-dynamics and density-functional theory. *Phys. Rev. Lett.* **55**, p. 2471
38. V. May and O. Kühn Eds. (2004) Charge and energy transfer dynamics in molecular systems. Wiley-VH: 2nd edition
39. M. Tachiya (1993) Generalization of the marcus equation for the electron-transfer rate. *J. Phys. Chem.* **97**, p. 5911
40. R. W. Zwanzig (1954) High-temperature equation of state by a perturbation method.1. Nonpolar gases. *J. Chem. Phys.* **22**, p. 1420
41. J. P. Valleau and G. M. Torrie (1977) In *Modern Theoretical Chemistry*, Berne; B. J. Ed., vol. 5 Plenum, New York
42. D. Frenkel and B. Smit Eds. (1996) *Understanding Molecular Simulation – From Algorithms to Applications.* Academic Press: San Diego
43. J. G. Kirkwood (1935) Statistical Mechanics of Fluid Mixtures. *J. Chem. Phys.* **3**, p. 300
44. M. Souaille and B. Roux (2001) Extension to the weighted histogram analysis method: combining umbrella sampling with free energy calculations. *Comp. Phys. Comm.* **135**, p. 40
45. M. Tachiya (1989) Relation between the electron-transfer rate and the free-energy change of reaction. *J. Phys. Chem.* **93**, p. 7050
46. A. J. Bard and L. R. Faulkner, Eds. (2001) *Electrochemical Methods.* John Wiley & Sons, 2nd ed.
47. CPMD Version 3.x, The CPMD consortium, `http://www.cpmd.org`, MPI für Festkörperforschung and the IBM Zurich Research Laboratory
48. N. Troullier and J. Martins (1991) Efficient pseudopotentials for plane-wave calculations. *Phys. Rev. B* **43**, p. 1993
49. J. VandeVondele, M. Krack, F. Mohamed, M. Parrinello, T. Chassaing, and J. Hutter (2005) QUICKSTEP: Fast and accurate density functional calculations using a mixed Gaussian and plane waves approach. *Comp. Phys. Comm.* **167**, p. 103
50. The CP2K developers group, http://cp2k.berlios.de/ (2005)
51. C. Hartwigsen, S. Goedecker, and J. Hutter (1998) Relativistic separable dual-space Gaussian pseudopotentials from H to Rn. *Phys. Rev. B* **58**, p. 3641
52. M. Sprik and G. Ciccotti (1998) Free energy from constrained molecular dynamics. *J. Chem. Phys.* **109**, p. 7737
53. P. L. Geissler, C. Dellago, and D. Chandler (1999) Chemical dynamics of the protonated water trimer analyzed by transition path sampling. *Phys. Chem. Chem. Phys.* **1**, p. 1317
54. P. L. Geissler, C. Dellago, and D. Chandler (1999) Kinetic pathways of ion pair dissociation in water. *J. Phys. Chem. B* **103**, p. 3706
55. G. Hummer, L. R. Pratt, and A. E. Garcia (1998) Molecular theories and simulation of ions and polar molecules in water. *J. Phys. Chem. A* **102**, p. 7885

Advanced Car–Parrinello Techniques: Path Integrals and Nonadiabaticity in Condensed Matter Simulations

D. Marx

Lehrstuhl für Theoretische Chemie, Ruhr–Universität Bochum 44780 Bochum, Germany
Dominik.Marx@theochem.rub.de



Dominik Marx

D. Marx: *Advanced Car–Parrinello Techniques: Path Integrals and Nonadiabaticity in Condensed Matter Simulations*, Lect. Notes Phys. **704**, 507–539 (2006)
DOI 10.1007/3-540-35284-8_19 © Springer-Verlag Berlin Heidelberg 2006

1 Setting the Stage: Basic Car–Parrinello Molecular Dynamics 509

2 Advanced Techniques: Beyond Classical Nuclei and Electronic Ground-States 512

2.1 Introduction 512
2.2 Dealing with Quantum Nuclei 512
2.3 Dealing with Excited Electronic States 522

3 Summary and Outlook 534

References 535

Extensions of Car–Parrinello (CP) ab initio molecular dynamics are presented for efficient treatments of nuclear quantum effects and electronically nonadiabatic processes in the realm of condensed matter simulations. Ab initio path integrals, being a combination of CP propagation of the electrons in conjunction with path integral MD sampling of the nuclei, allow to investigate quantum phenomena, such as the influence of zero-point motion and proton tunneling, in chemically complex systems. Nonadiabatic ab initio simulations rely on the coupling of the Kohn-Sham ground state, S_0, and the first excited electronic state, S_1, obtained within the restricted open-shell Kohn-Sham (ROKS) approach using Tully's surface hopping algorithm. The efficient evaluation of the nonadiabatic couplings together with an "on-the-fly" updating scheme makes possible nonadiabatic ab initio simulations of systems of similar complexity as those typically studied by ground-state CP methods. This method is thus ideally suited to study photoinduced reactions of large molecular systems, particularly in condensed phases.

1 Setting the Stage: Basic Car–Parrinello Molecular Dynamics

The reign of traditional molecular dynamics *and* electronic structure methods has been greatly extended by the family of techniques that is now called "*ab initio* molecular dynamics" [83]. The basic idea underlying every *ab initio* molecular dynamics method is to compute the forces acting on the nuclei from electronic structure calculations that are performed "on-the-fly" as the molecular dynamics trajectory is generated. In this way, the electronic variables are not integrated out beforehand, but are considered as active degrees of freedom. This implies that, given a suitable approximate solution of the many-electron problem, also "chemically complex" systems including chemical reactions in condensed matter can be handled by molecular dynamics. A non-obvious approach to cut down the computational expenses of molecular dynamics which includes explicitly the electrons has been proposed twenty years ago by Car and Parrinello [18]. Here, only the basic notions will be reviewed in a concise form, see for instance [44, 53, 83, 95, 99, 107] for various in-depth review articles and detailed Lecture Notes, such that two techniques going beyond the standard Car–Parrinello scheme can be introduced in Sect. 2.

The basic idea of the Car–Parrinello approach [18] can be viewed to exploit the quantum-mechanical adiabatic time-scale separation of fast electronic and slow nuclear motion by transforming that into classical-mechanical adiabatic energy-scale separation in the framework of dynamical systems theory. In order to achieve this goal the two-component quantum/classical problem is mapped onto a two-component purely classical problem with two separate energy scales at the expense of loosing the explicit time-dependence of the quantum subsystem dynamics. Furthermore, the central quantity, the energy of the electronic subsystem $\langle \Psi_0 | \mathcal{H}_e | \Psi_0 \rangle$ evaluated with some wavefunction Ψ_0,

is certainly a function of the classical nuclear positions $\{\mathbf{R}_I\}$ for a given electronic Hamiltonian

$$\mathcal{H}_\mathrm{e} = -\sum_i \frac{\hbar^2}{2m_\mathrm{e}}\nabla_i^2 + \sum_{i<j}\frac{\mathrm{e}^2}{|\mathbf{r}_i-\mathbf{r}_j|} - \sum_{I,i}\frac{\mathrm{e}^2 Z_I}{|\mathbf{R}_I-\mathbf{r}_i|} + \sum_{I<J}\frac{\mathrm{e}^2 Z_I Z_J}{|\mathbf{R}_I-\mathbf{R}_J|} \quad (1)$$

within the Born–Oppenheimer approximation. But at the same time it can also be considered to be a functional of the wavefunction Ψ_0 and thus of a set of (one-particle) orbitals $\{\phi_i\}$ used to build up the (many-body) wavefunction, which is for instance a single Slater determinant $\Psi_0 = \det\{\phi_i\}$ in the case of Kohn–Sham [34,93] or Hartree–Fock [120] electronic structure theories. Now, in classical mechanics the force on the nuclei is obtained from the derivative of a Lagrangian with respect to the nuclear positions. This suggests that a functional derivative with respect to the orbitals, which are interpreted as classical fields, might yield the force on the orbitals, given a suitable Lagrangian. In addition, possible constraints within the set of orbitals have to be imposed, such as e.g. orthonormality (or generalized orthonormality conditions that include an overlap matrix).

Car and Parrinello postulated the following class of Lagrangians [18]

$$\mathcal{L}_\mathrm{CP} = \underbrace{\sum_I \frac{1}{2}M_I\dot{\mathbf{R}}_I^2 + \sum_i \frac{1}{2}\mu_i\left\langle\dot{\phi}_i\middle|\dot{\phi}_i\right\rangle}_{\text{kinetic energy}} - \underbrace{\langle\Psi_0|\mathcal{H}_\mathrm{e}|\Psi_0\rangle}_{\text{potential energy}} + \underbrace{\textit{constraints}}_{\text{orthonormality}} \quad (2)$$

to serve this purpose. The corresponding Newtonian equations of motion are obtained from the associated Euler–Lagrange equations

$$\frac{d}{dt}\frac{\partial\mathcal{L}}{\partial\dot{\mathbf{R}}_I} = \frac{\partial\mathcal{L}}{\partial\mathbf{R}_I} \quad (3)$$

$$\frac{d}{dt}\frac{\delta\mathcal{L}}{\delta\dot{\phi}_i^\star} = \frac{\delta\mathcal{L}}{\delta\phi_i^\star} \quad (4)$$

like in classical mechanics, but here for both the nuclear positions and the orbitals; note $\phi_i^\star = \langle\phi_i|$ and that the constraints are holonomic. Following this line of ideas, generic Car–Parrinello equations of motion are found to be of the form

$$M_I\ddot{\mathbf{R}}_I(t) = -\frac{\partial}{\partial\mathbf{R}_I}\langle\Psi_0|\mathcal{H}_\mathrm{e}|\Psi_0\rangle + \frac{\partial}{\partial\mathbf{R}_I}\{\textit{constraints}\} \quad (5)$$

$$\mu_i\ddot{\phi}_i(t) = -\frac{\delta}{\delta\phi_i^\star}\langle\Psi_0|\mathcal{H}_\mathrm{e}|\Psi_0\rangle + \frac{\delta}{\delta\phi_i^\star}\{\textit{constraints}\} \quad (6)$$

where μ_i $(=\mu)$ are the "fictitious masses" or inertia parameters assigned to the orbital degrees of freedom; the units of the mass parameter μ are energy times a squared time for reasons of dimensionality. Note that the constraints

within the total wavefunction lead to "constraint forces" in the equations of motion. Note also that these constraints

$$constraints = constraints \ \ (\{\phi_i\}, \{\mathbf{R}_I\}) \tag{7}$$

might be a function of both the set of orbitals $\{\phi_i\}$ and the nuclear positions $\{\mathbf{R}_I\}$. These dependencies have to be taken into account properly in deriving the Car–Parrinello equations of motion following from (2) using (3)–(4), see e.g. [65] for a case with an additional dependence of the wavefunction constraint on nuclear positions.

A particularly efficient and thus useful combination is to express the generic Car–Parrinello Lagrangian (2) in the framework of Kohn–Sham density functional theory [34, 93] as originally done [18], although other electronic structure methods such as Hartree–Fock or correlated (post–HF) methods [120] could be used as well. This Lagrangian reads

$$\begin{aligned}\mathcal{L}_{\mathrm{CP}} = &\sum_I \frac{1}{2} M_I \dot{\mathbf{R}}_I^2 + \sum_i \frac{1}{2} \mu_i \left\langle \dot{\phi}_i(\mathbf{r}) \middle| \dot{\phi}_i(\mathbf{r}) \right\rangle - E^{\mathrm{KS}}\left[\{\phi_i\}, \{\mathbf{R}_I\}\right] \\ &+ \sum_{ij} \Lambda_{ij} \left(\langle \phi_i(\mathbf{r}) \, | \phi_j(\mathbf{r}) \rangle - \delta_{ij} \right)\end{aligned} \tag{8}$$

where $E^{\mathrm{KS}}\left[\{\phi_i\}, \{\mathbf{R}_I\}\right]$ denotes the Kohn–Sham functional [34, 93] and the Lagrange multipliers $\{\Lambda_{ij}\}$ impose the orthogonality of the Kohn–Sham orbitals $\{\phi_i\}$. Within the software package `CPMD` [56] the Kohn–Sham problem is solved in terms of a pseudopotential/plane wave representation [99] as explained in detail in [83].

According to the Car–Parrinello equations of motion, the nuclei evolve in time at a certain (instantaneous) physical temperature $\propto \sum_I M_I \dot{\mathbf{R}}_I^2$, whereas a "fictitious temperature" $\propto \sum_i \mu_i \langle \dot{\phi}_i | \dot{\phi}_i \rangle$ is associated to the electronic degrees of freedom. In this terminology, "low electronic temperature" or "cold electrons" means that the electronic subsystem is close to its instantaneous minimum energy $\min_{\{\phi_i\}} \langle \Psi_0 | \mathcal{H}_{\mathrm{e}} | \Psi_0 \rangle$, i.e. close to the exact Born–Oppenheimer surface. Thus, a ground-state wavefunction optimized for the initial configuration of the nuclei will stay close to its ground state also during time evolution if it is kept at a sufficiently low temperature.

The remaining task is to separate in practice nuclear and electronic motion such that the fast electronic subsystem stays cold also for long times but still follows the slow nuclear motion adiabatically (or instantaneously). Simultaneously, the nuclei are nevertheless kept at a much higher temperature. This can be achieved in nonlinear classical dynamics via decoupling of the two subsystems and (quasi-) adiabatic time evolution. This is possible if the power spectra stemming from both dynamics do not have substantial overlap in the frequency domain so that energy transfer from the "hot nuclei" to the "cold electrons" becomes practically impossible on the relevant time scales. This amounts in other words to imposing and maintaining a metastability

condition in a complex dynamical system for sufficiently long times. How and to which extent this is possible in practice was investigated in detail in an important investigation based on well-controlled model systems [96, 97] (see also Sects. 3.2 and 3.3 in [107]), with more mathematical rigor in [10], and in terms of a generalization to a second level of adiabaticity in [80].

2 Advanced Techniques: Beyond Classical Nuclei and Electronic Ground-States

2.1 Introduction

Within the standard Car–Parrinello technique [18] just introduced, classical nuclei evolve adiabatically in the electronic ground state in the microcanonical ensemble. This combination allows already a multitude of applications, see e.g. [83] for an overview, but many circumstances exist where the underlying approximations are unsatisfactory or even break down. Among these cases are situations where

- light nuclei are involved in crucial steps of a process, such as in studies of proton transfer processes, hydrogen-bonded systems, or muonium impurities;
- dynamical motion occurs in excited states, such as photoinduced defects in solids or chromophores after photoexcitation events.

A way to perform *ab initio* simulations beyond the classical approximation for nuclei with the help of path integrals is presented in Sect. 2.2, whereas Sect. 2.3 is devoted to introduce techniques that allow simulations beyond the adiabatic approximation. Both techniques were implemented in the `CPMD` software package [56, 83].

2.2 Dealing with Quantum Nuclei

Introduction

Up to this point the nuclei were approximated as classical point particles as customarily done in standard molecular dynamics. There are, however, many situations where quantum dispersion broadening and tunneling effects play an important role and cannot be neglected if the simulation aims at being realistic – which is the generic goal of *ab initio* simulations. The *ab initio* path integral technique [72] and its extension to quasiclassical time evolution [80] is able to cope with such situations at finite temperatures. The central idea is to quantize the nuclei using Feynman's path integrals and at the same time to include the electronic degrees of freedom akin to *ab initio* molecular dynamics – that is "on-the-fly". The main ingredients and approximations underlying the *ab initio* path integral approach [72, 74, 77, 122] are

- the adiabatic separation of electrons and nuclei where the electrons are kept in their ground state without any coupling to electronically excited states (Born–Oppenheimer or "clamped-nuclei" approximation),
- using a particular approximate electronic structure theory in order to calculate the interactions,
- approximating the continuous path integral for the nuclei by a finite discretization (Trotter factorization) and neglecting the indistinguishability of identical nuclei (Boltzmann statistics), and
- using finite supercells with periodic boundary conditions and finite sampling times (finite-size and finite-time effects) as usual.

Thus, quantum effects such as zero-point motion and tunneling as well as thermal fluctuations are included at some preset temperature without further simplifications consisting e.g. in quasiclassical or quasiharmonic approximations, restricting the Hilbert space, or in artificially reducing the dimensionality of the problem.

Ab Initio Path Integrals: Statics

For the purpose of introducing *ab initio* path integrals [72] it is convenient to start directly with Feynman's formulation of quantum-statistical mechanics in terms of path integrals as opposed to Schrödinger's formulation in terms of wavefunctions. For a general introduction to path integrals the reader is referred to standard textbooks [35, 36, 60], whereas their use in numerical simulations is discussed for instance in [19–21, 46, 78, 115, 124].

The derivation of the expressions for *ab initio* path integrals is based on the non-relativistic standard Hamiltonian

$$\mathcal{H} = -\sum_I \frac{\hbar^2}{2M_I}\nabla_I^2 + \mathcal{H}_\mathrm{e} \ , \tag{9}$$

where the quantum-mechanical kinetic energy operator of the nuclei has been introduced in addition to the electronic part (1). The corresponding canonical partition function of a collection of interacting nuclei with positions $\mathbf{R} = \{\mathbf{R}_I\}$ and electrons $\mathbf{r} = \{\mathbf{r}_i\}$ can be written as a path integral

$$\mathcal{Z} = \oint' \mathcal{D}\mathbf{R} \oint' \mathcal{D}\mathbf{r} \exp\left[-\frac{1}{\hbar}\int_0^{\hbar\beta} d\tau\, \mathcal{L}_\mathrm{E}\left(\{\dot{\mathbf{R}}_I(\tau)\}, \{\mathbf{R}_I(\tau)\}; \{\dot{\mathbf{r}}_i(\tau)\}, \{\mathbf{r}_i(\tau)\}\right)\right] \tag{10}$$

where

$$\begin{aligned}\mathcal{L}_\mathrm{E} &= T(\dot{\mathbf{R}}) + V(\mathbf{R}) + T(\dot{\mathbf{r}}) + V(\mathbf{r}) + V(\mathbf{R},\mathbf{r}) \\ &= \sum_I \frac{1}{2} M_I \left(\frac{d\mathbf{R}_I}{d\tau}\right)^2 + \sum_{I<J} \frac{\mathrm{e}^2 Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|} \\ &\quad + \sum_i \frac{1}{2} m_\mathrm{e} \left(\frac{d\mathbf{r}_i}{d\tau}\right)^2 + \sum_{i<j} \frac{\mathrm{e}^2}{|\mathbf{r}_i - \mathbf{r}_j|} - \sum_{I,i} \frac{\mathrm{e}^2 Z_I}{|\mathbf{R}_I - \mathbf{r}_i|} \ ,\end{aligned} \tag{11}$$

denotes the *Euclidean* Lagrangian. The primes in (10) indicate that the proper sums over all permutations corresponding to Bose–Einstein and/or Fermi–Dirac statistics have to be included. It is important to note that in (10) and (11) the positions $\{\mathbf{R}_I\}$ and $\{\mathbf{r}_i\}$ are not operators but simply *functions* of the imaginary time $\tau \in [0, \hbar\beta]$ which parameterizes fluctuations around the classical path. This implies that the dots denote here derivatives with respect to imaginary time τ as defined in (11). According to (10) exact quantum mechanics at finite temperature $T = 1/k_{\mathrm{B}}\beta$ is recovered if all closed paths $[\{\mathbf{R}_I\};\{\mathbf{r}_i\}]$ of "length" $\hbar\beta$ are summed up and weighted with the exponential of the Euclidean action measured in units of $\hbar$; atomic units will be used in the following. The partial trace over the electronic subsystem can formally be written down exactly

$$\mathcal{Z} = \oint' \mathcal{D}\mathbf{R} \exp\left[-\int_0^\beta d\tau \left(T(\dot{\mathbf{R}}) + V(\mathbf{R})\right)\right] \mathcal{Z}\left[\mathbf{R}\right] \ , \tag{12}$$

with the aid of the influence functional [35, 60]

$$\mathcal{Z}[\mathbf{R}] = \oint' \mathcal{D}\mathbf{r} \exp\left[-\int_0^\beta d\tau \left(T(\dot{\mathbf{r}}) + V(\mathbf{r}) + V(\mathbf{R},\mathbf{r})\right)\right] \ . \tag{13}$$

Note that $\mathcal{Z}[\mathbf{R}]$ is a complicated and unknown functional for a given nuclear path configuration $[\{\mathbf{R}_I\}]$. As a consequence the interactions between the nuclei become highly nonlocal in imaginary time due to memory effects.

In the standard Born–Oppenheimer or "clamped nuclei" approximation, see [63] for instance, the nuclei are frozen in some configuration and the complete electronic problem is solved for this single static configuration. In addition to the nondiagonal correction terms that are already neglected in the adiabatic approximation, the diagonal terms are now neglected as well. Thus the potential for the nuclear motion is simply defined as the bare electronic eigenvalues obtained from a series of fixed nuclear configurations.

In the statistical mechanics formulation of the problem (12)–(13) the Born–Oppenheimer approximation amounts to a "quenched average": at imaginary time τ the nuclei are frozen at a particular configuration $\mathbf{R}(\tau)$ and the electrons explore their configuration space subject only to that single configuration. This implies that the electronic degrees of freedom at different imaginary times τ and τ' become completely decoupled. Thus, the influence functional $\mathcal{Z}[\mathbf{R}]$ has to be local in τ and becomes particularly simple; a discussion of adiabatic corrections in the path integral formulation can be found in [12]. For each τ the influence functional $\mathcal{Z}[\mathbf{R}]$ is given by the partition function of the electronic subsystem evaluated for the respective nuclear configuration $\mathbf{R}(\tau)$. Assuming that the temperature is small compared to the gap in the electronic spectrum $\{E_k\}$ as a result of solving

$$\mathcal{H}_{\mathrm{e}}\Psi_k = E_k(\mathbf{R})\Psi_k(\mathbf{r};\mathbf{R}) \ , \tag{14}$$

for nuclear positions $\mathbf{R}(\tau)$ at all imaginary times $\tau \in [0, \beta]$, only the electronic ground state with energy $E_{k=0}(\mathbf{R}(\tau))$ is populated. This electronic ground state dominance leads to the following simple expression

$$\mathcal{Z}[\mathbf{R}]_{\mathrm{BO}} = \exp\left[-\int_0^\beta d\tau E_0(\mathbf{R}(\tau))\right] \quad , \tag{15}$$

which yields the final result

$$\mathcal{Z}_{\mathrm{BO}} = \oint \mathcal{D}\mathbf{R} \exp\left[-\int_0^\beta d\tau \left(T(\dot{\mathbf{R}}) + V(\mathbf{R}) + E_0(\mathbf{R})\right)\right] \quad . \tag{16}$$

Here nuclear exchange is neglected by assuming that the nuclei are distinguishable so that they can be treated within Boltzmann statistics, which corresponds to the Hartree approximation for the *nuclear* density matrix. The presentation given here follows [74] and alternative derivations were given in Sect. 2.3 of [20] and in the appendix of [85]. There, a wavefunction basis instead of the position basis as in (13) was formally used in order to evaluate the influence functional due to the electrons.

The partition function (16) together with the Coulomb Hamiltonian (9) leads after applying the lowest-order Trotter factorization [60] to the following discretized expression

$$\begin{aligned}\mathcal{Z}_{\mathrm{BO}} = \lim_{P\to\infty} \prod_{s=1}^{P} \prod_{I=1}^{N} \left[\left(\frac{M_I P}{2\pi\beta}\right)^{3/2} \int d\mathbf{R}_I^{(s)}\right] \\ \times \exp\left[-\beta \sum_{s=1}^{P} \left\{\sum_{I=1}^{N} \frac{1}{2} M_I \omega_P^2 \left(\mathbf{R}_I^{(s)} - \mathbf{R}_I^{(s+1)}\right)^2 + \frac{1}{P} E_0\left(\{\mathbf{R}_I\}^{(s)}\right)\right\}\right]\end{aligned} \tag{17}$$

for the path integral with $\omega_P^2 = P/\beta^2$. Thus, the continuous parameter $\tau \in [0, \beta]$ is discretized using P so-called Trotter slices or "time slices" $s = 1, \ldots, P$ of "duration" $\Delta\tau = \beta/P$. The paths

$$\begin{aligned}\left\{\{\mathbf{R}_I\}^{(s)}\right\} &= \left(\{\mathbf{R}_I\}^{(1)}; \ldots; \{\mathbf{R}_I\}^{(P)}\right) \\ &= \left(\mathbf{R}_1^{(1)}, \ldots, \mathbf{R}_N^{(1)}; \ldots; \mathbf{R}_1^{(P)}, \ldots, \mathbf{R}_N^{(P)}\right)\end{aligned} \tag{18}$$

have to be closed due to the trace condition, i.e. they are periodic in imaginary time τ which implies $\mathbf{R}_I(0) \equiv \mathbf{R}_I(\beta)$ and thus $\mathbf{R}_I^{(P+1)} = \mathbf{R}_I^{(1)}$. Note that (17) is an *exact* reformulation of (16) in the limit of an infinitely fine discretization $P \to \infty$ of the paths.

The effective classical partition function (17) with a fixed discretization P is isomorphic to that for N polymers each comprised by P monomers [19, 21, 46]. Each quantum degree of freedom is found to be represented by a ring polymer or necklace. The intrapolymeric interactions stem from the kinetic

energy $T(\dot{\mathbf{R}})$ and consist of harmonic nearest-neighbor couplings $\propto \omega_P$ along the closed chain. The interpolymeric interaction is given by the scaled potential $E_0^{(s)}/P$ which is only evaluated for configurations $\{\mathbf{R}_I\}^{(s)}$ at the *same* imaginary time slice s.

In order to evaluate operators based on an expression like (17) most numerical path integral schemes utilize Metropolis Monte Carlo sampling with the effective potential

$$V_{\text{eff}} = \sum_{s=1}^{P} \left\{ \sum_{I=1}^{N} \frac{1}{2} M_I \omega_P^2 \left(\mathbf{R}_I^{(s)} - \mathbf{R}_I^{(s+1)} \right)^2 + \frac{1}{P} E_0 \left(\{\mathbf{R}_I\}^{(s)} \right) \right\} \tag{19}$$

of the isomorphic classical system [19–21, 46, 78, 115, 124]. Molecular dynamics techniques were also proposed in order to sample configuration space, see [11, 54, 94, 101, 102] for pioneering work and [124] for an authoritative review. Formally a Lagrangian can be obtained from the expression (19) by extending it

$$\mathcal{L}_{\text{PIMD}} = \sum_{s=1}^{P} \left\{ \sum_{I=1}^{N} \left(\frac{\mathbf{P}_I^{(s)}}{2M_I'} - \frac{1}{2} M_I \omega_P^2 \left(\mathbf{R}_I^{(s)} - \mathbf{R}_I^{(s+1)} \right)^2 \right) - \frac{1}{P} E_0 \left(\{\mathbf{R}_I\}^{(s)} \right) \right\} \tag{20}$$

with $N \times P$ fictitious momenta $\mathbf{P}_I^{(s)}$ and corresponding (unphysical) fictitious masses M_I'. At this stage the time dependence of positions and momenta and thus the time evolution in *phase space* as generated by (20) has no physical meaning. The sole use of "time" is to parameterize the deterministic dynamical exploration of *configuration space*. The trajectories of the positions in configuration space, can, however, be analyzed similar to the ones obtained from the stochastic dynamics that underlies the Monte Carlo method.

The crucial ingredient in *ab initio* [72, 74, 77, 122] as opposed to standard [19–21, 46, 78, 115, 124] path integral simulations consists in computing the interactions E_0 "on-the-fly" like in *ab initio* molecular dynamics. The first implementation [72] was based on the Car–Parrinello algorithm [18] combined with Kohn–Sham density functional theory [34, 93] which leads to the following extended Lagrangian

$$\begin{aligned} \mathcal{L}_{\text{AIPI}} = \frac{1}{P} \sum_{s=1}^{P} \Bigg\{ & \sum_i \mu \left\langle \dot{\phi}_i^{(s)} \middle| \dot{\phi}_i^{(s)} \right\rangle - E^{\text{KS}} \left[\{\phi_i\}^{(s)}, \{\mathbf{R}_I\}^{(s)} \right] \\ & + \sum_{ij} \Lambda_{ij}^{(s)} \left(\left\langle \phi_i^{(s)} \middle| \phi_j^{(s)} \right\rangle - \delta_{ij} \right) \Bigg\} \\ & + \sum_{s=1}^{P} \left\{ \sum_I \frac{1}{2} M_I' \left(\dot{\mathbf{R}}_I^{(s)} \right)^2 - \sum_{I=1}^{N} \frac{1}{2} M_I \omega_P^2 \left(\mathbf{R}_I^{(s)} - \mathbf{R}_I^{(s+1)} \right)^2 \right\}, \end{aligned} \tag{21}$$

where the interaction energy $E^{\text{KS}}[\{\phi_i\}^{(s)}, \{\mathbf{R}_I\}^{(s)}]$ is the Kohn–Sham energy evaluated for the nuclear configuration at time slice s; note that here and in

the following the dots denote derivatives with respect to propagation time t and that $E_0^{\mathrm{KS}} = \min E^{\mathrm{KS}}$. The standard Car–Parrinello Lagrangian (8) is recovered in the limit $P = 1$ which corresponds to classical nuclei. Mixed classical/quantum systems can be treated by representing an arbitrary subset of the nuclei in (21) with only one time slice [125].

This simplest formulation of *ab initio* path integrals, however, is insufficient for the following reason: ergodicity of the trajectories and adiabaticity in the sense of Car–Parrinello simulations are not guaranteed. It is known since the very first molecular dynamics computer experiments that quasiharmonic systems (such as coupled stiff harmonic oscillators subject to weak anharmonicities, i.e. the famous Fermi–Pasta–Ulam chains) can easily lead to nonergodic behavior in the sampling of phase space [42]. Similarly "microcanonical" path integral molecular dynamics simulations might lead to an insufficient exploration of configuration space depending on the parameters [54]. The severity of this nonergodicity problem is governed by the stiffness of the harmonic intrachain coupling $\propto \omega_P$ and the anharmonicity of the overall potential surface $\propto E^{\mathrm{KS}}/P$ which establishes the coupling of the modes. For a better and better discretization P the harmonic energy term dominates according to $\sim P$ whereas the mode-mixing coupling decreases like $\sim 1/P$. This problem can be cured by attaching so-called Nosé–Hoover chain thermostats [69] to all path integral degrees of freedom [83, 121, 122], see Sect. 2.2.

The second issue is related to the separation of the power spectra associated to nuclear and electronic subsystems during Car–Parrinello *ab initio* molecular dynamics which is instrumental for maintaining adiabaticity, see Sect. 1. In *ab initio* molecular dynamics with classical nuclei the highest phonon or vibrational frequency $\omega_{\mathrm{n}}^{\max}$ is dictated by the physics of the system [83]. This means in particular that an upper limit is given by stiff intramolecular vibrations which do not exceed $\omega_{\mathrm{n}}^{\max} \leq 5000\ \mathrm{cm}^{-1}$ or 150 THz. In *ab initio* path integral simulations, on the contrary, $\omega_{\mathrm{n}}^{\max}$ is given by ω_P which actually diverges with increasing discretization as $\sim \sqrt{P}$. The simplest counteraction would be to compensate this artifact by decreasing the fictitious electron mass μ until the power spectra are again separated for a fixed value of P and thus ω_P. This, however, would lead to a prohibitively small time step because $\Delta t^{\max} \propto \sqrt{\mu}$. This dilemma can be solved by thermostatting the electronic degrees of freedom as well [72, 74, 122] as proposed earlier for systems with small or vanishing band gaps such as metals [9].

Finally, it is known that diagonalizing the harmonic spring interaction in (21) leads to more efficient propagators [121,122]. One of these transformation and the resulting Nosé–Hoover chain thermostatted equations of motion will be outlined in the following section, see in particular (33)–(39). In addition to keeping the average temperature fixed it is also possible to generate path trajectories in the isobaric-isothermal NpT ensemble [71,124]. Instead of using Car–Parrinello fictitious dynamics in order to evaluate the interaction energy in (20), which is implemented in the `CPMD` package [56,83], it is evident that also an iterative minimization of the Kohn–Sham energy functional could

be used. This route eliminates the adiabaticity problem and was taken up e.g. in [1, 23, 59, 86, 87, 116, 117].

A final observation concerning parallel supercomputers might be useful. It is evident from the Lagrangian (21) and the resulting equations of motion (33)–(39) that most of the numerical workload comes from calculating the *ab initio* forces on the nuclei. Given a fixed path configuration (18) the P underlying electronic structure problems are independent from each other and can be solved without communication on P nodes of a distributed memory machine. Communication is only necessary to send the final result, essentially the forces, to a special node that computes the quantum kinetic contribution to the energy and integrates finally the equations of motions. It is even conceivable to distribute this task on different supercomputers, i.e. "meta-computing" is within reach for such calculations. Thus, the algorithm is "embarrassingly parallel" provided that the memory per node is sufficient to solve the complete Kohn–Sham problem at a given time slice. If this is not the case the electronic structure calculation itself has to be parallelized on another hierarchical level as implemented in `CPMD` [56, 83].

Ab Initio Path Centroids: Dynamics

Initially the molecular dynamics approach to path integral simulations was invented merely as a trick in order to sample configuration space similar to the Monte Carlo method. This perception changed recently with the introduction of the so-called "centroid molecular dynamics" technique [13], see [14–16, 103–106, 132] for background information. In a nutshell it is found that the time evolution of the centers of mass or centroids

$$\mathbf{R}_I^c(t) = \frac{1}{P} \sum_{s'=1}^{P} \mathbf{R}_I^{(s')}(t) \tag{22}$$

of the closed Feynman paths that represent the quantum nuclei contains quasiclassical information about the true quantum dynamics. The centroid molecular dynamics approach can be shown to be exact for harmonic potentials and to have the correct classical limit. The path centroids move in an effective potential which is generated by all the other modes of the paths at the given temperature. This effective potential thus includes the effects of quantum fluctuations on the (quasiclassical) time evolution of the centroid degrees of freedom. Roughly speaking the trajectory of the path centroids can be regarded as a classical trajectory of the system, which is approximately "renormalized" due to quantum effects.

The original centroid molecular dynamics technique [13–16, 132] relies on the use of model potentials as the standard time-independent path integral simulations. This limitation was overcome independently in [80, 98] by combining *ab initio* path integrals with centroid molecular dynamics. The resulting technique, *ab initio* centroid molecular dynamics can be considered

as a quasiclassical generalization of standard *ab initio* molecular dynamics. At the same time, it preserves the virtues of the *ab initio* path integral technique [72,74,77,122] to generate exact time-independent quantum equilibrium averages.

Here, the so-called adiabatic formulation [16, 17, 70] of *ab initio* centroid molecular dynamics [80] is discussed. In close analogy to *ab initio* molecular dynamics with classical nuclei also the effective centroid potential is generated "on-the-fly" as the centroids are propagated. This is achieved by singling out the centroid coordinates in terms of a normal mode transformation [24] and accelerating the dynamics of all non-centroid modes artificially by assigning appropriate fictitious masses. At the same time, the fictitious electron dynamics *à la* Car–Parrinello is kept in order to calculate efficiently the *ab initio* forces on *all* modes from the electronic structure. This makes it necessary to maintain two levels of adiabaticity in the course of simulations, see Sect. 2.1 of [80] for a theoretical analysis of that issue.

The partition function (17), formulated in the so-called "primitive" path variables $\{\mathbf{R}_I\}^{(s)}$, is first transformed [122,124] to a representation in terms of the normal modes $\{\mathbf{u}_I\}^{(s)}$, which diagonalize the harmonic nearest-neighbor harmonic coupling [24]. The transformation follows from the Fourier expansion of a cyclic path

$$\mathbf{R}_I^{(s)} = \sum_{s'=1}^{P} \mathbf{a}_I^{(s')} \exp\left[2\pi i (s-1)(s'-1)/P\right] \quad , \tag{23}$$

where the coefficients $\{\mathbf{a}_I\}^{(s)}$ are complex. The normal mode variables $\{\mathbf{u}_I\}^{(s)}$ are then given in terms of the expansion coefficients according to

$$\begin{aligned} \mathbf{u}_I^{(1)} &= \mathbf{a}_I^{(1)} \\ \mathbf{u}_I^{(P)} &= \mathbf{a}_I^{((P+2)/2)} \\ \mathbf{u}_I^{(2s-2)} &= \mathrm{Re}\ (\mathbf{a}_I^{(s)}) \\ \mathbf{u}_I^{(2s-1)} &= \mathrm{Im}\ (\mathbf{a}_I^{(s)}) \quad . \end{aligned} \tag{24}$$

Associated with the normal mode transformation is a set of normal mode frequencies $\{\lambda\}^{(s)}$ given by

$$\lambda^{(2s-1)} = \lambda^{(2s-2)} = 2P\left[1 - \cos\left(\frac{2\pi(s-1)}{P}\right)\right] \tag{25}$$

with $\lambda^{(1)} = 0$ and $\lambda^{(P)} = 4P$. Equation (23) is equivalent to direct diagonalization of the matrix

$$\mathbf{A}_{ss'} = 2\delta_{ss'} - \delta_{s,s'-1} - \delta_{s,s'+1} \tag{26}$$

with the path periodicity condition $\mathbf{A}_{s0} = \mathbf{A}_{sP}$ and $\mathbf{A}_{s,P+1} = \mathbf{A}_{s1}$ and subsequent use of the unitary transformation matrix $\mathbf{U}$ to transform from the "primitive" variables $\{\mathbf{R}_I\}^{(s)}$ to the normal mode variables $\{\mathbf{u}_I\}^{(s)}$

$$\mathbf{R}_I^{(s)} = \sqrt{P} \sum_{s'=1}^{P} \mathbf{U}_{ss'}^{\dagger} \mathbf{u}_I^{(s')}$$
$$\mathbf{u}_I^{(s)} = \frac{1}{\sqrt{P}} \sum_{s'=1}^{P} \mathbf{U}_{ss'} \mathbf{R}_I^{(s')} \quad . \tag{27}$$

The eigenvalues of $\mathbf{A}$ when multiplied by P are precisely the normal mode frequencies $\{\lambda\}^{(s)}$. Since the transformation is unitary, its Jacobian is unity. Finally, it is convenient to define a set of normal mode masses

$$M_I^{(s)} = \lambda^{(s)} M_I \tag{28}$$

that vary along the imaginary time axis $s = 1, \ldots, P$, where $\lambda^{(1)} = 0$ for the centroid mode $\mathbf{u}_I^{(1)}$.

Based on these transformations the Lagrangian corresponding to the *ab initio* path integral expressed in normal modes is obtained [122]

$$\begin{aligned}\mathcal{L}_{\text{AIPI}} = \frac{1}{P} \sum_{s=1}^{P} \Bigg\{ & \sum_i \mu \left\langle \dot{\phi}_i^{(s)} \middle| \dot{\phi}_i^{(s)} \right\rangle - E^{\text{KS}} \left[\{\phi_i\}^{(s)}, \left\{ \mathbf{R}_I \left(\mathbf{u}_I^{(1)}, \ldots, \mathbf{u}_I^{(P)} \right) \right\}^{(s)} \right] \\ & + \sum_{ij} \Lambda_{ij}^{(s)} \left(\left\langle \phi_i^{(s)} \middle| \phi_j^{(s)} \right\rangle - \delta_{ij} \right) \Bigg\} \\ & + \sum_{s=1}^{P} \left\{ \sum_{I=1}^{N} \frac{1}{2} M_I'^{(s)} \left(\dot{\mathbf{u}}_I^{(s)} \right)^2 - \sum_{I=1}^{N} \frac{1}{2} M_I^{(s)} \omega_P^2 \left(\mathbf{u}_I^{(s)} \right)^2 \right\} \quad , \end{aligned} \tag{29}$$

where the masses $M_I'^{(s)}$ will be defined later, see (40). As indicated, the electronic energy $E^{(s)}$ is always evaluated in practice in terms of the "primitive" path variables $\{\mathbf{R}_I\}^{(s)}$ in Cartesian space. The necessary transformation to switch forth and back between "primitive" and normal mode variables is easily performed as given by the relations (27).

The chief advantage of the normal mode representation (27) for the present purpose is that the lowest-order normal mode $\mathbf{u}_I^{(1)}$

$$\mathbf{u}_I^{(1)} = \mathbf{R}_I^{\text{c}} = \frac{1}{P} \sum_{s'=1}^{P} \mathbf{R}_I^{(s')} \tag{30}$$

turns out to be identical to the centroid $\mathbf{R}_I^{\text{c}}$ of the path that represents the Ith nucleus. The centroid force can also be obtained from the matrix $\mathbf{U}$ according to [122]

$$\frac{\partial E}{\partial \mathbf{u}_I^{(1)}} = \frac{1}{P} \sum_{s'=1}^{P} \frac{\partial E^{(s')}}{\partial \mathbf{R}_I^{(s')}} \tag{31}$$

since $\mathbf{U}_{1s} = \mathbf{U}_{s1}^{\dagger} = 1/\sqrt{P}$ and the remaining normal mode forces are given by

$$\frac{\partial E}{\partial \mathbf{u}_I^{(s)}} = \frac{1}{\sqrt{P}} \sum_{s'=1}^{P} \mathbf{U}_{ss'} \frac{\partial E^{(s')}}{\partial \mathbf{R}_I^{(s')}} \quad \text{for } s = 2, \ldots, P \tag{32}$$

in terms of the "primitive" forces $-\partial E^{(s)}/\partial \mathbf{R}_I^{(s)}$. Here, E on the left-hand-side with no superscript (s) refers to the average electronic energy $E = (1/P)\sum_{s=1}^{P} E^{(s)}$ from which the forces have to be derived. Thus, the force (31) acting on each *centroid* variable $\mathbf{u}_I^{(1)}$, $I = 1, \ldots N$, is exactly the force averaged over imaginary time $s = 1, \ldots, P$, i.e. the *centroid force* on the Ith nucleus as already given in (2.21) of [122]. This is the desired relation which allows in centroid molecular dynamics the centroid forces to be simply obtained as the average force which acts on the lowest-order normal mode (30). The non-centroid normal modes $\mathbf{u}_I^{(s)}$, $s = 2, 3, \ldots, P$ of the paths establish the effective potential in which the centroid moves.

At this stage the equations of motion for adiabatic *ab initio* centroid molecular dynamics [80] can be obtained from the Euler–Lagrange equations. These equations of motion read

$$M_I'^{(1)} \ddot{\mathbf{u}}_I^{(1)} = -\frac{1}{P} \sum_{s=1}^{P} \frac{\partial E\left[\{\phi_i\}^{(s)}, \{\mathbf{R}_I\}^{(s)}\right]}{\partial \mathbf{R}_I^{(s)}} \tag{33}$$

$$\begin{aligned} M_I'^{(s)} \ddot{u}_{I,\alpha}^{(s)} = &-\frac{\partial}{\partial u_{I,\alpha}^{(s)}} \frac{1}{P} \sum_{s'=1}^{P} E\left[\{\phi_i\}^{(s')}, \left\{\mathbf{R}_I\left(\mathbf{u}_I^{(1)}, \ldots, \mathbf{u}_I^{(P)}\right)\right\}^{(s')}\right] \\ &-M_I^{(s)} \omega_P^2 u_{I,\alpha}^{(s)} - M_I'^{(s)} \dot{\xi}_{I,\alpha,1}^{(s)} \dot{u}_{I,\alpha}^{(s)}, \; s = 2, \ldots, P \end{aligned} \tag{34}$$

$$\begin{aligned} \mu \ddot{\phi}_i^{(s)} = &-\frac{\delta E\left[\{\phi_i\}^{(s)}, \{\mathbf{R}_I\}^{(s)}\right]}{\delta \phi_i^{\star(s)}} \\ &+\sum_j \Lambda_{ij}^{(s)} \phi_j^{(s)} - \mu \dot{\eta}_1^{(s)} \dot{\phi}_i^{(s)}, s = 1, \ldots, P \end{aligned} \tag{35}$$

where $u_{I,\alpha}^{(s)}$ denotes the Cartesian components of a given normal mode vector $\mathbf{u}_I^{(s)} = (u_{I,1}^{(s)}, u_{I,2}^{(s)}, u_{I,3}^{(s)})$. In the present scheme, independent Nosé–Hoover chain thermostats [69] of length K are coupled to all non-centroid mode degrees of freedom $s = 2, \ldots, P$

$$Q^{\mathrm{n}} \ddot{\xi}_{I,\alpha,1}^{(s)} = \left[M_I'^{(s)} \left(\dot{u}_{I,\alpha}^{(s)}\right)^2 - k_{\mathrm{B}}T\right] - Q^{\mathrm{n}} \dot{\xi}_{I,\alpha,1}^{(s)} \dot{\xi}_{I,\alpha,2}^{(s)} \tag{36}$$

$$\begin{aligned} Q^{\mathrm{n}} \ddot{\xi}_{I,\alpha,k}^{(s)} = &\left[Q^{\mathrm{n}} \left(\dot{\xi}_{I,\alpha,k-1}^{(s)}\right)^2 - k_{\mathrm{B}}T\right] \\ &-Q^{\mathrm{n}} \dot{\xi}_{I,\alpha,k}^{(s)} \dot{\xi}_{I,\alpha,k+1}^{(s)} \left(1 - \delta_{kK}\right), k = 2, \ldots, K \end{aligned} \tag{37}$$

and all orbitals at a given imaginary time slice s are thermostatted by one such thermostat chain of length L

$$Q_1^{\mathrm{e}}\ddot{\eta}_1^{(s)} = 2\left[\sum_i \mu \left\langle \phi_i^{(s)} \middle| \phi_i^{(s)} \right\rangle - T_{\mathrm{e}}^0\right] - Q_1^{\mathrm{e}}\dot{\eta}_1^{(s)}\dot{\eta}_2^{(s)} \tag{38}$$

$$Q_l^{\mathrm{e}}\ddot{\eta}_l^{(s)} = \left[Q_{l-1}^{\mathrm{e}}\left(\dot{\eta}_{l-1}^{(s)}\right)^2 - \frac{1}{\beta_{\mathrm{e}}}\right] - Q_l^{\mathrm{e}}\dot{\eta}_l^{(s)}\dot{\eta}_{l+1}^{(s)}\left(1-\delta_{lL}\right) \quad , l = 2,\ldots,L\,; \tag{39}$$

note that for standard *ab initio* path integral runs as discussed in the previous section the centroid mode should be thermostatted as well. The desired fictitious kinetic energy of the electronic subsystem T_{e}^0 can be determined based on a short equivalent classical Car–Parrinello run with $P = 1$ and using again the relation $1/\beta_{\mathrm{e}} = 2T_{\mathrm{e}}^0/6N_{\mathrm{e}}'$ where N_{e}' is the number of orbitals. The parameters $\{Q_l^{\mathrm{e}}\}$ are masses associated to the orbital thermostats, see [83] for a more detailed discussion, whereas the single mass/inertia parameter Q^{n} for the nuclei is determined by the harmonic interaction and is given by $Q^{\mathrm{n}} = k_{\mathrm{B}}T/\omega_P^2 = \beta/P$. The characteristic thermostat frequency of the electronic degrees of freedom ω_{e} should again lie above the frequency spectrum associated to the fictitious nuclear dynamics. These is the method that is implemented in the `CPMD` package [56, 83].

An important issue for adiabatic *ab initio* centroid molecular dynamics [80] is how to establish the time-scale separation of the non-centroid modes compared to the centroid modes. This is guaranteed if the *fictitious* normal mode masses $M_I'^{(s)}$ are taken to be

$$\begin{aligned} M_I'^{(1)} &= M_I \\ M_I'^{(s)} &= \gamma\, M_I^{(s)},\ s = 2,\ldots,P\,, \end{aligned} \tag{40}$$

where M_I is the *physical* nuclear mass, $M_I^{(s)}$ are the normal mode masses (28), and γ is the "centroid adiabaticity parameter"; note that this corrects a misprint of the definition of $M_I'^{(s)}$ for $s \geq 2$ in [80]. By choosing $0 < \gamma \ll 1$, the required time-scale separation between the centroid and non-centroid modes can be controlled so that the motion of the non-centroid modes is artificially accelerated, see Sect. 3 in [80] for a systematic study of the γ-dependence. Thus, the centroids with associated physical masses move quasiclassically in real-time in the centroid effective potential, whereas the fast dynamics of all other nuclear modes $s > 1$ is fictitious and serves only to generate the centroid effective potential "on-the-fly". In this sense γ (or rather γM_I) is similar to μ, the electronic adiabaticity parameter in Car–Parrinello molecular dynamics.

2.3 Dealing with Excited Electronic States

Introduction

Recent advances in density functional theory offer an efficient route to approximatively solving electronic structure problems even for excited states and thus opened the way for molecular dynamics involving excited states,

see [28] for a review focussing on this very topic. Here, the following scenarios implemented in CPMD [56, 83] will be considered:

- a single excited many-electron state is completely decoupled from all other states on the relevant timescale and thus evolves adiabatically;
- two such many-electron states interact explicitly via nonadiabatic couplings.

The first situation applies, for instance, to large-gap molecular systems which undergo an ultrafast chemical reaction in a single excited state as a result of a vertical HOMO/LUMO or instantaneous one-particle/one-hole photoexcitation, the initial steps of cis-trans isomerizations or tautomerizations being popular examples. The second scenario is a nontrivial generalization of the first one to cases where decay from a higher-lying state to a lower-lying state is crucial, important cases being radiationless decay mechanisms via conical intersections as encountered after cis-trans isomerization or tautomerization reactions have taken place.

A Single Excited State: Restricted Open-Shell Kohn–Sham Dynamics

For large-gap systems with well separated electronic states it might be desirable to single out a particular state in order to allow the nuclei to move on the associated excited state potential energy surface. Approaches that rely on fractional occupation numbers such as ensemble density functional theories and free energy functionals are difficult to adapt for cases where the symmetry and/or spin of the electronic state should be fixed [34].

A method that combines Roothaan's symmetry-adapted wavefunctions [109] with Kohn–Sham density functional theory was proposed in [43] and used to simulate a photoisomerization via molecular dynamics, see [51, 92] for further analysis of the method and for refinement. Viewed from Kohn–Sham theory this approach consists in building up the spin density of an open-shell system based on a symmetry-adapted wavefunction that is constructed from spin-restricted determinants, see also [48]. Viewed from the restricted open-shell Hartree–Fock (ROHF) theory *à la* Roothaan it amounts essentially to replacing Hartree–Fock exchange by an approximate exchange-correlation density functional, thus the name restricted open-shell Kohn–Sham (ROKS) method. Thus, the central idea is to impose symmetry constraints on the many-electron wavefunction by constructing a symmetry-adapted *multi-determinantal* wavefunction,

$$\Psi_{\mathrm{SA}} = \sum_{\mu} a_{\mu} \Phi_{\mu} \ , \tag{41}$$

from spin-restricted single Slater determinants, the so-called "microstates" Φ_{μ}. Following Roothaan [109], the expansion coefficients, a_{μ}, are chosen to be the Clebsch–Gordan coefficients for a given symmetry. The energy is then given by

$$E_{\mathrm{SA}} = \sum_{\mu} c_{\mu}(\{a_{\mu}\}) E_{\mu} \quad \text{with} \quad \sum_{\mu} c_{\mu} = 1 \ , \tag{42}$$

where E_{μ} is the total energy of the μth microstate. Within density functional theory this procedure leads to an explicitly orbital-dependent functional which was formulated for the first-excited singlet state S_1 in [43]. The relation of this approach to previous theories is discussed in some detail in [43]. In particular, the success of the closely-related Ziegler–Rauk–Baerends ΔSCF so-called "sum methods" [25, 119, 135] was an important stimulus. More recently, several papers [37–40, 49, 50, 55, 88] appeared that are similar in spirit to the method of [43]. In particular the approach of [37] and subsequent refinements can be viewed as a generalization to arbitrary spin states of the special case (S_1 state) worked out in [43]. In addition, the generalized methods [37–40] were derived within the framework of (ensemble) density functional theory, whereas the wavefunction perspective was the starting point in [43]. However, it should be acknowledged that is was demonstrated more recently that there is no Hohenberg–Kohn theorem for excited states [45]. Thus, *rigorously speaking*, no *exact* excited state functional of the excited state density exists even for a given level of excitation, whereas the excited state energy clearly is a functional of the *ground state* density by virtue of the Hohenberg–Kohn theorem [34, 93]. On the other hand there is definitively now a wealth of supporting evidence that *pragmatic approaches* to excited states within density functional theory are useful practical tools to extend *ab initio* molecular dynamics beyond the ground state.

In the following, the ROKS method is outlined with the focus to perform molecular dynamics in the S_1 state. Promoting one electron from the HOMO to the LUMO ("particle-hole excitation") in a closed-shell system with $2n$ electrons assigned to n doubly occupied orbitals (that is spin-restricted orbitals that have the same spatial part for both spin up α and spin down β electrons) leads to four different excited wavefunctions or determinants, see Fig. 1 for a sketch. Two states $|t_1\rangle$ and $|t_2\rangle$ are energetically degenerate triplets t whereas the two states $|m_1\rangle$ and $|m_2\rangle$ are not eigenfunctions of the total spin operator, $\hat{\mathcal{S}}^2$, and thus degenerate "mixed states" m. Note that the m states do not correspond – as is well known – to singlet states due to significant spin contamination despite the suggestive occupation pattern in Fig. 1 and that they are called "broken-symmetry states" mainly in the literature on magnetic interactions [57, 91, 100].

However, suitable Clebsch–Gordan projections of the mixed states $|m_1\rangle$ and $|m_2\rangle$ yield another triplet state $|t_3\rangle$ and the desired first excited singlet or S_1 state $|s_1\rangle$. Here, the ansatz [43] for the total energy of the S_1 state is given by

$$E_{S_1}^{\mathrm{ROKS}}[\{\phi_i\}] = 2E_m^{\mathrm{KS}}[\{\phi_i\}] - E_t^{\mathrm{KS}}[\{\phi_i\}] \tag{43}$$

where the energies of the mixed and triplet determinants

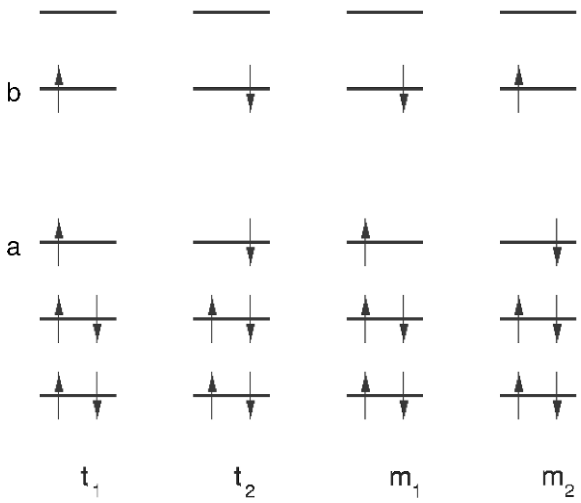


Fig. 1. Four possible determinants $|t_1\rangle$, $|t_2\rangle$, $|m_1\rangle$ and $|m_2\rangle$ as a result of the promotion of a single electron from the HOMO to the LUMO of a closed shell system yielding the singly occupied orbitals ϕ_a and ϕ_b, see text for further details. Taken from [43]

$$E_m^{\mathrm{KS}}[\{\phi_i\}] = T_\mathrm{s}\left[\{\phi_i\}\right] + \int d\mathbf{r}\ V_{\mathrm{ext}}(\mathbf{r})n(\mathbf{r}) + \frac{1}{2}\int d\mathbf{r}\ V_\mathrm{H}(\mathbf{r})n(\mathbf{r}) + E_{\mathrm{xc}}[n_m^\alpha, n_m^\beta] \tag{44}$$

$$E_t^{\mathrm{KS}}[\{\phi_i\}] = T_\mathrm{s}\left[\{\phi_i\}\right] + \int d\mathbf{r}\ V_{\mathrm{ext}}(\mathbf{r})n(\mathbf{r}) + \frac{1}{2}\int d\mathbf{r}\ V_\mathrm{H}(\mathbf{r})n(\mathbf{r}) + E_{\mathrm{xc}}[n_t^\alpha, n_t^\beta] \tag{45}$$

are expressed in terms of (restricted) Kohn–Sham spin-density functionals constructed from the set $\{\phi_i\}$. The associated S_1 wavefunction is given by

$$|\,s_1\left[\{\phi_i\}\right]\rangle = \frac{1}{\sqrt{2}}\left\{|\,m_1\left[\{\phi_i\}\right]\rangle + |\,m_2\left[\{\phi_i\}\right]\rangle\right\}\ , \tag{46}$$

whereas the antisymmetric linear combination yields

$$|\,t_3\left[\{\phi_i\}\right]\rangle = \frac{1}{\sqrt{2}}\left\{|\,m_1\left[\{\phi_i\}\right]\rangle - |\,m_2\left[\{\phi_i\}\right]\rangle\right\} \tag{47}$$

a third triplet wavefunction. Importantly, the "microstates" m_1 and m_2 are both constructed from the same set $\{\phi_i\}$ of $n+1$ spin-*restricted* orbitals. Using this particular set of orbitals the total density

$$n(\mathbf{r}) = n_m^\alpha(\mathbf{r}) + n_m^\beta(\mathbf{r}) = n_t^\alpha(\mathbf{r}) + n_t^\beta(\mathbf{r}) \tag{48}$$

is of course identical for both the m and t determinants whereas their spin densities clearly differ, see Fig. 2. Thus, the decisive difference between the

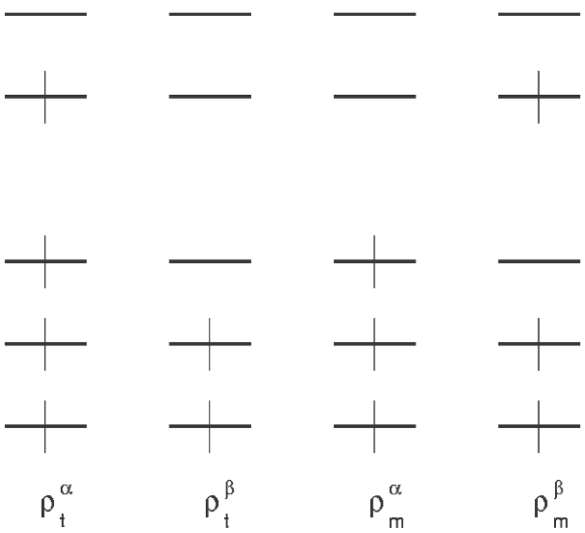


Fig. 2. Four patterns of spin densities n_t^α, n_t^β, n_m^α, and n_m^β corresponding to the two spin-restricted determinants $|t\rangle$ and $|m\rangle$ sketched in Fig. 1, see text for further details. Taken from [43]

m and t functionals (44) and (45), respectively, comes exclusively from the exchange-correlation functional $E_{\rm xc}$, whereas kinetic, external and Hartree energy are identical by construction. Note that this basic philosophy can be generalized to other spin-states by adapting suitably the microstates and the corresponding coefficients in (43) and (46).

Having defined a density functional for the first excited singlet state the corresponding Kohn–Sham equations are obtained by varying the ROKS functional (43)

$$\frac{\delta}{\delta\phi_k^\star}\left\{2E_m^{\rm KS}\left[\{\phi_i\}\right]-E_t^{\rm KS}\left[\{\phi_i\}\right]-\sum_{i,j=1}^{n+1}\Lambda_{ij}(\langle\phi_i\mid\phi_j\rangle-\delta_{ij})\right\}=0 \qquad (49)$$

subject to the orthonormality constraint. Following this procedure the equation for the doubly occupied orbitals $i=1,\ldots,n-1$ reads

$$\begin{aligned}&\Big\{-\frac{1}{2}\nabla^2+V_{\rm H}(\mathbf{r})+V_{\rm ext}(\mathbf{r})\\&+V_{\rm xc}^\alpha[n_m^\alpha(\mathbf{r}),n_m^\beta(\mathbf{r})]+V_{\rm xc}^\beta[n_m^\alpha(\mathbf{r}),n_m^\beta(\mathbf{r})]\\&-\frac{1}{2}V_{\rm xc}^\alpha[n_t^\alpha(\mathbf{r}),n_t^\beta(\mathbf{r})]-\frac{1}{2}V_{\rm xc}^\beta[n_t^\alpha(\mathbf{r}),n_t^\beta(\mathbf{r})]\Big\}\phi_i(\mathbf{r})=\sum_{j=1}^{n+1}\Lambda_{ij}\phi_j(\mathbf{r}) \qquad (50)\end{aligned}$$

whereas

$$\Big\{ \frac{1}{2}\Big[-\frac{1}{2}\nabla^2 + V_{\rm H}(\mathbf{r}) + V_{\rm ext}(\mathbf{r})\Big] + V^{\alpha}_{\rm xc}[n^{\alpha}_m(\mathbf{r}), n^{\beta}_m(\mathbf{r})] - \frac{1}{2}V^{\alpha}_{\rm xc}[n^{\alpha}_t(\mathbf{r}), n^{\beta}_t(\mathbf{r})]\Big\}\phi_{\rm a}(\mathbf{r}) = \sum_{j=1}^{n+1} \Lambda_{{\rm a}j}\phi_j(\mathbf{r}) \ , \quad (51)$$

and

$$\Big\{ \frac{1}{2}\Big[-\frac{1}{2}\nabla^2 + V_{\rm H}(\mathbf{r}) + V_{\rm ext}(\mathbf{r})\Big] + V^{\beta}_{\rm xc}[n^{\alpha}_m(\mathbf{r}), n^{\beta}_m(\mathbf{r})] - \frac{1}{2}V^{\alpha}_{\rm xc}[n^{\alpha}_t(\mathbf{r}), n^{\beta}_t(\mathbf{r})]\Big\}\phi_{\rm b}(\mathbf{r}) = \sum_{j=1}^{n+1} \Lambda_{{\rm b}j}\phi_j(\mathbf{r}) \ . \quad (52)$$

are two *different* equations for the two singly-occupied open-shell orbitals $\phi_{\rm a}$ and $\phi_{\rm b}$, respectively, see Fig. 1. Note that these Kohn–Sham–like equations

$$H^{\rm ROKS}_{{\rm e},i}\phi_i(\mathbf{r}) = \sum_{j=1}^{n+1} \Lambda_{ij}\phi_j(\mathbf{r}) \quad (53)$$

feature a Kohn–Sham–like Hamiltonian as defined in (50)–(52), which is however orbital-dependent, hence the orbital index i, by virtue of its exchange-correlation potential $V^{\alpha}_{\rm xc}[n^{\alpha}_m, n^{\beta}_m] = \delta E_{\rm xc}[n^{\alpha}_m, n^{\beta}_m]/\delta n^{\alpha}_m$, analogous definitions hold for the β and t cases. Solving these equations self-consistently yields a set of relaxed orbitals, including the two singly occupied ones.

The set of equations (50)–(52) could be solved by diagonalization of the corresponding ROKS Hamiltonian or alternatively by direct minimization of the associated total energy functional. The algorithm proposed in [47], which allows to properly and efficiently minimize such orbital-dependent functionals including the orthonormality constraints, was implemented in the `CPMD` package [56, 83]. Based on this minimization technique Born–Oppenheimer molecular dynamics simulations can be performed in the first excited singlet state.

As was well-known the original approach [43] causes problems in situations where the excited state has the same spatial symmetry as the corresponding ground state, such as encountered in $\pi - \pi^{\star}$ transitions of conjugated systems. The angle between the two singly occupied orbitals can be rotated arbitrarily if it is not fixed by symmetry constraints, thus changing the S_1 energy as a function of orbital rotation [92]; viewed from ensemble density functional theory these collapse phenomena have been traced back to the choice of coefficients for the symmetry adapted wavefunction, (41). In order to largely cure this problem, a generalization of the original approach was devised [51], checked [51] for various $\pi - \pi^{\star}$ transitions, and used [90] in an *ab initio* molecular dynamics study of photoinduced isomerizations of 1,3-butadiene and cyclohexadiene. In order to impose explicitly the generalized Brillouin theorem and thus to prevent the collapse to a lower energy within the same symmetry class an additional constraint is introduced [51]

$$H_{\mathrm{e},i}^{\mathrm{ROKS}}\phi_i(\mathbf{r}) = \sum_{j=1}^{n+1} \left\{ \Lambda_{ij}\phi_j(\mathbf{r}) + \Theta_{ij}\left(\Lambda_{ji}^{\star} - \Lambda_{ij}\right) \right\} \tag{54}$$

via another set of real Lagrange multipliers $\{\Theta_{ij}\}$ as inspired from analogous techniques within open-shell Hartree–Fock theory, see [51] for details and references. It is found that the specific choice $\{\Theta_{ij} = 1/2\}$ yields the Goedecker–Umrigar algorithm [47] used in the original ROKS implementation [43]. This works well if the off-diagonal multipliers vanish for symmetry reasons, such as for $n - \pi^{\star}$ transitions [43]. Otherwise unphysical solutions may exist that are lower in energy but do not fulfill the two variational conditions independently. These are characterized by a rotation between the two open shell orbitals ϕ_{a} and ϕ_{b}, leading to singly occupied orbitals that do not obey the molecular symmetry. In addition, the energy collapses to that of the triplet state. To avoid such spurious solutions, it is sufficient to choose $\Theta_{i\mathrm{b}}$ differently from $\Theta_{j\mathrm{a}}$. It is found that the choice $\Theta_{i\mathrm{b}} = -1/2$ and $\Theta_{j\mathrm{a}} = +1/2$ leads to good convergence [51] comparable to that of the Goedecker–Umrigar algorithm [47]. This generalization allows to treat $\pi - \pi^{\star}$ transitions reasonably well as checked for homologous series with conjugated double bonds using polyenes, cyanines and protonated imines [51].

An alternative but closely related formulation of open-shell singlets was provided in [49,50], implemented in the `CPMD` package [56,83] and compared to ROKS [92]. This so-called restricted open-shell singlet (ROSS) method [49,50] is defined by the energy functional

$$\begin{aligned} E_{S_1}^{\mathrm{ROSS}}\left[\{\phi_i\}\right] = T_{\mathrm{s}}\left[\{\phi_i\}\right] + \int d\mathbf{r}\ V_{\mathrm{ext}}(\mathbf{r})n(\mathbf{r}) + \frac{1}{2}\int d\mathbf{r}\ V_{\mathrm{H}}(\mathbf{r})n(\mathbf{r}) \\ + E_{\mathrm{x}}[n_t^{\alpha}, n_t^{\beta}] + 2K[\phi_{\mathrm{a}}, \phi_{\mathrm{b}}] + E_{\mathrm{c}}[n_m^{\alpha}, n_m^{\beta}]\ , \end{aligned} \tag{55}$$

which is minimized subject to the usual orthogonality constraint, see (49). Thus, in both ROSS and ROKS the orbitals are obtained variationally and the efficient computation of gradients with respect to nuclear degrees of freedom, as needed for both optimization and molecular dynamics, is obtained straightforwardly. The term $K = \langle\phi_{\mathrm{a}}\phi_{\mathrm{b}}|\phi_{\mathrm{a}}\phi_{\mathrm{b}}\rangle$ is the usual exchange energy due to the electrons in the two singly occupied orbitals ϕ_{a} and ϕ_{b} from Fig. 1. Within Hartree–Fock theory, $2K$ is the (non-relaxed) singlet-triplet splitting, thus adding $2K$ explicitly to the lower-lying triplet energy should yield another approximation to the singlet state based on an "exact exchange" ansatz. Correlation is taken care of by adding the correlation functional E_{c} evaluated with the spin densities of the mixed state. The ROSS method has been implemented in the `CPMD` package as well and compared to ROKS data in [92]; the single exchange integral is calculated in Fourier space whereas the overlap density is evaluated on the real-space grid.

Much experience concerning the strengths and limitations of the ROKS approach [43] to first excited singlet states accumulated since the first benchmarks were published in 1998. There are certainly notorious cases where

ROKS/ROSS fails, in particular concerning excited state-ordering and when close-lying excitations are present [92]; it is noted in passing that very accurate excited-states forces are available [58] at the expense of increasing the computuational effort such that *ab initio* molecular dynamics is out of reach. However, when ROKS/ROSS is applicable, the typical quality of excited state *structures* seems to be quite similar to corresponding ground state density functional theory. Thus, ROKS/ROSS structures compare in general very favorably to quantum-chemical benchmark data [32, 43, 51, 92], and, as shown later, to optimized structures obtained from time-dependent density functional theory [92]. Clearly, there is the known systematic underestimation of the ROKS/ROSS excitation energies in the range of about 1 eV, which appears to be an essentially constant shift of the S_0/S_1 energy gap as demonstrated for conjugated π-systems [51]; ROSS excitation energies seem to be in slightly better agreement with analogous time-dependent density functional data. It is noted in passing that the qualitative failure of ROKS highlighted in [114] for a specific case, formaldimine $H_2C{=}NH$, could be traced back to quite unusual benchmark parameters used, see [32, 41] and additional unpublished data [31]. Still, it must be stressed that *ab initio* molecular dynamics involving excited states requires in general very careful and extensive checking of the performance of the electronic structure treatment for every specfic problem of interest!

Two Coupled Excited States: Explicit Nonadiabatic Dynamics

A wealth of processes involve neither only a single nor very many electronic states, but rather only "a few but important" states, typical examples being the vast classes of photo(bio)physical processes and photo(bio)chemical reactions. The development of nonadiabatic *ab initio* molecular dynamics methods is further motivated by organic photochemistry [61] and the recent experimental advances in ultrafast photochemistry [33, 134]. Many photochemical processes take place on a subpicosecond time scale driven by conical intersections or avoided crossings of different potential energy surfaces causing the Born–Oppenheimer approximation to break down [61,64,89,108,133]. Thus, *ab initio* molecular dynamics methods that go "beyond the Born–Oppenheimer approximation" by including explicitly nonadiabatic couplings between at least two states in a fully dynamical sense are required, see [28] for a review devoted to these topics.

To this end it is referred to the equations of motion of Ehrenfest molecular dynamics

$$M_I \ddot{\mathbf{R}}_I(t) = -\nabla_I \sum_k |c_k(t)|^2 E_k \tag{56}$$

$$= -\sum_k |c_k(t)|^2 \nabla_I E_k - \sum_{k,l} c_k^\star c_l \left(E_k - E_l\right) \mathbf{d}_I^{kl} \tag{57}$$

$$i\hbar \dot{c}_k(t) = c_k(t) E_k - i\hbar \sum_l c_l(t) D^{kl} \tag{58}$$

with the nonadiabatic coupling elements

$$D^{kl} = \int d\mathbf{r}\, \Psi_k^\star \frac{\partial}{\partial t} \Psi_l = \sum_I \dot{\mathbf{R}}_I \int d\mathbf{r}\, \Psi_k^\star \nabla_I \Psi_l = \sum_I \dot{\mathbf{R}}_I \mathbf{d}_I^{kl} \ , \tag{59}$$

where the adiabatic basis, (14), has been used in order to expand the electronic wavefunction

$$\Psi(\{\mathbf{r}_i\}, \{\mathbf{R}_I\}; t) = \sum_{k=0}^{\infty} c_k(t) \Psi_k(\{\mathbf{r}_i\}; \{\mathbf{R}_I\}) \ ; \tag{60}$$

see [83] for a detailed derivation of these equations and for background material. These equations of motion propagate the expansion coefficients $c_k(t)$ whose square modulus, $|c_k(t)|^2$, can be interpreted as state occupation numbers, i.e. as probabilities of finding the system in the adiabatic state Ψ_k at time t.

In most case, however, these equations are immediately simplified by completely neglecting the nonadiabatic coupling elements and vectors, D^{kl} and $\mathbf{d}_I^{kl}$, as defined in (59) and by restricting the propagation to the ground state Ψ_0. When is this adiabatic approximation, i.e. $|c_0(t)|^2 = 1 \,\forall\, t$, justified? The quantitative answer is embodied in the so-called Massey parameter [26, 84]

$$\xi = \frac{\tau_\mathrm{n}}{\tau_\mathrm{e}} = \frac{\Delta E\, L}{\hbar \dot{R}} \ , \tag{61}$$

which is the ratio of the passage time of the nuclei, τ_n, and the characteristic timescale of electronic motion, τ_e, approximated in terms of the energy difference ΔE between two electronic states, a characteristic length L, and the velocity of the nuclei $\dot{R}$. Thus, nonadiabatic effects are negligible only in the limit $\xi \gg 1$, i.e. for large energy spacings and small velocities, so that propagation of only the ground-state adiabatic wavefunction, Ψ_0, is justified.

As discussed in Sect. 2.3, efficient *adiabatic* Car–Parrinello simulations have become possible also in the first excited state using the restricted open-shell Kohn–Sham (ROKS) approach [43]. Thus, *nonadiabatic* extensions of *ab initio* molecular dynamics can be devised by coupling the ROKS S_1 excited state to the Kohn–Sham ground state S_0 as introduced in [27], see [28] for a review of the entire field of "on-the-fly" approaches to excited state dynamics. As shown in Sect. 2.3, the S_1 restricted open-shell singlet wavefunction is constructed by linearly combining the mixed determinants, m_1 and m_2 from Fig. 1,

$$\Psi_1 = \frac{1}{\sqrt{2}} \Big\{ |\phi_1^{(1)} \bar{\phi}_1^{(1)} \phi_2^{(1)} \bar{\phi}_2^{(1)} \cdots \phi_n^{(1)} \bar{\phi}_{n+1}^{(1)}\rangle + |\phi_1^{(1)} \bar{\phi}_1^{(1)} \phi_2^{(1)} \bar{\phi}_2^{(1)} \cdots \bar{\phi}_n^{(1)} \phi_{n+1}^{(1)}\rangle \Big\} \tag{62}$$

where the "ket" notation signifies Slater determinants made up of Kohn–Sham orbitals, $\phi_i^{(1)}$ (spin up) and $\bar{\phi}_i^{(1)}$ (spin down); the total number of electrons is $2n$ as in Sect. 2.3. Within the ROKS approach, see (49), a *single* set of orbitals $\{\phi_i^{(1)}\}$ is determined that minimizes the energy functional $E_{S_1}[\{\phi_i^{(1)}\}]$ for the first excited state. Due to this optimization the entire set of orbitals $\{\phi_i^{(1)}\}$ will, in general, differ from the set of orbitals $\{\phi_i^{(0)}\}$ that defines the ground state wavefunction, Ψ_0,

$$\Psi_0 = |\phi_1^{(0)} \bar{\phi}_1^{(0)} \phi_2^{(0)} \bar{\phi}_2^{(0)} \cdots \phi_l^{(0)} \bar{\phi}_l^{(0)}\rangle \ , \tag{63}$$

by minimizing the Kohn–Sham energy functional and, as a consequence, the two state functions, Ψ_0 and Ψ_1, are nonorthogonal

$$S_{01} = S_{10} = S \ , \quad S_{kk} = 1 \ . \tag{64}$$

giving rise to the overlap matrix $\{S_{kl}\}$. Using these two adiabatic states in a truncated expansion, see (60), of the electronic wavefunction

$$\Psi = a_0(t)\Psi_0 \exp\left[-\frac{i}{\hbar}\int E_0 dt\right] + a_1(t)\Psi_1 \exp\left[-\frac{i}{\hbar}\int E_1 dt\right] \tag{65}$$

and inserting it into the Ehrenfest equations yields

$$\sum_l a_l p_l (H_{kl} - E_l S_{kl}) = i\hbar \left\{ \sum_l \dot{a}_l p_l S_{kl} + \sum_l a_l p_l D^{kl} \right\} \ , \tag{66}$$

where the Hamiltonian matrix elements are given by

$$H_{kk} = \langle \Psi_k | H_{\mathrm{e}}^{(\mathrm{RO})\mathrm{KS}} | \Psi_k \rangle = E_k \tag{67}$$

$$H_{01} = H_{10} = E_0 S \ ; \tag{68}$$

the phase factors

$$p_l = \exp\left[-\frac{i}{\hbar}\int E_l dt\right] \tag{69}$$

have been introduced for convenience [27, 28] and the nonadiabatic couplings D^{kl} are defined in (59). Solving (66) for the time-dependent expansion coefficients $\dot{a}_0$ and $\dot{a}_1$ leads to the set of coupled differential equations

$$\dot{a}_0 = \frac{1}{S^2 - 1}\left[i a_1 \frac{p_1}{p_0} S(E_0 - E_1) + a_1 D^{01} \frac{p_1}{p_0} - a_0 D^{10} S \right] \tag{70}$$

$$\dot{a}_1 = \frac{1}{S^2 - 1}\left[a_0 D^{10} \frac{p_0}{p_1} - a_1 D^{01} S - i a_1 S^2 (E_0 - E_1) \right] \ , \tag{71}$$

which can be integrated numerically using standard schemes. It is computationally attractive to work with the nonadiabatic coupling *elements*, D^{kl}, instead of the nonadiabatic coupling *vectors*, $\mathbf{d}_I^{lk}$, as defined in (59). The orbital velocities required in the former case, $\partial_t|\Psi_l\rangle$, are explicitly available within the Car–Parrinello method, whereas the nuclear gradient required for the latter, $\nabla_I|\Psi_l\rangle$, must be evaluated in addition. An iterative Born–Oppenheimer–style implementation of nonadiabatic *ab initio* molecular dynamics is also straightforwardly possible within the outlined framework. This can be done either using a simple finite difference approach to the coupling elements or by a scheme to directly compute the nonadiabatic coupling vectors as developed in [8].

If both electronic state functions $\{\Psi_0, \Psi_1\}$ were eigenfunctions of the Kohn–Sham Hamiltonian, $|a_0|^2$ and $|a_1|^2$ would be their respective occupation numbers. A look at the normalization integral of the electronic wavefunction, Ψ,

$$\langle\Psi|\Psi\rangle = |a_0|^2 + |a_1|^2 + 2S\,\Re\left(a_0^\star a_1 \frac{p_1}{p_0}\right) = 1 \tag{72}$$

shows that the definition of state populations in this basis is ambiguous due to the nonvanishing overlap $S > 0$. Thus, an auxiliary, orthonormal basis $\{\Psi_0', \Psi_1'\}$ is introduced [27] to re-expand Ψ

$$\Psi = c_0\Psi_0' + c_1\Psi_1' \;, \tag{73}$$

which allows to obtain proper occupations numbers

$$|c_0|^2 = |a_0|^2 + S^2|a_1|^2 + 2S\,\Re\left(a_0^\star a_1 \frac{p_1}{p_0}\right) \tag{74}$$

$$|c_1|^2 = (1 - S^2)|a_1|^2 \tag{75}$$

satisfying the sum rule $|c_0|^2 + |c_1|^2 = 1$ such that the density matrix $\rho_{kl}(t) = c_k^\star(t)c_l(t)$ can be defined, see [28] for details. It is noted that these reorthogonalization problems, encountered when using ROKS including orbital relaxation, do not occur when all states $\{\Psi_k\}$, i.e. also the excited states, are eigenfunctions of the *same* Kohn–Sham Hamiltonian. This is the case when no orbital relaxation is performed within ROKS or when the "broken-symmetry determinant" [57, 91, 100] is used, see Sect. 2.3, which corresponds to the full "particle-hole" excitation without any orbital relaxation and thus to the m states in Fig. 1. A somewhat more sophisticated approach [8] would be to make use of "Slater's transition-state" determinant, which can be viewed as a simple approximation to ensemble density functional theory [34] by introducing fractional occupation numbers of exactly 1/2 for the two involved particle-hole orbitals, i.e. ϕ_a and ϕ_b in Fig. 1.

At this stage, at least two different avenues to obtain a *nuclear* trajectory coupled to the time-evolution of the electrons accessing more than one electronic state can be followed. Clearly, nonadiabatic Ehrenfest molecular

dynamics according to (57)–(58) is a first option. However, several severe deficiencies of this approach are well documented, see for example [26,28,129,130] and references cited therein. At the root of many of these problems is the mean-field character of Ehrenfest molecular dynamics [26]: the force acting on the nuclei is obtained by *averaging* over all adiabatic states $\{\Psi_k\}$ used to expand the electronic wavefunction Ψ according to their instantaneous occupation numbers $|c_k(t)|^2$, as most directly represented by (56). In particular, a system that was initially prepared in a pure adiabatic state will be in a mixed state when leaving the region of strong nonadiabatic coupling. In general, the pure adiabatic character of the wavefunction cannot be recovered even in the asymptotic regions of configuration space. In cases where the differences in the adiabatic potential energy landscapes are pronounced, it is clear that an average potential will be unable to describe all reaction channels adequately. In particular, if one is interested in a reaction branch whose occupation number is very small, the average path is likely to diverge from the "true trajectory". Furthermore, the total wavefunction may contain significant contributions from adiabatic states that are energetically inaccessible in classical mechanics due to lack of sufficient energy.

A powerful alternative is Tully molecular dynamics [127, 128, 131] in particular when using the ingenious "fewest switches" surface hopping algorithm [128] to account for nonadiabatic effects, see for instance [26,28,129,130] for reviews and [118] for a comparison to other methods and a discussion of its limitations. Fundamental to Tully molecular dynamics is the idea that the system is always in a particular pure state. Thus, at any moment in time, the system is propagated on some adiabatic state Ψ_k, which is selected according to its state population $|c_k|^2$ by a statistical hopping criterion. Changing adiabatic state occupations can thus result in nonadiabatic transitions between different adiabatic potential energy surfaces. At variance with the Ehrenfest approach where "the best path" is generated, an ensemble of trajectories must be generated and analyzed in order to extract observables. This approach to nonadiabatic *ab initio* molecular dynamics [27, 28] is obtained by integrating simultaneously the following equations

$$M_I \ddot{\mathbf{R}}_I(t) = -\nabla_I E_k \tag{76}$$

$$i\hbar \dot{c}_k(t) = c_k(t) E_k - i\hbar \sum_l c_l(t) D^{kl} \tag{77}$$

together with solving the ground-state Kohn–Sham equations and (50)–(52) to obtain the energies and wavefunctions, where the nonadiabatic coupling elements, (59), are defined as $D^{kl} = \int d\mathbf{r}\, \Psi_k^\star \partial_t \Psi_l$. Importantly, the forces acting on the nuclei, (76), are obtained from the gradient of the energy E_k of a *single* adiabatic state Ψ_k, at variance with the Ehrenfest approach where *all* states are averaged according to (56). The particular state populated is stochastically selected in each step of the molecular dynamics propagation using the fewest switches hopping criterion [128]

$$P_k(t) = -\frac{\dot{\rho}_{kk}(t)\delta t}{\rho_{kk}(t)} \tag{78}$$

to hop from state Ψ_k to any other state in the time interval $[t, t+\delta t]$; here $\rho_{kl}(t) = c_k^\star(t)c_l(t)$ is the density matrix obtained in the auxiliary basis (74)–(75) and δt is the molecular dynamics time step. More details of the method can be found in [26–28].

The computational effort per timestep increases linearly with the number of adiabatic states involved as compared to adiabatic Car–Parrinello or Born–Oppenheimer propagation. However, the timestep must be as small as the one typically used in Ehrenfest dynamics [83], which is of the order of $\delta t \sim 10^{-16}$ s in order to resolve the very fast dynamics of the electrons. In a straightforward implementation this would increase the cost of nonadiabatic simulations per picosecond by another factor of 10 or 100 compared to the two corresponding adiabatic schemes; interpolation between the time steps is used to cut down this overhead. Finally, it is required within Tully molecular dynamics to generate as many as possible independent stochastic trajectories in order to sample averages. Despite the computationally demanding character of nonadiabatic *ab initio* molecular dynamics several case studies [29,30,66,67] clearly demonstrate both its power and application range in the realm of photophysics and photochemistry including the condensed phase.

3 Summary and Outlook

The original Car–Parrinello *ab initio* molecular dynamics approach [18], which relies on *adiabatic* propagation of *classical* nuclei on the electronic *ground-state* potential energy surface, has proven to be an extremly powerful atomistic simulation technique not only in physics but also in chemistry and more recently in biology [83]. Since its invention twenty year ago, in 1985, a host of extensions and generalizations have been worked out [83], most of which are now implemented in the `CPMD` software package [56,83]. Here, the focus was to go beyond both the classical nuclei approximation and the Born–Oppenheimer approximation. Quantization of the nuclear degrees of freedom is achieved by using Feynman's path integral representation of quantum mechanics for the nuclei, whereas nonadiabatic effects are included based on a restricted open-shell Kohn–Sham density functional in conjunction with Tully's surface hopping algorithm. Both techniques carry over the spirit of the Car–Parrinello approach to new shores. Thus, large and chemically complex systems including in particular condensed matter are efficiently accessible akin to the original approach, including however quantum effects on the nuclei and electronically nonadiabatic effects. These advances allow to answer questions about the importance of proton tunneling in hydrogen-bonded systems, the effects of zero-point motion on shallow potential energy landscapes of molecular complexes and clusters, nonradiative decay mechanisms via conical intersections in nucleobases

to name only a few. Thus, using *ab initio* path integrals [72,74,80,122], molecular complexes [62,72,73,75,76,81,85,123,125], clusters [110,111,113,116,117], solids [2–7], and liquids [22,79,82,126] have been investigated in great detail. The more recent nonadiabatic technique [27,28] is still in the process of unfolding its potential, but significant applications have been published in the realm of photobiophysical and photobiochemical processes in the gas phase, in microsolvated clusters, and in aqueous solution [27,29,30,66,67].

Acknowledgments

It gives me great pleasure, to thank the many colleagues and co-workers, with whom I have been honoured to collaborate with in the covered fields. I would especially like to single out Nikos Doltsinis, Jürg Hutter, Michele Parrinello, and Mark Tuckerman. Last but not least, this research would not have been possible without the financial help of funding organisations (RUB, DFG, VW-Stiftung, FCI, AVH, DAAD) and German Supercomputer Resources (NIC, HLRS, SSCK, HLRB, RV-NRW).

References

1. R. N. Barnett, H.-P. Cheng, H. Hakkinen, and U. Landman (1995) *J. Phys. Chem.* **99**, p. 7731
2. M. Benoit, D. Marx, and M. Parrinello (1998) *Comp. Mat. Sci.* **10**, p. 88
3. M. Benoit, D. Marx, and M. Parrinello (1998) *Nature* **392**, p. 258; see also *J. Teixeira, Nature* (1998) **392**, p. 232
4. M. Benoit, D. Marx, and M. Parrinello (1999) *Solid State Ionics* **125**, p. 23
5. M. Benoit and D. Marx (2005) *Chem. Phys. Chem.* **6**, p. 1738
6. S. Biermann, D. Hohl, and D. Marx (1998) *J. Low Temp. Phys.* **110**, p. 97
7. S. Biermann, D. Hohl, and D. Marx (1998) *Solid State Commun.* **108**, p. 337
8. S. R. Billeter and A. Curioni (2005) *J. Chem. Phys.* **122**, p. 034105
9. P. E. Blöchl and M. Parrinello (1992) *Phys. Rev. B* **45**, p. 9413
10. F. A. Bornemann and C. Schütte (1998) *Numer. Math.* **78**, p. 359
11. D. J. E. Callaway and A. Rahman (1982) *Phys. Rev. Lett.* **49**, p. 613
12. J. Cao and B. J. Berne (1993) *J. Chem. Phys.* **99**, p. 2902
13. J. Cao and G. A. Voth (1993) *J. Chem. Phys.* **99**, p. 10070
14. J. Cao and G. A. Voth (1994) *J. Chem. Phys.* **100**, p. 5106
15. J. Cao and G. A. Voth (1994) *J. Chem. Phys.* **101**, p. 6157
16. J. Cao and G. A. Voth (1994) *J. Chem. Phys.* **101**, p. 6168
17. J. Cao and G. J. Martyna (1996) *J. Chem. Phys.* **104**, p. 2028
18. R. Car and M. Parrinello (1985) Unified Approach for Molecular Dynamics and Density-Functional Theory. *Phys. Rev. Lett.* **55**, pp. 2471–2474; this article belongs to the most quoted "Physical Review Letters, Top 10", see `http://www.aps.org/apsnews/topten.html`, and led in 1996 to the definition of its own PACS number, 71.15.Pd.

19. D. M. Ceperley (1995) *Rev. Mod. Phys.* **67**, p. 279; for errata and updates see http://archive.ncsa.uiuc.edu/Science/CMP/papers/cep95a/cep95a.html
20. Ch. Chakravarty (1997) *Int. Rev. Phys. Chem.* **16**, p. 421
21. D. Chandler (1991) In *Liquids, Freezing, and Glass Transition*, eds. J. P. Hansen, D. Levesque, and J. Zinn-Justin Elsevier, Amsterdam
22. B. Chen, I. Ivanov, M. L. Klein, and M. Parrinello (2003) *Phys. Rev. Lett.* **91**, p. 215503
23. H.-P. Cheng, R. N. Barnett, and U. Landman (1995) *Chem. Phys. Lett.* **237**, p. 161
24. R. D. Coalson (1986) *J. Chem. Phys.* **85**, p. 926
25. C. Daul (1994) *Int. J. Quantum Chem.* **52**, p. 867
26. N. L. Doltsinis, Nonadiabatic Dynamics: Mean-Field and Surface Hopping, in [53]
27. N. L. Doltsinis and D. Marx (2002) *Phys. Rev. Lett.* **88**, p. 166402
28. N. L. Doltsinis and D. Marx (2002) *J. Theor. Comput. Chem.* **1**, p. 319; see http://www.theochem.rub.de/go/cprev.html
29. N. L. Doltsinis (2004) *Mol. Phys.* **102**, p. 499
30. N. L. Doltsinis (2004) *Faraday Discuss.* **127**, p. 231
31. N. L. Doltsinis and K. Fink (2004) Unpublished calculations
32. N. L. Doltsinis and K. Fink (2005) *J. Chem. Phys.* **122**, p. 087101; Comment to [114], for Reply see [41]
33. W. Domcke and G. Stock (1997) *Adv. Chem. Phys.* **100**, p. 1
34. R. M. Dreizler and E. K. U. Gross (1990) Density-Functional Theory. Springer, Berlin Heidelberg
35. R. P. Feynman and A. R. Hibbs (1965) Quantum Mechanics and Path Integrals. McGraw-Hill, New York
36. R. P. Feynman (1972) Statistical Mechanics. Addison-Wesley, Redwood City
37. M. Filatov and S. Shaik (1998) *Chem. Phys. Lett.* **288**, p. 689
38. M. Filatov and S. Shaik (1999) *J. Chem. Phys.* **110**, p. 116
39. M. Filatov and S. Shaik (1999) *Chem. Phys. Lett.* **304**, p. 429
40. M. Filatov and S. Shaik (2000) *Chem. Phys. Lett.* **332**, p. 409
41. C. Filippi and F. Buda (2005) *J. Chem. Phys.* **122**, p. 087102; Reply to Comment [32]
42. J. Ford (1992) *Phys. Rep.* **213**, p. 271
43. I. Frank, J. Hutter, D. Marx, and M. Parrinello (1998) *J. Chem. Phys.* **108**, p. 4060
44. G. Galli and M. Parrinello (1991) In *Computer Simulations in Materials Science*, p. 282, eds. M. Meyer and V. Pontikis Kluwer, Dordrecht
45. R. Gaudoin and K. Burke (2004) *Phys. Rev. Lett.* **93**, p. 173001; Publisher's Note (2005) *Phys. Rev. Lett.* **94**, p. 029901
46. M. J. Gillan (1990) In *Computer Modelling of Fluids, Polymers, and Solids*, eds. C. R. A. Catlow, S. C. Parker, and M. P. Allen Kluwer, Dordrecht
47. S. Goedecker and C. J. Umrigar (1997) *Phys. Rev. A* **55**, p. 1765
48. A. Görling (2000) *Phys. Rev. Lett.* **85**, p. 4229
49. J. Gräfenstein, E. Kraka, and D. Cremer (1998) *Chem. Phys. Lett.* **288**, p. 593
50. J. Gräfenstein and D. Cremer (2000) *Phys. Chem. Chem. Phys.* **2**, p. 2091
51. S. Grimm, C. Nonnenberg, and I. Frank (2003) *J. Chem. Phys.* **119**, p. 11574
52. E. K. U. Gross and W. Kohn (1990) *Adv. Quant. Chem.* **21**, p. 255

53. J. Grotendorst, D. Marx, and A. Muramatsu, eds., Quantum Simulations of Complex Many-Body Systems: From Theory to Algorithms. (John von Neumann Institute for Computing, Forschungszentrum Jülich 2002); for Electronic Version and Audio-Visual Lecture Notes see: `http://www.theochem.rub.de/go/cprev.html`
54. R. W. Hall and B. J. Berne (1984) *J. Chem. Phys.* **81**, p. 3641
55. A. Hellmann, B. Razaznejad, and B. I. Lundqvist (2004) *J. Chem. Phys.* **120**, p. 4593
56. J. Hutter et al., `CPMD`: *Car-Parrinello Molecular Dynamics: An Ab Initio Electronic Structure and Molecular Dynamics Program*; IBM Zurich Research Laboratory (1990–2005) and Max-Planck-Institut für Festkörperforschung (1997–2001); for downloads see: `http://www.cpmd.org/`
57. F. Illas, I. de P. R. Moreira, C. de Graaf, and V. Barone (2000) *Theor. Chem. Acc.* **104**, p. 265
58. S. Ismail Beigi and S. G. Louie (2003) *Phys. Rev. Lett.* **90**, p. 076401
59. H. Kitamura, S. Tsuneyuki, T. Ogitsu, and T. Miyake (2000) *Nature* **404**, p. 259
60. H. Kleinert (2004) Path Integrals in Quantum Mechanics, Statistics, Polymer Physics, and Financial Markets. World Scientific, Singapore
61. M. Klessinger and J. Michl (1995) *Excited States and Photochemistry of Organic Molecules.* VCH, New York
62. L. Knoll, Z. Vager, and D. Marx (2003) *Phys. Rev. A* **67**, p. 022506
63. W. Kołos (1970) *Adv. Quant. Chem.* **5**, p. 99
64. H. Köppel, W. Domcke, and L. S. Cederbaum (1984) *Adv. Chem. Phys.* **57**, p. 59
65. K. Laasonen, A. Pasquarello, R. Car, C. Lee, and D. Vanderbilt (1993) *Phys. Rev. B* **47**, pp. 10–142
66. H. Langer and N. L. Doltsinis (2004) *Phys. Chem. Chem. Phys.* **6**, p. 2742
67. H. Langer, N. L. Doltsinis, and D. Marx (2005) *Chem. Phys. Chem.* **6**, p. 1734
68. M. A. L. Marques and E. K. U. Gross (2004) *Annu. Rev. Phys. Chem.* **55**, p. 427
69. G. J. Martyna, M. L. Klein, and M. Tuckerman (1992) *J. Chem. Phys.* **97**, p. 2635
70. G. J. Martyna (1996) *J. Chem. Phys.* **104**, p. 2018
71. G. J. Martyna, A. Hughes, and M. E. Tuckerman (1999) *J. Chem. Phys.* **110**, p. 3275
72. D. Marx and M. Parrinello (1994) *Z. Phys. B* (Rapid Note) **95**, p. 143; a misprinted sign in the Lagrangian is correctly given in (21) of the present review
73. D. Marx and M. Parrinello (1995) *Nature* **375**, p. 216
74. D. Marx and M. Parrinello (1996) *J. Chem. Phys.* **104**, p. 4077
75. D. Marx and M. Parrinello (1996) *Science* **271**, p. 179
76. D. Marx and A. Savin (1997) *Angew. Chem. Int. Ed. Engl.* **36**, p. 2077
77. D. Marx (1998) In *Classical and Quantum Dynamics in Condensed Phase Simulations.* Chap. 15, eds. B. J. Berne, G. Ciccotti, and D. F. Coker World Scientific, Singapore
78. D. Marx and M. H. Müser (1999) *J. Phys. Condens. Matter* **11**, p. R117
79. D. Marx, M. E. Tuckerman, J. Hutter, and M. Parrinello (1999) *Nature* **397**, p. 601; see also J. T. Hynes (1999) *Nature* **397**, p. 565

80. D. Marx, M. E. Tuckerman, and G. J. Martyna (1999) *Comput. Phys. Commun.* **118**, p. 166; the misprinted definition of the fictitious normal mode masses given in (2.51) is corrected in (40) of the present review. However, the correct definition was implemented in the `CPMD` package [56] so that all data reported in the paper are unaffected
81. D. Marx and M. Parrinello (1999) *Science* **286**, p. 1051
82. D. Marx, M. E. Tuckerman, and M. Parrinello (2000) *J. Phys.: Condens. Matter* **12**, p. A153
83. D. Marx and J. Hutter, "Ab Initio Molecular Dynamics: Theory and Implementation," In *Modern Methods and Algorithms of Quantum Chemistry*, pp. 301–449 (first edition, paperback, `ISBN 3-00-005618-1`) or pp. 329–477 (second edition, hardcover, `ISBN 3-00-005834-6`), ed. J. Grotendorst, (John von Neumann Institute for Computing, Forschungszentrum Jülich 2000); see: `http://www.theochem.rub.de/go/cprev.html`
84. H. S. W. Massey (1949) *Rep. Prog. Phys.* **12**, p. 248
85. S. Miura, M. E. Tuckerman, and M. L. Klein (1998) *J. Chem. Phys.* **109**, p. 5290
86. T. Miyake, T. Ogitsu, and S. Tsuneyuki (1998) *Phys. Rev. Lett.* **81**, p. 1873
87. T. Miyake, T. Ogitsu, and S. Tsuneyuki (1999) *Phys. Rev. B* **60**, pp. 14–197
88. Á. Nagy (1998) *Phys. Rev. A* **57**, p. 1672
89. H. Nakamura (2002) *Nonadiabatic Transition: Concepts, Basic Theories and Applications*. World Scientific, Singapore
90. C. Nonnenberg, S. Grimm, and I. Frank (2003) *J. Chem. Phys.* **119**, p. 11585
91. L. Noodleman (1981) *J. Chem. Phys.* **74**, p. 5737
92. M. Odelius, D. Laikov, and J. Hutter (2003) *J. Mol. Struc. (Theochem.)* **630**, p. 163
93. R. G. Parr and W. Yang (1989) *Density-Functional Theory of Atoms and Molecules*. Oxford University Press, Oxford
94. M. Parrinello and A. Rahman (1984) *J. Chem. Phys.* **80**, p. 860
95. M. Parrinello (1997) *Solid State Commun.* **102**, p. 107
96. G. Pastore, E. Smargiassi, and F. Buda (1991) *Phys. Rev. A* **44**, p. 6334
97. G. Pastore (1996) In *Monte Carlo and Molecular Dynamics of Condensed Matter Systems*, Chap. 24, p. 635, eds. K. Binder and G. Ciccotti Italien Physical Society SIF, Bologna
98. M. Pavese, D. R. Berard, and G. A. Voth (1999) *Chem. Phys. Lett.* **300**, p. 93
99. M. C. Payne, M. P. Teter, D. C. Allan, T. A. Arias, and J. D. Joannopoulos (1992) *Rev. Mod. Phys.* **64**, p. 1045
100. Á. J. Pérez Jiménez, J. M. Pérez Jordá, and F. Illas (2004) *J. Chem. Phys.* **120**, p. 18
101. J. Polonyi and H. W. Wild (1983) *Phys. Rev. Lett.* **51**, p. 2257
102. B. De Raedt, M. Sprik, and M. L. Klein (1984) *J. Chem. Phys.* **80**, p. 5719
103. R. Ramírez, T. López Ciudad, and J. C. Noya (1998) *Phys. Rev. Lett.* **81**, p. 3303
104. R. Ramírez and T. López Ciudad (1999) *J. Chem. Phys.* **111**, p. 3339
105. R. Ramírez and T. López Ciudad (1999) *Phys. Rev. Lett.* **83**, p. 4456
106. R. Ramírez and T. López Ciudad, Dynamic Properties via Fixed Centroid Path Integrals, in [53]
107. D. K. Remler and P. A. Madden (1990) *Molec. Phys.* **70**, p. 921
108. M. A. Robb, M. Garavelli, M. Olivucci, and F. Bernardi (2000) *Rev. Comp. Ch.* **15**, p. 87

109. C. C. J. Roothaan (1960) *Rev. Mod. Phys.* **32**, p. 179
110. R. Rousseau and D. Marx (1998) *Phys. Rev. Lett.* **80**, p. 2574
111. R. Rousseau and D. Marx (1999) *J. Chem. Phys.* **111**, p. 5091
112. E. Runge and E. K. U. Gross (1984) *Phys. Rev. Lett.* **52**, p. 997
113. F. Della Sala, R. Rousseau, A. Görling, and D. Marx (2004) *Phys. Rev. Lett.* **92**, p. 183401
114. F. Schautz, F. Buda, and C. Filippi (2004) *J. Chem. Phys.* **121**, p. 5836; see also a Comment [32] and the Reply [41]
115. K. E. Schmidt and D. M. Ceperley (1992) In *The Monte Carlo Method in Condensed Matter Physics*, p. 205, ed. K. Binder Springer-Verlag, Berlin Heidelberg
116. I. Štich, D. Marx, M. Parrinello, and K. Terakura (1997) *Phys. Rev. Lett.* **78**, p. 3669
117. I. Štich, D. Marx, M. Parrinello, and K. Terakura (1997) *J. Chem. Phys.* **107**, p. 9482
118. G. Stock and M. Thoss (2005) *Adv. Chem. Phys.* **131**, p. 243
119. A. C. Stückl, C. A. Daul, and H. U. Güdel (1997) *Int. J. Quantum Chem.* **61**, p. 579
120. A. Szabo and N. S. Ostlund (1989) *Modern Quantum Chemistry – Introduction to Advanced Electronic Structure Theory* (McGraw-Hill Publishing Company, New York
121. M. E. Tuckerman, B. J. Berne, G. J. Martyna, and M. L. Klein (1993) *J. Chem. Phys.* **99**, p. 2796
122. M. E. Tuckerman, D. Marx, M. L. Klein, and M. Parrinello (1996) *J. Chem. Phys.* **104**, p. 5579
123. M. E. Tuckerman, D. Marx, M. L. Klein, and M. Parrinello (1997) *Science* **275**, p. 817
124. M. E. Tuckerman and A. Hughes (1998) In *Classical and Quantum Dynamics in Condensed Phase Simulations*, Chap. 14, p. 311, eds. B. J. Berne, G. Ciccotti, and D. F. Coker World Scientific, Singapore
125. M. E. Tuckerman and D. Marx (2001) *Phys. Rev. Lett.* **86**, p. 4946; see animation of quantum fluctuations and tunneling proton `http://www.theochem.rub.de/go/malon.html`
126. M. E. Tuckerman, D. Marx, and M. Parrinello (2002) *Nature* **417**, p. 925; see also R. Ludwig (2003) *Angew. Chem. Int. Ed.* **42**, p. 258
127. J. C. Tully and R. K. Preston (1971) *J. Chem. Phys.* **55**, p. 562
128. J. C. Tully (1990) *J. Chem. Phys.* **93**, p. 1061
129. J. C. Tully (1998) In *Modern Methods for Multidimensional Dynamics Computations in Chemistry*, ed. D. L. Thompson World Scientific, Singapore
130. J. C. Tully (1998) In *Classical and Quantum Dynamics in Condensed Phase Simulations*. Chap. 21, p. 489, eds. B. J. Berne, G. Ciccotti, and D. F. Coker World Scientific, Singapore
131. P. V. Parandekar and J. C. Tully (2005) *J. Chem. Phys.* **122**, p. 094102
132. G. A. Voth (1996) *Adv. Chem. Phys.* **93**, p. 135
133. D. R. Yarkony (1996) *Rev. Mod. Phys.* **68**, p. 985
134. A. H. Zewail (1994) *Femtochemistry: Ultrafast Dynamics of the Chemical Bond.* World Scientific, Singapore
135. T. Ziegler, A. Rauk, and E. J. Baerends (1977) *Theor. Chim. Acta* **43**, p. 261

Evolutionary Design in Biological Physics and Materials Science

M. Yang[1], J.-M. Park[1,2], and M.W. Deem[1]

[1] Rice University, 6100 Main Street—MS 142, Houston, TX, 77005-1892 USA
mwdeem@rice.edu
[2] Department of Physics, The Catholic University of Korea, Puchon 420–743, Korea
jmanpark@catholic.ac.kr



Michael W. Deem

M. Yang et al.: *Evolutionary Design in Biological Physics and Materials Science*, Lect. Notes Phys. **704**, 541–562 (2006)
DOI 10.1007/3-540-35284-8_20
© Springer-Verlag Berlin Heidelberg 2006

1 Introduction 543

1.1 The Polytopic Vaccination Experiment 544

2 The Statistical Mechanical Model of the Immune Response to Cancer 547

2.1 Generalized NK Model 547
2.2 TCR Selection Dynamics 552
2.3 Tumor-Related Parameters 555

3 Results 555

4 Discussion 555

5 Conclusions 559

References 559

In this chapter we provide a thorough discussion of the theoretical description of the multi-site approach to cancer vaccination. The discussion is somewhat demanding from a biological point of view. References to primary biological publications are given. A general reference on immunology is [1].

What distinguishes cancer from other more treatable diseases is perhaps the random, multi-strain nature of the disease. Here we apply tools from statistical mechanics to model cancer vaccine design. The difficulty of controlling cancer by many of the standard therapies has led to substantial interest in control by the immune system. Escape of cancer from the immune system can be viewed as a percolation transition, with the immune system killing being the parameter controlling whether the cancers cells proliferate. The model we develop suggests that vaccination with the different strains of cancer in different physical regions leads to an improved immune response against each strain. Our approach captures the recognition characteristics between the T cell receptors and tumor, the primary dynamics due to T cell resource competition, and elimination of tumor cells by the selected T cells.

1 Introduction

The refractory nature of cancer to many standard therapies has led to substantial efforts to achieve immune control. Here we propose a mechanism by which the immunodominance hierarchy that allows a growing tumor to escape from immune surveillance may be broken. We focus on mitigating the deleterious effects of immunodominance and on achieving an effective strategy in the face of central and peripheral tolerance. Our approach captures the recognition characteristics between the T cell receptors (TCRs) and tumor, the primary dynamics due to TCR resource competition, and elimination of tumor cells by TCRs. The hypothesis that polytopic vaccination induces independent selection of T cells for each epitope of the vaccine in distinct lymph nodes is consistent with the experimental data. Polytopic administration of a therapeutic cancer vaccine may sculpt a broader immune response and mitigate immunodominance. We suggest that by inducing a T cell response to each cancer-associated epitope in a distinct lymph node, vaccine efficacy is increased and immunodominance is reduced. Whether the cancer-associated epitopes are related or unrelated, polytopic vaccination appears to be a promising therapeutic strategy.

Our immune system protects us against a broad spectrum of possible cancers [2–4]. Work with interferon-γ receptor knockout mice, which exhibited extremely high incidences of spontaneous cancers, suggests a daily combat of cancers by the immune system [5,6]. Several limitations of the cellular immune system have been reported, however, that stem from the cross-reactivity of T cell receptors. One obstacle to a robust immune response is immunodominance, in which dominant cancer-associated epitopes suppress the generation of CTL (cytotoxic T lymphocyte) activity toward other non-dominant

epitopes of the same cancerous cell [7]. Immunodominance may thereby prevent the development of an immune responses to more than the dominant epitope among the multiple tissue-specific and tissue-associated cancer antigens [8]. Immunodominance may also inhibit development of an immune response to the new tissue-specific antigens that develop during the course of tumor progression [9]. Many experimental observations are consistent with the hypothesis that outgrowth of patient's tumors reflects Darwinian selection of tumor cells that have acquired escape mechanisms from immune recognition [5,10–12]. It remains a challenge to fully elucidate the mechanisms holding back tumor-specific immunity [13].

A given cancer typically has several tissue-specific or tissue-related epitopes that are recognized by the immune system. Typically, one of these epitopes generates the most strong immune response, i.e. is dominant. This dominant epitope inhibits the immune response to the subdominant epitopes and this immunodominance phenomenon reduces the diversity of the immune response to such a disease [7]. The essence of this immunodominance phenomenon is competition of TCRs for epitope on antigen presenting cells [7, 14]. This immunodominance phenomenon is important to understand because it can render a multivalent vaccine effectively monovalent.

Immunodominance is one general mechanism by which cancer cells may escape, either by mutation of the dominant epitope or by loss of the MHC class I allele that expresses the dominant epitope [15]. Cross-presentation of the lost dominant epitope on surrounding cells often sustains the futile immune response [8]. Indeed, cancerous cells of many types are exceptionally adept at evading the immune response [4]. It has been noticed that not only the quantity, but also the quality of the T cell response induced by therapeutic vaccination is important for clinical efficacy [16]. For these reasons, it has been suggested that multiple immune-stimulation strategies will be necessary to avoid escape from the immune response by cancer [4, 17].

By sculpting the diversity of the effector T cell receptor (TCR) repertoire, immune evasion of tumor cells can be reduced. Here we propose a mechanism by which the immunodominance hierarchy that allows a growing tumor to escape from immune surveillance may be broken. Our approach captures the essence of the interaction between TCRs and antigenic epitopes and the primary selection dynamics of TCRs within lymph nodes due to TCR resource competition [18].

1.1 The Polytopic Vaccination Experiment

Immunodominance implies that the immune response to multiple cancer-specific epitopes may be incomplete because the response may be directed primarily against the dominant epitope. That tumors are suppressed by *in vitro*-derived T cells directed against two or more epitopes [19–21] suggests the failure may be in the development of the immune response rather than in the intrinsic lytic ability of the T cells.

One suggestion for breaking the effectively monovalent response to a multivalent cancer vaccine is to inject the different vaccine epitopes or strains in different physical regions of the patient [8]. When this is done for the 1591-A/B system, a T cell response is generated against the dominant epitope A, and a response is generated against the subdominant epitope B as well [19]. Specifically, cells expressing the A epitope are lysed by vaccine A or by vaccines AB and B injected in the same site. Cells expressing only the subdominant B epitope are not. Cells expressing only the subdominant B epitope are lysed by vaccine B or by vaccines AB and B injected in different sites, not by vaccine A nor by vaccines AB and B injected in the same site. Thus, it appears that vaccination with individual tumor epitopes at separate sites rather than with multiple epitopes at one site may be needed to prevent tumor escape and recurrence of cancer.

These experiments have not been widely cited by other researchers. Moreover the mechanism for the reduction in immunodominance has not been described. We seek to shed some light on the possible mechanism by polytopic vaccination may reduce immunodominance. With such a mechanism in hand, further experiments to confirm or refute the hypothesis can be performed.

We investigate the hypothesis that polytopic, or multi-site, administration of a therapeutic cancer vaccine may sculpt a broader immune response. We here study the proposed mechanism of polytopic vaccination by developing a sequence-level model of immune response to polytopic cancer vaccines [22]. Our approach captures the recognition characteristics between the T cell receptors (TCRs) and tumor, the primary dynamics due to TCR resource competition, and elimination of tumor cells by effector TCRs. We focus the discussion on reducing the deleterious effects of immunodominance and on strategies that may achieve an effective strategy in the face of central and peripheral tolerance. The model we develop complements the long and difficult process of experimental vaccine development. The model takes explicit account of the dynamics of the 10^8 different T-cell sequences that exist within an individual.

We study the interactions between the response of the immune system to the various cancer tissue-specific epitopes [23–25]. For specificity, we consider $V = 2$ or $V = 4$ tissue-specific epitopes. Each antigen is an epitope of 9 amino acids [26]. The primary immune response lasts for 10 rounds of T cell division. After this period of time, there is a high concentration of T cell receptors specific for the tissue-specific epitopes. In single-site vaccination, the immune system simultaneously shapes the T cell repertoire based upon all epitopes. In polytopic vaccination, on the other hands, the immune system responds to each epitope independently in distinct lymph nodes, and only after some number of days is there a significant mixing of the evolved T cell repertoires. We denote the day after which the T cell repertoires for the V distinct epitopes begin to compete as mixing day. We present results for a range of this parameter, mixing day.

The dependence on the spatial separation of the vaccination sites is an interesting question. There are three times to consider. The first is the time it takes for the injected vaccine to localize to the draining lymph node. This is on the order of 1–12 hours [27]. The second is the time for the T cells to grow to such a concentration that they begin to leave the lymph node. This time is known to be 4–5 days [1]. Finally, there is the time it takes the T cells that have left a lymph node to transport along the major lymphatics to the thoracic duct (or right lymphatic duct for lymph fluid from the right upper arm and head), where the T cells then mix with the blood. The characteristic timescale for complete circulation of the lymph system is an additional 4 days [28]. The time for a significant number of T cells to reach the bloodstream from the lymph can be a bit shorter, on the order of 2 additional days, for T cells leaving lymph nodes located closer to the heart [29,30]. Thus, for vaccination that drains to lymph nodes relatively far from the heart, the time it takes for T cells from different injection sites to mix through the lymph is on the order of 7–9 days. We term this time the "mixing day."

The focus of our study is a comparison of vaccination at distant lymph nodes, e.g. mixing day = 7–9, to vaccination at the same lymph node, e.g. mixing day = 0. Drainage to a specific lymph node can be enhanced in polytopic vaccination by using antigen-bearing dendritic cells [31,32]. While interactions due to transport can be important for lymph nodes very close together on a lymphatic, for polytopic vaccination we will choose well-separated sites, so that these interactions need not be considered. Typical values of mixing day are shown in Fig. 1. Values for vaccination that drains to lymph nodes relatively far from the heart are in the range 7–9 days. For an effective polytopic vaccination, we would choose sites such that mixing day is large, e.g. mixing day = 9.

Since immunodominance is a competition of TCRs for alternative epitopes [7, 14], we capture this effect within our model by the inclusion of multiple, sub-dominant, epitopes for each strain. Immunodominance causes the T cell immune response to sequential exposure to two different epitopes to depend on the order of their presentation. For example, in the 1591 system, where the A epitope is dominant and the B epitope is subdominant, exposure to a mixed vaccine of A and B generates a response primarily against A. Exposure to a single vaccine of B, however, generates a response against B, even if a subsequent vaccine of A and B is given [19]. The same phenomenon occurs in the 8101 AB system [34]. We will model this phenomenon by performing primary responses to vaccines of A, B, or A and B in the same site. The results will be compared to primary vaccination of A and B in separate sites.

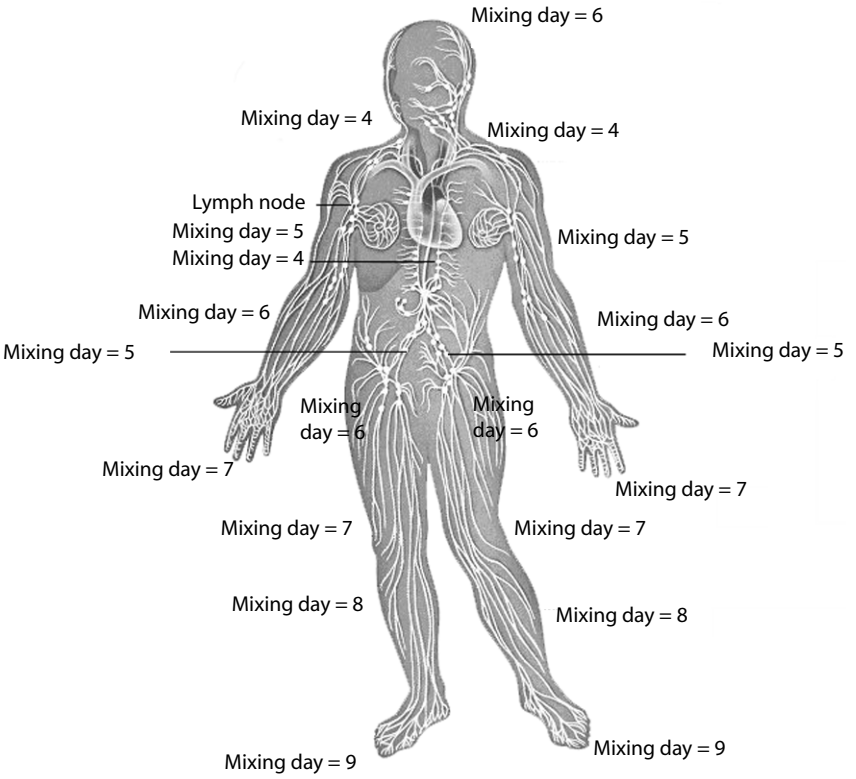


Fig. 1. For effective polytopic vaccination, well-separated sites on different limbs are used. The value of the parameter mixing day for vaccination to draining lymph nodes at different distances from the heart [33]. Humans have several hundred lymph nodes. For effective polytopic vaccination, well-separated sites on different limbs are used

2 The Statistical Mechanical Model of the Immune Response to Cancer

2.1 Generalized NK Model

We use a spin glass model to represent the interaction between the T cell receptors and the epitopes [22, 35, 36]. This model captures the essence of the correlated ruggedness of the interaction energy in the variable space, the variables being the T cell amino acid sequences and the identity of the disease proteins, and the correlations being mainly due to the physical structure of the T cell receptors. The random energy model allows study of the sequence-level dynamics of the immune/antigen system, which would otherwise be an intractable problem at the atomic scale, with 10^4 atoms per T cell receptor,

10^8 antibodies per individual, 6×10^9 individuals, and many possible cancer epitopes.

The generalized NK model, while a simplified description of real proteins, captures much of the thermodynamics of protein folding and ligand binding. In this model N is the number of amino acids in a secondary structure, and K is the range of the local interaction. In the model, a specific T cell repertoire is represented by a specific set of amino acid sequences. Moreover, a specific instance of the random parameters within the model represents a specific epitope. An immune response that finds a T cell that produces a T cell receptor with a high affinity constant to a specific epitope corresponds in the model to finding a sequence having a low energy for a specific parameter set.

Use of a sequence-level model allows for a broad range of predictions. We can, for example, predict altered peptide ligand experiments [22]. Such a model also allows one to predict the reduced immune response to a mutated cancer antigen. Such a model, therefore, has a broader range of applicability than a model with an *ad hoc* assumption about the distribution of binding constants.

The generalized NK model for the T cell response considers four different kinds of interactions: interactions within a subdomain of the TCR (U^{sd}), interactions between subdomains of the TCR ($U^{\mathrm{sd-sd}}$), interactions between the TCR and the peptide ($U^{\mathrm{pep-sd}}$), and direct binding interaction between the TCR and peptide (U^{c}) [22,36]. A figure of the TCR-peptide MHCI complex is shown in Fig. 2. The direct interactions are distinguished as resulting from a limited number of "hot spot" interactions. The model returns the free energy (U) as a function of the TCR (a_j) and epitope (a_j^{pep}) amino acid sequence.

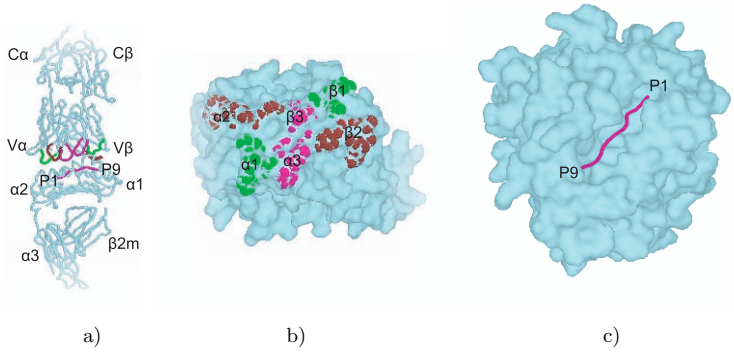


Fig. 2. Three-dimensional structures of the TCR-MHC-peptide complex. PDB accession number 2CKB. (**a**) Backbone tube diagram of the ternary complex of mouse TCR bound to the class I MHC H-2Kb molecule and peptide (magenta tube numbered P1–P9) (**b**) CDR regions of mouse TCR α and β chains viewed from above, showing the surface that is involved in binding the MHC-peptide complex. (**c**) Molecular surface of the class I MHC H-2Kb molecule and peptide (magenta tube numbered P1–P9) viewed from the above

Briefly:

$$U = \sum_{i=1}^{M} U_{\alpha_i}^{\mathrm{sd}} + \sum_{i>j=1}^{M} U_{ij}^{\mathrm{sd-sd}} + \sum_{i=1}^{M} U_{i}^{\mathrm{pep-sd}} + \sum_{i=1}^{N_{\mathrm{b}}} \sum_{j=1}^{N_{\mathrm{CON}}} U_{ij}^{\mathrm{c}} \ . \tag{1}$$

$$U_{\alpha_i}^{\mathrm{sd}} = \frac{1}{\sqrt{M(N-K+1)}} \times \sum_{j=1}^{N-K+1} \sigma_{\alpha_i}(a_j, a_{j+1}, \cdots, a_{j+K-1}) \ . \tag{2}$$

$$U_{ij}^{\mathrm{sd-sd}} = \sqrt{\frac{2}{DM(M-1)}} \times \sum_{k=1}^{D} \sigma_{ij}^{k}(a_{j_1}^{i}, \cdots, a_{j_{K/2}}^{i}; a_{j_{K/2+1}}^{j}, \cdots, a_{j_K}^{j}) \ . \tag{3}$$

$$U_{i}^{\mathrm{pep-sd}} = \sqrt{\frac{1}{DM}} \times \sum_{k=1}^{D} \sigma_{i}^{k}(a_{j_1}^{\mathrm{pep}}, \cdots, a_{j_{K/2}}^{\mathrm{pep}}; a_{j_{K/2+1}}^{i}, \cdots, a_{j_K}^{i}) \ . \tag{4}$$

$$U_{ij}^{\mathrm{c}} = \frac{1}{\sqrt{N_{\mathrm{b}} N_{\mathrm{CON}}}} \sigma_{ij}(a_{j_1}^{\mathrm{pep}}, a_{j_2}) \ . \tag{5}$$

Here $M = 6$ is the number of TCR secondary structural subdomains, $N_{\mathrm{b}} = 3$ is the number of hot-spot amino acids that directly bind to the TCR, and $N_{\mathrm{CON}} = 3$ is the number of T cell amino acids contributing directly to the binding of each peptide amino acid. $N = 9$ is the number of amino acids in a subdomain, and $K = 4$ is the range of local interaction within a subdomain. All subdomains belong to one of $L = 5$ different types (e.g., helices, strands, loops, turns, and others). The σ are matrices with random elements that mediate the interactions. The quenched Gaussian random number σ_{α_i} is different for each value of its argument for a given subdomain type, α_i. All σ values in the model are Gaussian random numbers and have zero mean

and unit variance. The sigma values are different for each value of the argument, subscript, or superscript. The variable α_i defines the type of secondary structure for the i^{th} subdomain, $1 \leq \alpha_i \leq L$. The number of interactions between secondary structures is $D = 2$. The σ_{ij}^k and the interacting amino acids, $j_1, \ldots, j_K$, are selected at random for each interaction (i, j, k). The σ_i^k and the interacting amino acids, $j_1, \ldots, j_K$, are selected at random in the peptide and TCR subdomain for each interaction (i, k). The contributing amino acids, j_1, j_2, and the unit-normal weight of the binding, σ_{ij}, are chosen at random for each interaction (i, j), with N_{b} possible values for j_1, and NM possible values for j_2. To consider all 20 amino acids within the random energy model, (1), and to consider the differing effects of conservative and non-conservative mutations, we set the random σ for amino acid i that belongs to group j as $\sigma = w_j + w_i/2$, where the w are Gaussian random numbers with zero average and unit standard deviation. There are 5 groups, with 8 amino acids in the neutral and polar plus cystein group, 2 amino acids in the negative and polar group, 3 amino acids in the positive and polar group, 4 amino acids in the nonpolar without ring group, and 3 amino acids in the nonpolar with ring group. Table 1 lists all the parameter values for the Generalized NK model used here.

The binding constant is related to the energy by

$$K = e^{a-bU} \ . \tag{6}$$

We determine the values of a, b in each instance of the ensemble by fixing the geometric average TCR:p-MHC I affinity to be $K = 10^4$ l/mol and minimum affinity to be $K = 10^2$ l/mol [37] for the $V \times N_{\text{size}} = V \times 10^8/10^5 = V \times 1000$ naive TCRs that respond to all V epitopes [37, 38]. This means that for the highest affinity TCR, K fluctuates between 10^5 l/mol and 10^7 l/mol for the different epitopes [39]. The distribution is shown in Fig. 3.

Specific lysis is a measure of the probability that an activated T cell will recognize an cell that is expressing a particular peptide-MHC I complex. It is given by [22]

$$L = \frac{zE/T}{1 + zE/T} \ , \tag{7}$$

where E/T is the effector to target ratio. The quantity z is the average clearance probability of one TCR:

$$z = \frac{1}{N_{\text{size}}} \sum_{i=1}^{N_{\text{size}}} \min(1, K_i/10^6) \ . \tag{8}$$

Note the clearance probability is related to both affinity and clonal sizes. Typical naive values will be approximately 5/1000, and typical values after a primary response will be in the range 0.1 to 1.0, due to the increased copy number of the selected T cells. Specific lysis is a standard measure of the immunological response to a vaccine, which correlates well with vaccine efficacy [40]. On

Table 1. Parameter values for the Generalized NK model

	Parameter	Value	Definition
TCR			
	a_j		Identity of amino acid at sequence position j
	M	6	Number of secondary structure subdomains
	N	9	Number of amino acids in each subdomain
	L	5	Number of subdomain types (e.g., helices, strands, loops, turns and others)
	α_i	$1 \leq \alpha_i \leq L$	Type of secondary structure for the i^{th} subdomain
	K	4	Local interaction range within a subdomain
	σ_{α_i}		Local interaction coupling within a subdomain, for subdomain type α_i
	D	2	Number of interactions between subdomains
	σ_{ij}^k		Nonlocal interaction coupling between secondary structures
Epitope			
	a_j^{pep}		Identity of amino acid of epitope at sequence position j
	N	9	Number of amino acids in epitope
	V	1–4	Number of epitopes on tumor cell
TCR–Epitope			
	N_{b}	3	Number of hot-spot amino acids in the epitope
	N_{CON}	3	Number of amino acids in TCR that each hot spot interacts with
	σ_{ij}		Interaction coupling between TCR and epitope
	σ_i^k		Interaction coupling between TCR secondary structure and epitope
Random couplings			
	w		Gaussian random number with zero average and unit standard deviation
	σ	$w_j + w_i/2$	Value of coupling for amino acid i of non-conservative type j

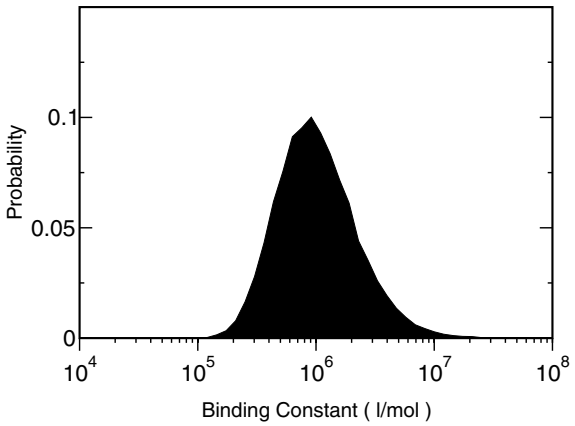


Fig. 3. Probability distribution of the binding constant of the highest affinity TCR. The highest affinity fluctuates between 10^5 l/mol and 10^7 l/mol, in agreement with experiment [39]. The median upper affinity is 10^6 l/mol

the order of 1–3 TCR/peptide-MHCI interactions are enough for killing [41], and the TCR affinity is highly correlated with proliferation [18, 42]. Binding constants larger than roughly 10^6 l/mol do not increase the lysis, hence the bound in (8).

2.2 TCR Selection Dynamics

The cellular immune system performs a search of T cell receptors (TCRs) that recognize antigenic peptide ligands bound to the MHCI molecule. The T cells are activated by ligand binding to the multiple identical TCRs on the T cell membrane [39, 43, 44]. TCRs are constructed from modular elements, and each individual has an approximate diversity of 10^8 different receptors [45]. TCRs undergo rounds of selection for increased avidity [46–49]. TCRs do not undergo any further mutation during the immune response. Those TCRs that are stochastically selected during the primary response become memory cells [50, 51]. While there may well be other details which affect the T cell selection process, selection for increased affinity has been shown to be an important factor [18, 51–54]. A schematic of the selection model of the T cell immune system maturation is shown in Fig. 4.

The naive TCR repertoire is generated randomly from gene fragments. This is accomplished by constructing the TCRs from subdomain pools. Fragments for each of the L subdomain types are chosen randomly from 13 of the 100 lowest energy subdomain sequences. This diversity mimics the known TCR diversity, $(13 \times L)^M \approx 10^{11}$ [55]. Only 1 in 10^5 naive TCRs responds to any particular antigen, and there are only 10^8 distinct TCRs present at any one point in time in the human immune system [37, 38], so the primary response starts with a repertoire of $N_{\text{size}} = 10^3$ distinct TCRs. The initial

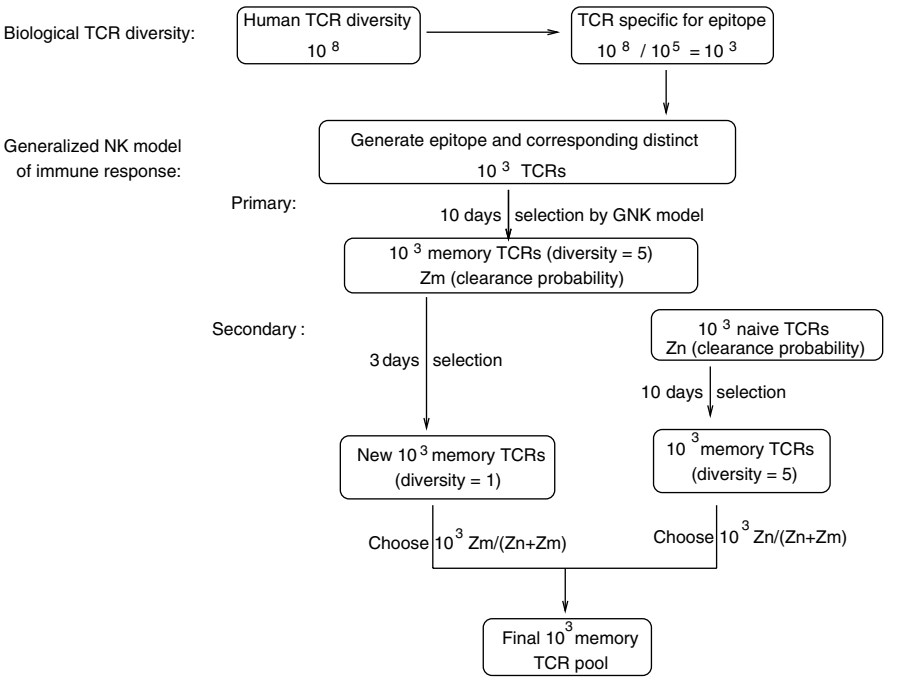


Fig. 4. Selection of TCRs in the primary response. The secondary dynamics is also shown, although it is not used in the present application to cancer

TCR repertoire is redetermined for each realization of the model – this must be done because the U^{sd} that defines the TCR repertoire is different in each instance of the ensemble.

The T cell-mediated response is driven by cycles of concentration expansion and selection for better binding constants. The primary response increases the concentration of selected TCRs by 1000 fold over 10 days, with a rough T cell doubling time of one day [1]. The diversity of the memory sequences is 0.5% of that of the naive repertoire [45]. Thus, while the number of distinct T cells selected by the expansion processes is 5 out of 1000, the copy number of each of the 5 clones increases. Specifically, 10 rounds of selection are performed during the primary response, with the top $x = 58\%$ of the sequences chosen at each round. This procedure mimics the concentration expansion factor of $10^3 \approx 2^{10}$ in the primary response and leads to 0.5% diversity of the memory repertoire, because $0.58^{10} \approx 0.5\%$ and 10 days of doubling leads to a concentration expansion of $2^{10} = 1024$. Thus, the repertoire killing the tumor consists of these 5 sequences, each at an average copy number of 200. After 10 rounds, about 1% of the T cells in the human immune system will be specific for the cancer epitopes. Each additional round will double this number, up to a maximum of 100%, since the total number of memory T cells is roughly constant. We will look at the effect of an additional couple of rounds, beyond

the 10 rounds that leads to the physiological 1000 × concentration expansion. There cannot be more than a couple of additional rounds, because the percentage of T cells specific for a disease does not typically exceed 1–5% in the human immune system [56].

For polytopic vaccination, the different epitopes are injected in different physical locations and evoke an immune response that evolves independently in different lymph nodes until mixing round, after which the lymph system is well-mixed. This is calculated by performing a response against each of the four epitopes independently, with selection among 10^3 TCRs for each epitope, until mixing round. After the mixing round, T cells from one of the V lymph nodes are able to renter the other $V - 1$ lymph nodes. So, all V lymph nodes will begin to contain roughly similar repertoires of T cells. Thus, we consider a representative lymph node, with a T cell capacity equal to that of one of the V lymph nodes, and in which the T cells from all V lymph nodes may be present. In other words, at the mixing round, 10^3 TCRs are randomly chosen from the $V \times 10^3$ partially evolved TCRs. These 10^3 TCRs are then evolved from mixing round until day 10. During the period from mixing round until day 10, the TCRs are ranked by the sum of the V binding constants when the top 58% selection is performed. In this way, we compute the independent response in different lymph nodes against each epitope that occurs early on and the combined response in a typical lymph node against all epitopes that occurs after the lymph system has mixed. To make comparison with Fig. 1, the conversion between mixing day and mixing round is needed. Since T cells divide 1–3 times a day [1], mixing round will be mixing day divided by the factor of 1–3. For this reason, we expect mixing round to be near the maximum value of 10 for polytopic vaccination. Single-site vaccination is computed with mixing round = 0, and polytopic vaccination is computed with mixing round = 9.

Implementation of our theory proceeds by computational simulation of the generalized NK model. First, the peptide and the altered peptide ligands are created. Then, the random terms such as the secondary structural types of the subdomains, the σ values, the binding sites, and the interaction sites of the generalized NK model are determined. Then the fraction and identity of the T cell repertoire that responds well to each epitope is identified. We take $N_{\text{size}} = 10^8/10^5 = 1000$ distinct T cells to participate in the naive response against each epitope. Then the values of the constants a and b in (6) are calculated. The primary response of 10 rounds of selection is then carried out. Finally, the secondary response, a linear combination of 10 rounds from the naive pool and 3 rounds from the memory pool, with the fraction of each depending on the relative binding constants for the altered peptide, is carried out. This secondary response is not modeled in the present application to cancer.

2.3 Tumor-Related Parameters

For each epitope, the sequence, model, and VDJ selection pools differ by p_{epitope} [36], where $p_{\text{epitope}} = [(\text{non-conservative} + \frac{1}{2}\ \text{conservative})\ \text{amino acid difference in epitope}]/[\text{total number of amino acids in epitope}]$. In other words, $V \times 10^3$ naive TCRs respond to the V epitopes, but on average only $10^3[1 + (V-1)p_{\text{epitope}}]$ of them are distinct. To generate results, an average over many instances of these random epitope sequences, models, and VDJ selection pools that differ by p_{epitope} is taken. That is, the V different cancer epitopes were chosen so that they differ by the requisite p_{epitope}. For each instance of the ensemble, V new random epitopes were generated. We display results for correlated and uncorrelated epitopes by using the two values of $p_{\text{epitope}} = 0.15, 1.00$.

3 Results

Figure 5a shows the typical level of immunodominance that occurs for a traditional two-component vaccine in the 1591-A/B system [19]. Figure 5b shows the reduced immunodominance that occurs if each component of the vaccine is injected into a different physical location, with a different draining lymph node. Briefly, there are two cancer strains. One is strain A and the other is strain B. It is assumed that the response to each strain is to a single epitope and that the two epitopes are unrelated. The epitope on strain A is dominant to the epitope on strain B. In Fig. 5a, epitope A and epitope B were injected together in the same flank of one animal, while they are injected in different flanks of one animal in Fig. 5b. After 1 month, spleen cells from the animal were restimulated with epitope A or B and the T cells assayed for specific lysis.

While epitope A is dominant, there can be a significant immune response to epitope B. In particular, if vaccination is done against epitope B, there is response only against epitope B, Fig. 6b. Conversely, if vaccination is done against epitope A, there is response only against epitope A, Fig. 6a. If single-site vaccination is done against both A and B, there is response primarily against A, Fig. 5a. Conversely, if polytopic vaccination is done against both A and B, there is a significant response against both A and B, Fig. 5b.

4 Discussion

To compare single-site and polytopic vaccination, we show in Fig. 7ac the immune responses to the four epitopes induced by single-component, multi-component single-site, and multi-component polytopic vaccinations. This figure shows that polytopic vaccination induces a significantly enhanced response, especially against the subdominant epitopes. The poor recognition

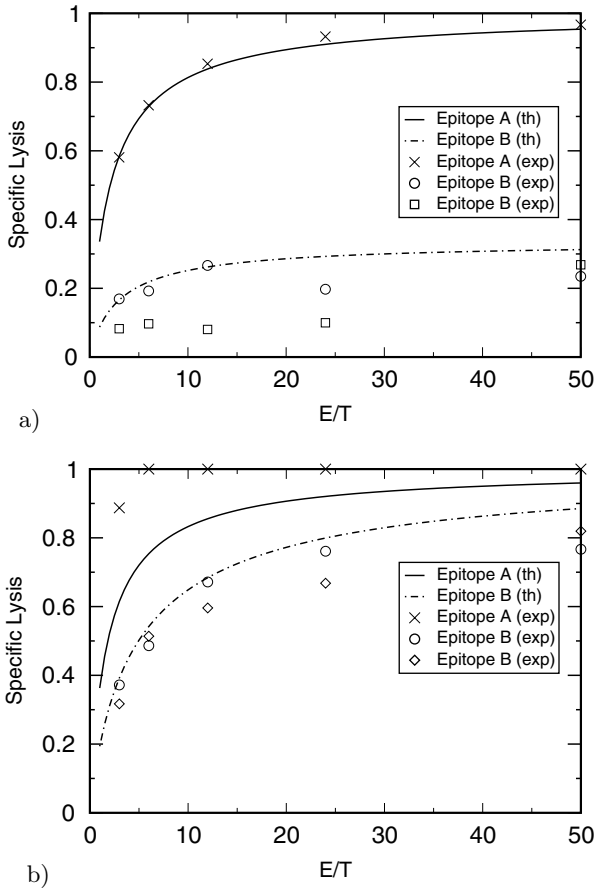


Fig. 5. Specific lysis of two-component single-site and polytopic vaccination. (**a**) The specific lysis arising from the T cell repertoire induced if two completely different ($p_{\text{epitope}} = 1$) cancer epitopes (A,B) are injected into the same site. Epitope A is dominant. Comparison between theory with no adjustable parameters (th) and experiment (exp) [19]. Two sets of experimental data for epitope B are shown. (**b**) The specific lysis arising from the T cell repertoire induced if two cancer epitopes are injected into different sites, with different draining lymph nodes. Note the reduced immunodominance. Data contain $\sim$+20% background, which was subtracted from the clearance probability

of the subdominant epitopes is an especially important mechanism for immune evasion by mutation: If the dominant epitope mutates and is lost, the response to the remaining epitopes is very low. This effect is shown in Table 2.

The diversity of the TCR repertoire plays a key role in inducing a long-lasting and broad immune response [38]. Some insight into the diversity of the T cell repertoire can be gained from the histograms of the T cell reper-

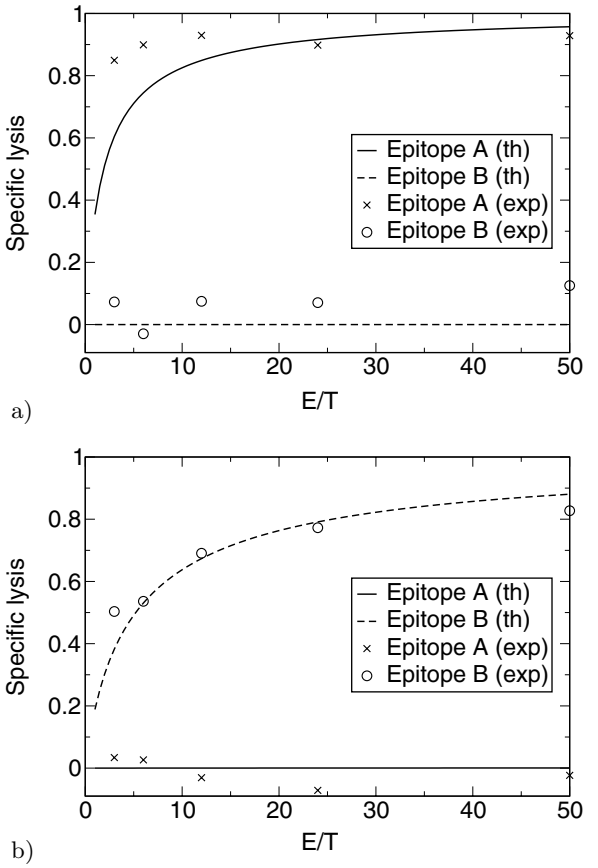


Fig. 6. Specific lysis for the 1591-$A^{\pm}B^{+}$ system. Epitope A is dominant. (**a**) Vaccination against epitope A. (**b**) Vaccination against epitope B. Curve is from the model, and data are from [19]. Data contain ∼+20% background, which was subtracted from the clearance probability

toire recognition of each epitope, Fig. 7bd. This figure shows that the T cell repertoire of the immune response is shaped by the epitopes presented on the tumor cell. This figure also shows that the method of delivering the vaccine affects the T cell repertoire induced by the immune response, Fig. 7bd. In particular, polytopic vaccination leads to a broader and more robust repertoire, Fig. 7bd. This broader response of the polytopic vaccination, in turn, leads to less immunodominance, Fig. 5.

The solid nature of tumors presents some challenges to immune control, but also some opportunities. It may be difficult, for example, for T cells to enter the solid tumor. Conversely, high intensity focused ultrasound can disrupt solid tumors and can enhance systemic antitumor cellular immunity [57]. Although the exact mechanism of this enhancement is unknown, one possi-

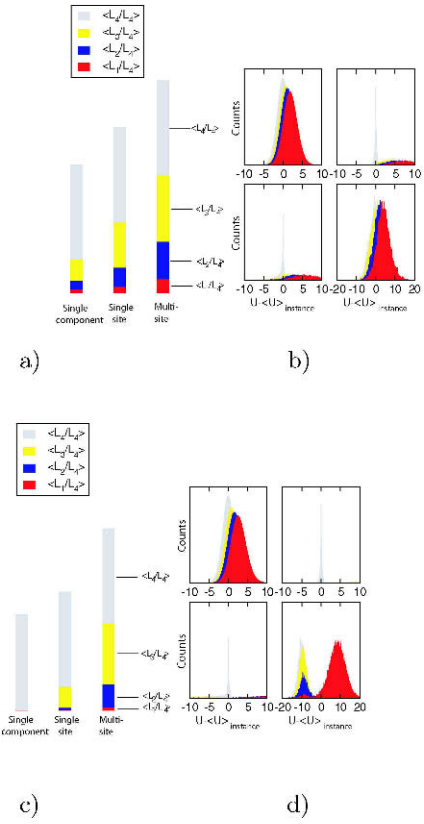


Fig. 7. Immunodominance hierarchy and energy profile of different vaccination protocols. (**a**) The immunodominance hierarchy if only the dominant epitope is injected (*left*), all four epitopes are injected in the same site (*middle*), and the four epitopes are injected into different draining lymph nodes (*right*). The four epitopes are related by $p_{\text{epitope}} = 0.15$. (**b**) The histograms of the naive (*upper left*), single-component (*upper right*), multi-component single-site (*lower left*), and multi-component polytopic (*lower right*) T cell repertoire binding energies for each cancer epitope (arbitrary consistent units). The four epitopes are related by $p_{\text{epitope}} = 0.15$. (**c**) The immunodominance hierarchies if $p_{\text{epitope}} = 1.0$. (**d**) The histograms if $p_{\text{epitope}} = 1.0$. Energy, U, is related to the binding constant, K, by (6). Here $E/T = 3.0$

bility is that the fragments of tumor after destruction can travel to different lymph nodes and thus induce a diverse TCR repertoire, in similar fashion to polytopic vaccination. More prosaically, the physical disruption of the tumor can allow easier entry of the T cells. Another feature of solid tumors is the enhanced probability of uptake of large proteins, due to the highly porous capillaries in tumors [58]. By using this property, stimulants of T cell activity may be localized to the tumors. The concentration of such stimulants could be increased further by conjugation with elastin-like polypeptides that undergo

Table 2. Probability of killing the tumor cell, P_{killing}, calculated from the specific lysis values of Fig. 7

P_{killing}	$p_{\text{epitope}} = 0.15$		$p_{\text{epitope}} = 1.00$	
	Normal	Dominant epitope lost	Normal	Dominant epitope lost
Single-component	0.704	0.202	0.627	0.00026
Multi-component single-site	0.693	0.349	0.562	0.120
Multi-component polytopic	0.707	0.467	0.602	0.334

a thermally triggered phase transition in heated tumors, causing selective aggregation in the tumor [58].

5 Conclusions

In this chapter, we have described a possible mechanism by which polytopic vaccination overcomes immunodominance in therapeutic T cell vaccines for cancer. Our approach captured the recognition characteristics between the TCRs and tumor, the primary dynamics due to the TCR resource competition, and elimination of tumor cells by effector TCRs. The ability of our model to reproduce standard cancer immunological data was demonstrated. Multi-component polytopic vaccination was shown to be a promising protocol to mitigate immunodominance.

Therapeutic cancer vaccine design is undergoing a resurgence not only because of its clinical importance, but also because of the many fascinating associated scientific issues. We hope that by suggesting here the mechanism of action for polytopic vaccination to be independent concentration expansion of T cells in distinct lymph, additional studies making use of the polytopic protocol will be spurred.

Acknowledgments

The authors thank Hans C. Schreiber for insightful discussions. This research was supported by the U.S. National Institutes of Health.

References

1. C. A. Janeway, P. Travers, M. Walport, and M. Shlomchik (2004) *Immunobiology: The Immune System in Health and Disease.* Taylor & Francis Inc., New York, 6th edition

2. E. Gilboa (2004) The promise of cancer vaccines. *Nature Rev. Cancer* **4**, pp. 401–411
3. O. J. Finn (2003) Cancer vaccines: Between the idea and the reality. *Nature Rev. Immunol.* **3**, pp. 630–641
4. M. Whelan, J. Whelan, N. Russell, and A. Dalgleish (2003) Cancer immunotherapy: An embarrassment of riches? *Drug Discov. Today* **8**, pp. 253–258
5. G. P. Dunn, L. J. Old, and R. D. Schreiber (2004) The three Es of cancer immunoediting. *Annu. Rev. Immunol.* **22**, pp. 329–360
6. D. Pardoll (2003) Does the immune system see tumors as foreign or self? *Annu. Rev. Immunol.* **21**, pp. 807–839
7. G. Roy-Prouix, M.-C. Meunier, A.-M. Lanteigne, and S. Brochu (2001) Immunodomination results from functional differences between competing CTLs. *Eur. J. Immunol.* **31**, pp. 2284–2292
8. H. Schreiber, T. H. Wu, J. Nachman, and W. M. Kast (2002) Immunodominance and tumor escape. *Cancer Biology* **12**, pp. 25–31
9. H. T. Khong and N. P. Restifo (2002) Natural selection of tumor variants in the generation of "tumor escape" phenotypes. *Nature Immunol.* **3**, pp. 999–1005
10. M. Campoli, C.-C. Chang, and S. Ferrone (2002) HLA class I antigen loss, tumor immune escape and immune selection. *Vaccine* **20**, pp. A40–A45
11. L. Müller, R. Kiessling, R. C. Rees, and G. Pawelec (2002) "Escape mechanisms in tumor immunity: An update. *J. Env. Path., Tox., and Onc.* **21**, pp. 277–330
12. B. Stockinger, T. Barthlott, and G. Kassiotis (2004) The concept of space and competition in immune regulation. *Immunology* **111**, pp. 241–247
13. L. Miller (2003) Destination: Tumor eradication. *Nature Immunol.* **4**, p. 813
14. J. Sprent and C. D. Surh (2002) T cell memory. *Annu. Rev. Immunol.* **20**, pp. 551–579
15. G. P. Dunn, A. T. Bruce, H. Ikeda, L. J. Old, and R. D. Schreiber (2002) Cancer immunoediting: From immunosurveillance to tumor escape. *Nature Immunol.* **3**, pp. 991–998
16. T. B. Stuge, S. P. Holmes, S. Saharan, A. Tuettenberg, M. Roederer, and J. S. Weber (2004) Diversity and recognition efficiency of T cell responses to cancer. *PLOS Medicine* **1**, pp. 149–160
17. M. A. Markiewicz and W. M. Kast (2004) Progess in the development of immunotherapy of cancer using ex vivo-generated dendritic cells expressing multiple tumor antigen epitopes. *Canc. Investi.* **22**, pp. 417–434
18. R. M. Kedl, J. W. Kappler, and P. Marrack (2003) Epitope dominance, competition and T cell affinity maturation. *Curr. Opin. Immunol.* **15**, pp. 120–127
19. C. Van Waes, P. A. Monach, J. L. Urban, R. D. Wortzel, and H. Schreiber (1996) Immunodominance deters the response to other tumor antigens thereby favoring escape. Prevention by vaccination with tumor variants selected with cloned cytolytic T cells *in vitro*. *Tissue Antigens* **47**, pp. 399–407
20. C. J. Melief and W. M. Kast (1994) Prospects for T cell immunotherapy of tumors by vaccination with immunodominant and subdominant peptides. *Ciba Found. Symp.* **187**, pp. 97–104
21. M. C. Feltkamp, G. R. Vreugdenhil, M. P. Vierboom, E. Ras, and S. H. van der Burg (1995) Cytotoxic T lymphocytes raised against a subdominant epitope offered as a synthetic peptide eradicate human papillomavirus type 16-induced tumors. *Eur. J. Immunol.* **25**, pp. 2638–2642
22. J.-M. Park and M. W. Deem (2004) Correlations in the T cell response to altered peptide ligands. *Physica A* **341**, pp. 455–470

23. P. A. Futreal, L. Coin, M. Marshall, T. Down, T. Hubbard, R. Wooster, N. Rahman, and M. R. Stratton (2004) A census of human cancer genes. *Nature Rev. Cancer* **4**, pp. 177–183
24. F. Michor, Y. Iwasa, and M. A. Nowak (2004) Dynamics of cancer progression. *Nature Rev. Cancer* **4**, pp. 197–205
25. M. T. Spiotto, D. A. Rowley, and H. Schreiber (2004) Bystander elimination of antigen loss variants in established tumors. *Nature Med.* **10**, pp. 294–298
26. M. A. Brehm, L. K. Serlin, and R. M. Welsh (2004) CD8 T cell responses to viral infections in sequence. *Cellular Microbiology* **6**, pp. 411–421
27. R. Ridolfi, A. Riccobon, R. Galassi, G. Giorgetti, M. Petrini, L. Fiammenghi, M. Stefanelli, L. Ridolfi, A. Moretti, G. Migliori, and G. Fiorentini (2004) Evaluation of in vivo labelled dendritic cell migration in cancer patients. *J. Translational Medicine* **22**, pp. 27–37
28. R. L. Fournier (1999) *Basic Transport Phenomena in Biomedical Engineering.* Taylor & Francis, Philadelphia
29. L. S. Caulty, T. Cookenham, R. J. Hogan, S. R. Crowe, and D. L. Woodland (2003) Renewal of peripheral $CD8^+$ memory T cells during secondary viral infection of antibody-sufficient mice. *J. Immunol.* **170**, pp. 5597–5606
30. T. R. Mempel, S. E. Henrickson, and U. H. von Adrian (2004) T-cell priming by dendritic cells in lymph nodes occurs in three distinct phases. *Nature* **427**, pp. 154–159
31. F. Sallusto, J. Geginat, and A. Lanzavecchia (2004) Cental memory and effector memory T cell subsets: Function, generation and maintenance. *Annu. Rev. Immunol* **22**, pp. 745–763
32. J. Jacob, J. Przylepa, C. Miller, and G. Kelsoe (1993) In situ studies of the primary immune response to (4-hydroxy-3-nitrophenyl)acetyl. III. The kinetics of V region mutation and selection in germinal center B cells. *J. Exp. Med.* **178**, pp. 1293–1307
33. Lymph Man. Image after www.ener-chi.com/Image_13.jpg
34. P. Dubey, R. C. Hendrickson, S. C. Meredith, C. T. Siegel, J. Shabanowitz, J. C. A. Skipper, V. H. Engelhard, D. F. Hunt, and H. Schreiber (1997) The immunodominant antigen of an ultraviolet-induced regressor tumor is generated by a somatic point mutation in the DEAD box helicase p68. *J. Exp. Med.* **185**, pp. 695–705
35. L. D. Bogarad and M. W. Deem (1999) A hierarchical approach to protein molecular evolution. *Proc. Natl. Acad. Sci. USA* **96**, pp. 2591–2595
36. M. W. Deem and H.-Y. Lee (2003) Sequence space localization in the immune system response to vaccination and disease. *Phys. Rev. Lett.* **91**, p. 068101
37. R. M. Zinkernagel and H. Hengartner (2001) Regulation of the immune response by antigen. *Science* **293**, pp. 251–253
38. A. W. Goldrath and M. J. Bevan (1999) Selecting and maintaining a diverse T-cell repertoire. *Nature* **402**, pp. 255–262
39. P. A. van der Merwe and S. J. Davis (2003) Molecular interactions mediating T cell antigen recognition. *Annu. Rev. Immunol.* **21**, pp. 659–684
40. H. L. Robinson and R. R. Amara (2005) T cell vaccines for microbial infections. *Nature Med.* **11**, pp. S25–S32
41. M. M. Davis, M. Krogsgaard, J. B. Huppa, C. Sumen, M. A. Prubhoo, D. J. Irvine, L. C. Wu, and L. Ehrlich (2003) Dynamics of cell surface molecules during T cell recognition. *Annu. Rev. Biochem.* **72**, pp. 717–742

42. Q. Ge, A. Bai, B. Jones, H. N. Eisen, and J. Chen (2004) Competition for self-peptide-MHC complexes and cytokines between naive and memory CD8+ T cells expressing the same or different T cell receptors. *Immunol.* **101**, pp. 3041–3046
43. B. A. Schodin, T. J. Tsomides, and D. M. Kranz (1996) Correlation between the number of T cell receptors required for T cell activation and TCR-ligand affinity. *Immunity* **5**, pp. 137–146
44. P. Marrack and J. Kappler (2004) Control of T cell viability. *Annu. Rev. Immunol.* **22**, pp. 765–787
45. T. P. Arstila, A. Casrouge, V. Baron, J. Even, J. Kanellopoulos, and P. Kourilsky (1999) Direct estimate of the human $\alpha\beta$ T cell receptor diversity. *Science* **286**, pp. 958–961
46. J. N. Blattman, D. J. D. Sourdive, K. Murali-Krishna, R. Ahmed, and J. D. Altman (2000) Evolution of the T cell repertoire during primary, memory, and recall responses to viral infection. *J. Immunol.* **165**, pp. 6081–6090
47. D. H. Busch and E. G. Pamer (1999) T cell affinity maturation by selective expansion during infection. *J. Exp. Med.* **189**, pp. 701–709
48. G. Kassiotis, R. Zamoyska, and B. Stockinger (2003) Involvement of avidity for major histocompatibility complex in homeostasis of naive and memory T cells. *J. Exp. Med.* **197**, pp. 1007–1016
49. W. C. Kieper, J. T. Burghardt, and C. D. Surh (2004) A role for TCR affinity in regulating naive T cell homeostasis. *J. Immunol.* **172**, pp. 40–44
50. D. J. D. Sourdive, K. Murali-Krishna, J. D. Altman, A. J. Zajac, J. K. Whitmire, C. Pannetier, P. Kourilsky, B. Evavold, A. Sette, and R. Ahmed (1998) Conserved T cell receptor repertoire in primary and memory CD8 T cell responses to an accute viral infection. *J. Exp. Med.* **188**, pp. 71–82
51. R. M. Welsh, L. K. Selin, and E. Szomolanyi-Tsuda (2004) Immunological memory to viral infections. *Annu. Rev. Immunol.* **22**, pp. 711–743
52. A. Lanzavecchia (2002) Lack of fair play in the T cell response. *Nature Immunol.* **3**, pp. 9–10
53. R. M. Kedl, B. C. Schaefer, J. W. Kappler, and P. Marrack (2002) T cells down-modulate peptide-MHC complexes on APCs *in vivo*. *Nature Immunol.* **3**, pp. 27–32
54. M. Berard and D. F. Tough (2002) Qualitative differences between naïve and memory T cells. *Immunology* **106**, pp. 127–138
55. C. Kesmir, J. A. M. Borghans, and R. J. de Boer (2000) Diversity of human $\alpha\beta$ T cell receptors. *Science* **288**, p. 1135
56. R. A. K. Singh, J. R. Rodgers, and M. A. Barry (2002) The role of T cell antagonism and original antigenic sin in genetic immunization. *J. Immunol.* **169**, pp. 6779–6786
57. F. Wu, Z.-B. Wang, P. Lu, Z.-L. Xu, W.-Z. Chen, H. Zhu, and C.-B. Jin (2004) Activated anti-tumor immunity in cancer patients after high intensity focused ultrasound ablation. *Ultrasound Med. & Biol.* **30**, pp. 1217–1222
58. D. Raucher and A. Chilkoti (2001) Enhance uptake of a thermally responsive polypeptide by tumor cells in response to its hyperthermia-mediated phase transition. *Cancer Res.* **61**, pp. 7163–7170

Monte-Carlo Methods in Studies of Protein Folding and Evolution

E. Shakhnovich

Department of Chemistry and Chemical Biology, Harvard University, 12 Oxford Street, Cambridge MA 02138, USA
Shakhnovich@chemistry.harvard.edu



Eugene Shakhnovich

E. Shakhnovich: *Monte-Carlo Methods in Studies of Protein Folding and Evolution*, Lect. Notes Phys. **704**, 563–593 (2006)
DOI 10.1007/3-540-35284-8_21
© Springer-Verlag Berlin Heidelberg 2006

1 Protein Design – Practical and Evolutionary Aspects 566

2 Prebiotic Discovery of Protein Folds – Wonderfolds and All That 570

3 From Coarse Grained to All-Atom Studies of Protein Folding 575

3.1 Long-Time Side-Chain and Backbone Dynamics – A Glassy Story 576

3.2 Analyzing Folding Nucleus at Atomic Detail 577

4 Sequence or Structure: Insight from High-Resolution Simulations 580

5 How Dynamically Realistic is MC? – Comparison with Discrete Molecular Dynamics Simulations 581

6 Towards Realistic Transferable Sequence-Based Potentials for Folding Simulations and Protein Design 582

7 Concluding Remarks 585

References 586

As was noted in recent review [1] studies in molecular biophysics (e.g. protein folding and evolution) underwent cyclic development. Initially protein folding was viewed as strictly experimental field belonging to realm of biochemistry where each protein is viewed as a unique system that requires its own detailed characterization – akin to any mechanism in biology. Introduction, in early nineties, of simplified models to the field and their success in explaining several key aspects of protein folding such as two-state folding of many proteins, nucleation mechanism and its relation to native state topology, has pretty much shifted thinking towards views motivated by physics. The "physics"-centered approach focuses on statistical mechanical aspect of folding problem by emphasizing universality of folding scenarios over uniqueness of folding pathways for each protein. This approach dominated theoretical thinking in the last decade (reviewed in [1–4] and its successes brought theory and experiment closer together transforming the protein folding field from branch of biochemistry to a truly interdisciplinary enterprise where physics, chemistry and biology meet. An important contribution to success of statistical-mechanical paradigm in protein folding was the use of stochastic methods, most notably Monte-Carlo. Indeed it was understood early on (since Levinthal formulated his famous argument about protein folding) that significant aspect of protein folding involves massive search in conformational space and stochastic methods would be a natural choice to achieve that goal.

Further indication of the key role of stochastic search methods in biophysics came with the discovery of specific requirements for sequence selection of foldable proteins [5–7]. That called for development of efficient methods to search sequence space for sequences that can successfully fold into desired structure. Theory predicted that such sequences should have special features in their free energy landscape – a large (intensive in number of aminoacid residues) gap between native state (that should be the lowest energy conformation) and lowest energy misfolded conformation that structurally dissimilar to the native state [8, 9]. The fraction of such sequences is exponentially small [10] so that search for them would (naively) constitute a "Levinhal paradox" in sequence space. Stochastic search – Monte-Carlo in sequence space appears to be an efficient method both for model proteins [7, 10] and, as was shown more recently by Kuhlman and coauthors for real proteins [11].

Finally there is another dimension to search problems in molecular biophysics – the question of how nature searches (and finds) protein structures, i.e. how protein universe evolved. Here stochastic modeling of evolutionary processes involving simultaneous sequence and structure changes promises new and exciting advances. In this review we will discuss earlier and more recent work where Monte-Carlo methods played decisive role in studies of conformational and evolutionary changes in biological macromolecules.

1 Protein Design – Practical and Evolutionary Aspects

The idea to select folding (large gap) sequences from canonical ensemble [6] immediately suggests a *practical* approach to find such sequences. Indeed any stochastic search in sequence space that converges to canonical distribution will do the job. Such method was first suggested in [5, 12] – Monte Carlo in sequence space. One issue that needs to be addressed in such search is that it can converge to homopolymeric sequences composed of residues that attract each other most strongly. Indeed such solution will certainly lead to low energy in the native conformation, but it is flawed. The reason is that in fact energy gap between the native state and set of misfolds need to be maximized, not just energy of the native state. The simplest solution to that problem was proposed in [5]: to run stochastic Monte-Carlo search in sequence space to minimize energy of the native state *under constraint of constant aminoacid composition*. This idea appeared successful in preventing the convergence to homopolymer sequences providing sequences with optimized energy gaps. The reason why such approach is successful was explained in [5]. The lowest energy conformations in the misfolded set depend primarily on aminoacid composition. At the same time energy of the native conformation for which search in sequence space is carried out depends on sequence. Therefore minimization of energy of the native conformation while keeping aminoacid composition constant provided simplest way to maximize energy gap.

This approach to sequence design while conceptually simplest is perhaps not the optimal because it, by construction, is not able to find also an optimal aminoacid composition. Several improvements were suggested. First, as a proxy of energy gap, Z-score in the native conformation [8, 13]:

$$Z(\{\sigma\}) = \frac{E_{NAT}(\{\sigma\}) - E_{av}(\{\sigma\})}{D_E(\{\sigma\})} \tag{1}$$

can be optimized in sequence space. Here $E_{NAT}(\{\sigma\})$ is energy of sequence $\{\sigma\}$ in the native ("target") conformation, $E_{av}(\{\sigma\})$ and $D_E(\{\sigma\})$ are average energy and its dispersion (over all M conformations) of sequence $\{\sigma\}$:

$$E_{av} = \frac{\sum\limits_{conf} E(\{\sigma\}, conf)}{M}; \quad D_E = \frac{\left(\sum\limits_{conf} (E(\{\sigma\}, conf) - E_{av})^2\right)^{1/2}}{M^{1/2}}$$

Apparently homopolymeric solutions do not optimize Z-score - rather $Z = 0$ for homopolymers because in this case $E_{NAT} = E_{av}$. Z-score optimization of sequences was first developed in [14] for lattice model proteins and was further extended to real proteins in [15].

Another approach to design optimal sequences was proposed in [16] where sequences $\{\sigma\}$ that maximize Boltzmann probability to be in the native state at a given temperature T:

$$p_{NAT}(T) = \frac{e^{-E_{NAT}\{\sigma\}/kT}}{\sum\limits_{conf} e^{-E(\{\sigma\}, conf)/kT}} \tag{2}$$

are sought.

Exact evaluation of sum over all conformation in the partition function in the denominator of (1.10) is not feasible. Instead an approximation based on cumulant expansion of partition function was used in [16]. This approach provides a possibility to design proteins with selected thermal properties – from mesophilic to hypethermophilic proteins. It also accounts for free energy difference between folded and unfolded states (the latter is accounted for via estimate of partition function).

Further developments of stochastic Monte-Carlo sequence design procedures followed two tracks. First, it was applied to design of model lattice proteins in [7] and real proteins (with extension to all atom model of a protein and significant development of force-fields to realistically represent protein energetics) by Kuhlman and Baker [11,17,18] and by Mayo and coworkers [19]. In particular, Kuhlman and Baker were able to design a sequence that folds into new fold [11]. This remarkable result provides strong support to the basic theoretical result that low energy in the native state is necessary and sufficient for sequence to be protein-like and foldable. Earlier this concept was proven in simulations [7] where sequence was designed to have large energy gap in an arbitrarily selected target and was shown to fold into that target (see Fig. 1).

The success of stochastic energy-based design procedures in providing foldable model and real protein sequences provided strong vindication of the energy gap and related concepts of protein thermodynamics.

The second direction of development and application of stochastic sequence selection methods is to consider them as simplest models of natural evolution. Along these lines two important sets of results were obtained. First, one can further exploit the analogy between statistics in sequence space and canonical ensemble to estimate the number of sequences that can fold into a given protein structure [10,12]. That was done for several proteins in [10] and for many more (using a different sequence sampling strategy and analysis) in [20]. Of particular interest are differences in the capacity of protein structures to accommodate different numbers of stable sequences, i.e. in their designabilities [21–24]. It was suggested for simple models [21,24] and made clear for real proteins [20] that different protein structures may have vastly different designabilities. Then the question is what is the structural determinants of protein designability? The theory developed by England and Shakhnovich [25] suggested that a particular property of a protein structure, namely traces of higher powers of its contact matrix (or equivalently, maximum eigenvalue of its contact matrix) may serve as a good predictor of protein designability. This prediction was tested on lattice model 27-mers whose conformations could be exhaustively enumerated (there are total 103346 of them unrelated by symmetry [26]). The structure with highest and lowest maximum eigenvalues of their contact matrices can be selected and their designabilities can be then directly compared by calculating $S(E)$ – (log) of the number of sequences that can fold into a given structure with energy E. This quantity can be calculated using analogy between statistics of sequences and statistical

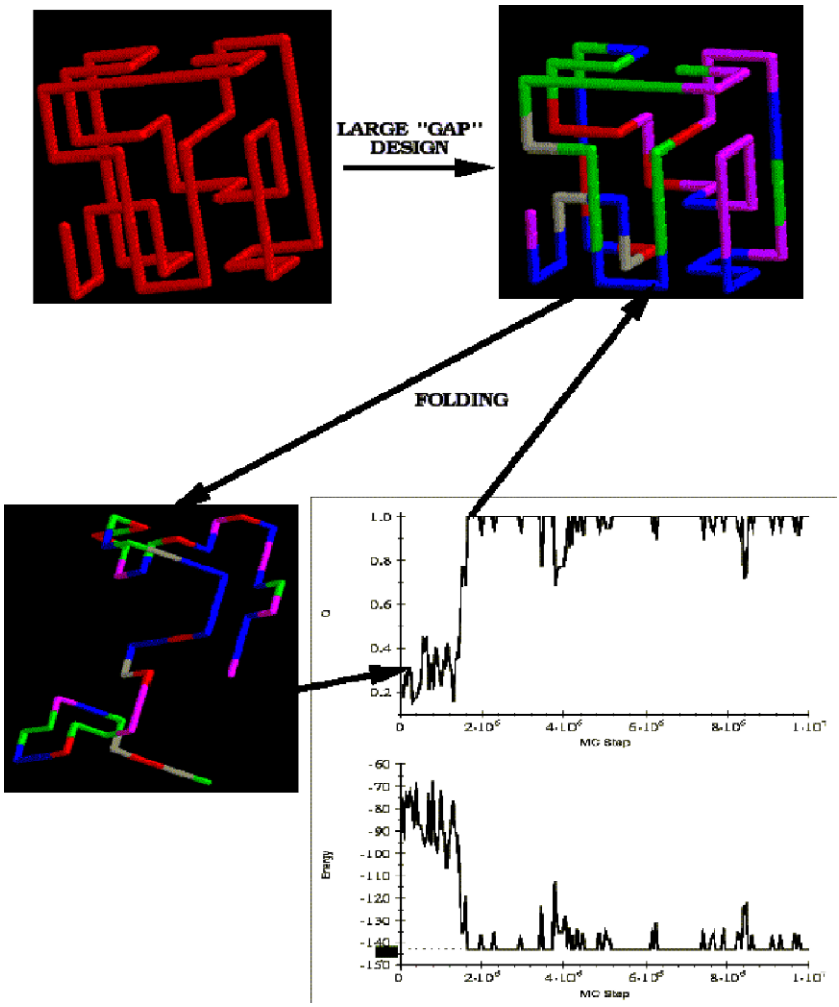


Fig. 1. Computational experiment showing that sequences designed with large energy gap fold cooperatively and rapidly into their native conformations [7]. First, a structure is chosen to serve as target, native conformation. Then sequences are designed (using Monte-Carlo search in sequence space with fixed composition) to have large energy difference (gap) between native conformation and set of structurally distinct misfolds. One of such sequences is memorized. Monte-Carlo folding simulations for this sequence start from an arbitrary random coil conformation and quickly and cooperatively converge to the target conformation for which the sequence was designed. The designed sequence has target conformation as its apparent global energy minimum as no conformations with energy lower than that of the target (native) conformation are found

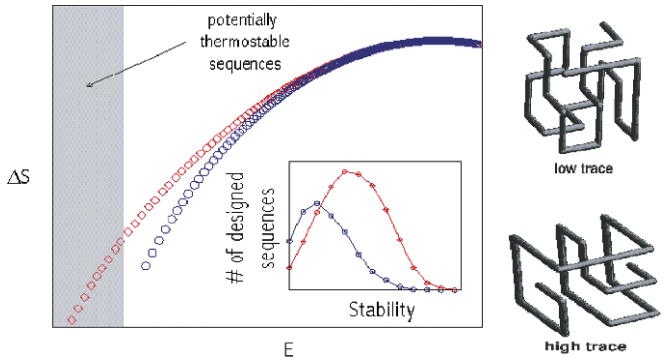


Fig. 2. Two lattice structures – having highest and lowest predicted (by traces of their contact matrices) designabilities – and statistics of sequences that can fold into these structures with given energy. ΔS is entropy (log) of the number of sequences that fold into a given structure with a given energy counted from fully unconstrained statistics (at $E = 0$). Dark grey points describe entropy of sequences designed for the low trace structure and light grey points are for high trace structure. The insert shows how many sequences can be stable (i.e. have high Boltzmann probability) in less and more designable structures respectively

mechanics of canonical ensemble (8) [10]. The comparison shown in Fig. 2 indeed indicates that structures that have higher maximal eigenvalue of their contact matrices (or, similarly, higher traces of powers of contact matrices) are indeed more designable: more sequences exist that can fold into such highly designable structures with low energy.

The analysis of sequence entropy curves presented in Fig. 2 points out to another interesting feature – that it is easier to find thermostable sequences for more designable structures than for less designable ones. Indeed sequences that have exceptionally low energy in their native states can be found only for more designable structures – light grey curve on Fig. 2 ends at higher energy than dark grey curve. This observation suggests possible direct implication for structural genomics: that proteomes from more thermostable organisms will be statistically enriched with more designable structures. The analysis confirmed this conjecture [27]. This finding is very important as it provides direct connection between theory of protein folding, structural genomics (proteomics) and evolution. Further connection between protein evolution and designability is revealed in comparison between gene families of different sizes. The idea that designability may affect size of gene families (so that more designable proteins can accommodate more sequences i.e. have gene families of greater size) was proposed by several researchers [21, 24, 28]. However in the absence of structural determinant of protein designability such proposals were hard to evaluate. Now, when structural determinant of protein designability is better understood direct test of the hypothesis that designability effects the size of gene family was carried out [29]. It was found in [29] that there

indeed exists statistically significant correlation between size of gene family and designability of protein structure that it encodes. However this correlation is limited because other factors such as evolutionary history affect the size of a gene family [30]. Indeed when factor of age of gene family is taken into account the correlation between designability and size of a gene family becomes more pronounced. Further it was found that more ancient proteins – i.e. the ones that are shared by all kingdoms of life – are significantly more designable. This suggests that evolution progressed towards discovery of less designable proteins. This finding can be explained by the observation that as evolution progressed search in sequence space was facilitated. This relaxed restrictions on structures for which viable sequences could be found. This trend is consistent with observations from simulations of evolution in lattice models [31].

2 Prebiotic Discovery of Protein Folds – Wonderfolds and All That

The concept of protein designability was introduced in nineties [21, 28] as an attempt to shed light on observed unequal distribution of protein folds. Revealing the cause of such unequal distribution is a key to our understanding of protein evolution from structural perspective. Two fundamentally distinct explanations can be hypothesized. First, it is possible that modern highly populated folds were selected by chance early in evolution and their dominance was preserved in the process of divergent evolution that resulted in modern protein universe [32]. The second possibility is that most abundant folds have certain intrinsic advantage, i.e. they emerged as a result of some form of purifying selection. Finkelstein and coworkers suggested that folds corresponding to populated superfamilies are highly designable, i.e. they can accommodate a large number of sequences [33]. Subsequent study by Li and coworkers [21, 34] demonstrated, using an exact protein model of 27-mer heteropolymer on the cubic lattice [35] that indeed some structures can adopt more sequences than others, i.e. they appear to be more designable. Later, it was discovered that sequences with highly designable native structures possess increased stability as compared to the rest of the sequence/structure pairs [34]. Several models of selection of more designable structures have been proposed to date. For example, Taverna and Goldstein showed that highly designable structures are even more populated if one considers not all of the sequences, but only those providing high "foldability" [30]. In this work, a power law dependence between designability of a structure and its occupancy by selected sequences has been found.

Despite these theoretical advances, their applicability to real proteins was uncertain since no transferable and reliable structural determinant of designability was found at that time. As described above in more detail, England and Shakhnovich [25] found the structural determinant of protein

designability - contact traces (traces of powers of a protein's contact matrix), well-approximated by powers of maximal eigenvalue λ_{max} of the contact matrices, - using an analogy between protein design and statistical mechanics of spin models. This property is directly transferable to real proteins since it is straightforward to calculate their contact matrices. Indeed, in recent study a statistically significant correlation between the structural determinant of designability and the size of gene families was found [29].

Recent analysis (Zeldovich, Berezovsky, ES, unpublished) shows that the requirement to maintain high thermostability results in emergence of superfamilies of most designable folds. This finding provides the missing link between modern physical views on protein structure fitness (designability) and possible prebiotic mechanisms that gave rise to uneven fold usage and emergence of superfamilies *en route* of protein evolution. Moreover, we show that the dependence of lattice fold usage on the structural determinant of designability quantitatively reproduces uneven fold usage in natural proteins.

This approach is based on the 27-mer lattice model of protein whose all compact conformations were enumerated [35], providing the basis for exact statistical-mechanical analysis. It presents the simplest model of convergent *prebiotic* evolution whereby simultaneous search in sequence and structure spaces are carried out to find stable model proteins. This approach is an extension of the earlier works [5, 12, 16, 36] discussed in detail in the previous chapter, where a design procedure based on stochastic optimization in sequence space with *fixed* target structure has been developed. Here we no longer require that the structure is fixed: a new native conformation (the one of lowest energy among all 103346 compact folds) is determined after each attempted "move" in sequence space (i.e. after each mutation). This crucial extension makes it possible to explore the issue of whether there are preferred native conformations of stable proteins and if so, what are their properties.

In this microscopic model of physical selection the search starts from proteins with random sequences and results in convergent repetitive discovery of proteins having special specific structures, "wonderfolds". Multiple independent runs of sequence design algorithm yield many unrelated sequences folding into the same structure, i.e. prototypic protein superfamilies. In contemporary proteins, superfamilies [37] are the sets of proteins with apparently non-homologous sequences (i.e. undetectable by sequence alignment programs such as BLAST) yet similar structures [15].

20000 runs of the sequence/structure design procedure starting each run from a randomly generated sequence are performed and the native structures for each of the 20000 evolved sequences are determined. Figure 3 shows the probability that a structure is a ground state for k sequences out the total of 20000 evolved sequences. For comparison the same plot for random sequences is shown in Fig. 3.

It turns out that some structures can adopt up to 19 evolved sequences (light grey curve), while no cases were observed when more than 5 random sequences had the same ground state structure (black curve). For random

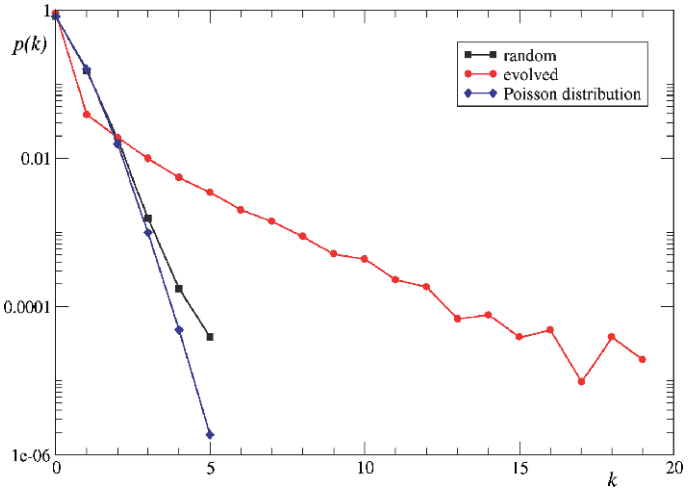


Fig. 3. Probability $p(k)$ of finding k sequences that fold into the same structure from a pool of 20000 sequences. For random sequences (*black curve*), no structures populated by more than 5 sequences are observed. On the contrary, evolved sequences (*light grey curve*) tend to crowd specific structures, with up to 19 sequences in each of the most popular structures, and the distribution dramatically deviates from the null model of random population of folds in the form of Poisson distribution (*dark grey curve*). The value of $p(19)$ for the Poisson distribution equals to $2.5 \cdot 10^{-14}$

sequences, this probability is approximately given by Poisson distribution

$$p(k) = e^{-\lambda}\lambda^k/k!$$

with $\lambda = 20000/103346$ (dark gray curve). At high k, the probability for selected structure/sequence pairs exponentially decreases with k, at a much slower rate than the Poisson distribution. Therefore, this procedure of sequence design efficiently uncovers special structures, wonderfolds, that serve as native structures to unusually large numbers of independently discovered evolved sequences. This result is in agreement with an earlier finding by Taverna and Goldstein [30], who demonstrated that selection for "foldability" (defined as the difference of the energy of the native state with respect to the average energy of all other states) biases the distribution of selected lattice proteins towards more designable structures.

To reveal a quantitative structural characteristic of advantageous folds, we considered the maximum eigenvalue (λ_{max}) of a structure's contact matrix, which serves as the structural determinant of its designability [25]. The scatter plot for the number of evolved sequences k that a structure adopted in the process of our sequence selection versus λ_{max} of its contact matrix is shown in Fig. 4a. It is important to note that λ_{max} is the only reliable predictor of structure's designability discovered so far. For example, there is no correlation between the number of evolved sequences and the contact order [38] of their

ground state structures reveals a very strong positive correlation between the λ_{max} of a structure and the maximum number of sequences found in the process of selection that fold into that structure. This result establishes a crucial connection between designability and formation of superfamilies: wonderfolds, being the most designable structures, form superfamilies in the process of sequence selection that favors highly stable proteins. It also shows an interesting interplay between statistical factors and structure-related bias towards wonderfolds: While low-k structures, those adopting only few sequences, can span a broad range of λ_{max}, the high-k structures, forming highly populated superfamilies are exclusively wonderfodls with high λ_{max}. Due to this interplay between selection and chance, the scatter plots between designability of a fold (its λ_{max}) and its gene family size (k) may be rather broad as suggested by Fig. 4a. However, if one bins the data in λ_{max} bins and evaluates correlation between the logarithm of average k in a bin and λ_{max} it would be extremely strong, with correlation coefficient $R = 0.98$ (Fig. 4b). Thus, selection leads to a very steep exponential dependence of the average structure occupancy k on the designability determinant λ_{max}. If designability and selection pressure were not the factors, the graph in Fig. 4b would be a horizontal line at the constant level of $\ln(20000/103346) = -1.7$. The pronounced slope of the graph is a very clear illustration of the variations in designability, where some (high-λ_{max}) structures are much more populated than others. It also serves as a direct demonstration that λ_{max} is the structural determinant of designability. Exactly the same effect – broad triangle-shaped scatter plot for raw data and very high correlation for binned data – is seen in the analysis of correlation between designability and gene family sizes in real data [29].

To qualitatively compare the designability of evolved lattice proteins with one of natural proteins, we estimated the designability of the structures of the protein domain universe graph (PDUG [32]), where protein domains are clustered according to their structural similarity. Figure 6 shows the dependence of the PDUG cluster size on the average λ_{max} of the structures in this cluster. The exponential increase of cluster size with λ_{max} in the lattice model (Fig. 4b) is in quantitative agreement with the data for natural proteins (Fig. 5). Thus, we have demonstrated how selection for stability gave rise to the very specific, exponential, dependence of the number of protein fold superfamilies on the structural determinant (λ_{max}) of designability of the folds.

These Monte-Carlo sequence/structure design calculations explicitly and rigorously demonstrated in an a priori way the emergence of a small number of structural patterns that are inherently more favorable with respect to the physical requirement of thermostability regardless of initial sequence from which sequence selection starts. Using a lattice model, we revealed here that native state energy optimization naturally leads to the convergence towards a limited number of conformations with very specific structural properties, separating them from the bulk of compact 27-mers. Moreover, the exponential dependence between the determinant of designability, gλ_{max} and occupancy of the lattice structures is recovered in real proteins.

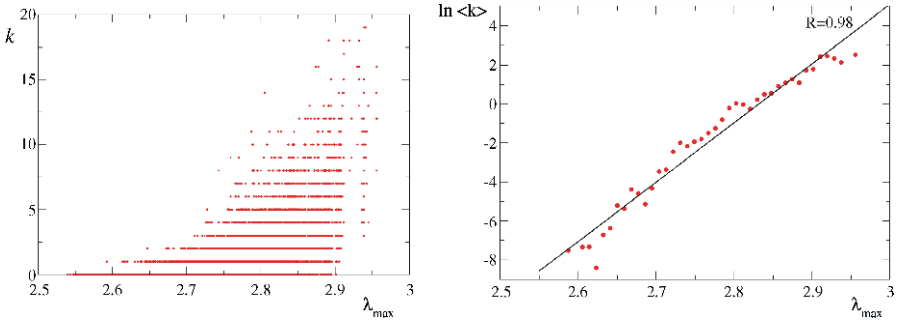


Fig. 4. (**a**) Dependence of the number of evolved sequences corresponding to a structure on the maximum eigenvalue (λ_{max}) of the structure's contact matrix. One can clearly see that only most designable structures – wonderfolds – form large superfamilies containing many (up to 19) dissimilar sequences. A notable property of compact 27-mers is the gap in the distribution of the eigenvalues at $\lambda_{max} \approx 2.92$, which naturally separates wonderfolds from the bulk of the structures. (**b**) Dependence of the logarithm of the average number of evolved sequences k within a bin $\lambda_{max,i} < \lambda_{max} < \lambda_{max,i} + \Delta\lambda$, with $\lambda_{max,i}$ corresponding to 50 bins spanning the range from 2.52 to 2.96. All 103346 compact structures (including those that were not found in evolution simulations, i.e. having $k = 0$) are considered here. A linear approximation (*black line*) is shown, correlation coefficient $R = 0.98$

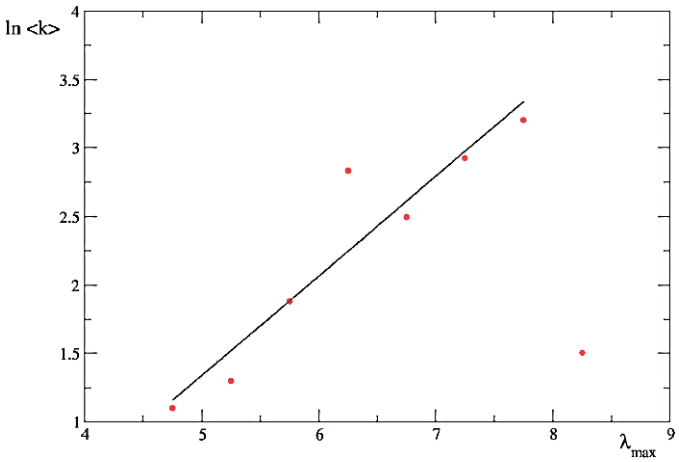


Fig. 5. Dependence of the logarithm of the average size of the PDUG cluster k within a bin $\lambda_{max,i} < \lambda_{max} < \lambda_{max,i} + \Delta\lambda$, with $\Delta\lambda = 0.50$. A linear approximation (*black line*) is shown, correlation coefficient $R = 0.95$

Native energy optimization in these simulations, where aminoacid composition is fixed in each run, is apparently equivalent to optimization of stability: Monte-Carlo search for model proteins sequences having the lowest possible ground state energy, provided evolved model proteins have a remarkably high thermostability. This can represent a good model for prebiotic selection of thermostable structures that can result in emergence of superfamilies around preferred most designable folds. Indeed, it has been convincingly demonstrated that proteomes from ancient archaea exhibit clear structural bias towards more designable folds [39] and, in particular, proteins from Last Universal Common Ancestor (LUCA) were found to be significantly more designable, than other, later evolved proteins [29].

3 From Coarse Grained to All-Atom Studies of Protein Folding

Studies of simple models indeed contributed considerably to our understanding of protein folding by emphasizing universal aspects of protein folding thermodynamics and kinetics. That helped to focus experimental and theoretical studies on key common milestones along protein folding pathways such as transition states (and its common structural features such as folding nucleus [40–44]), on- and off-pathway intermediates [45–48]. Importantly that many of the experimental studies were directly motivated by specific predictions and questions raised in theoretical studies. For example the analysis of nucleation in lattice models and in bioinformatics studies of protein families [40,49] as well as studies of Go-models of proteins [50–53] suggested that location of nucleus may be robust within families of homologous and analogous proteins as well with respect to changes in environment. A number of experimental studies were undertaken to test and further explore these insights from theory [54, 55]. Their results appeared to be in good agreement with theoretical predictions.

These and several other successes of theoretical studies of simple universal models advanced protein folding field drawing theory closer to experiment and opening direct channel of communication between theorists and experimentalists in the field.

However simplified models, while providing significant conceptual advances and insights are not able to capture man specific features of protein folding that are accessible to experimental studies. To address this limitation an all-atom Monte-Carlo approach to fold proteins has been developed and tested on a series of proteins using, first, Go model [52, 56, 57] and more recently full-sequence based model that does not require knowledge of native state of a protein [58].

3.1 Long-Time Side-Chain and Backbone Dynamics – A Glassy Story

The developed all-atom simulations tool made it possible to address several problems that a previous coarse-grained models were not able to approach due to their (over)simplified character. One of them is the issue of statistics and dynamics of side-chain packing – an aspect of protein folding that was recognized by many as being at its cornerstone [59,60]. The all-atom MC simulations proved useful in addressing this problem. First, a direct sampling of side-chain packing states was performed to resolve a long-standing issue [59]: how many side-chain packing arrangements are sterically compatible with a given backbone conformation? The analysis was performed for several models of sterics – from hard-shell to van-der-Waals soft-shell steric interactions with unexpected conclusion – that many (exponential in the number of side-chain degrees of freedom) conformations are compatible with a given backbone conformation [61]. Naturally this degeneracy is broken in real proteins by interactions so that native conformation of side-chains is energetically favored over alternatives (decoys). The side-chain packing decoys generated by this algorithm are used to develop atom-atom potentials for protein folding using potential optimization techniques [8, 62–64].

The large conformation space of packed side-chains suggests that there may be a peculiar dynamics of their packing during folding. Again all-atom folding simulations proved an invaluable tool to address this difficult question – the analysis of many individual trajectories for protein G folding allows to develop a very detailed picture of how side-chains get organized in folding process and the results are quite interesting. It appears that there is a broad distribution in time-scales for side-chain packing times even in an apparently two-state kinetics but side-chains that constitute the nucleus are the fastest to acquire their native conformation! [65]. This result was obtained in [65] in simulations of a new lattice model with side-chains as well as in analysis of trajectories of all-atom simulations of protein G.

Further analysis of protein G folding trajectories revealed a complex folding scenario whereby major features of protein topology and packing of nucleus side-chains get established first concurrently with nucleation while side-chain packing of the rest of the structure occurs over longer time scale and is accompanied with backbone fluctuations (see Fig. 6).

These longer time-scale fluctuations appear to be of peculiar character resembling glass transition dynamics with its signature power law relaxation of many characteristics such as total energy. The detailed analysis of then requires a new theoretical approach based on mode-coupling theory [66–68]. A general theoretical formalism based on mode-coupling theory applicable to homo- and heteropolymer dynamics has been developed in [68]. It was shown there that indeed in the low temperature regime a glass transition may occur that would feature a long-time non-exponential relaxation of energy. However

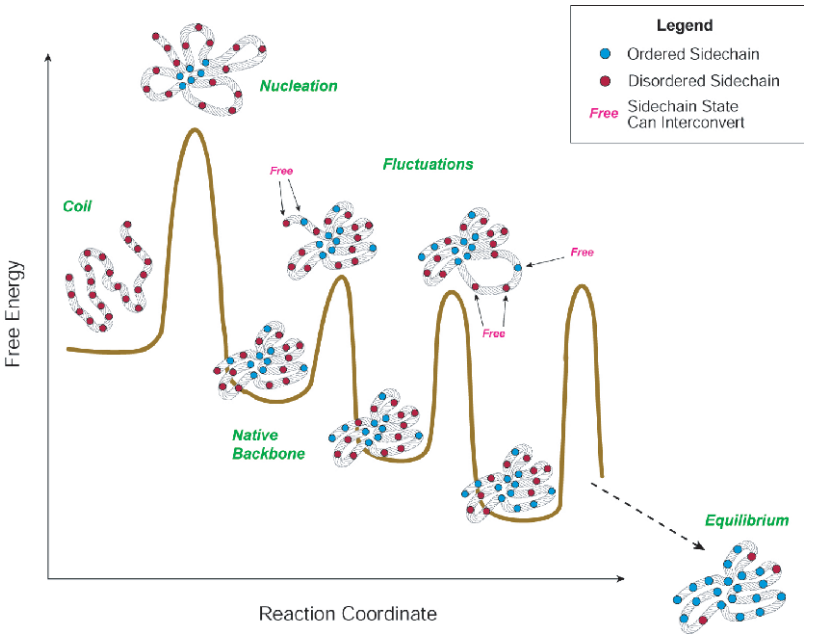


Fig. 6. A Schematic representation of full dynamic process of folding that includes side-chain organization. The main nucleation barrier is overcome first and it leads to establishment of the overall fold. Subsequent dynamics includes local fluctuations of the backbone accompanied by progressive freezing of side-chains. Barrier heights are shown for illustrative purposes only and may be not representative of real situation

this is only the initial step – a comprehensive theory that would treat directly side chain relaxation in proteins is a matter of future development.

3.2 Analyzing Folding Nucleus at Atomic Detail

Development of all-atom MC and MD simulations allowed researchers to overcome a major shortcoming of present experimental studies of protein folding – their relative disconnect from detailed structural interpretation. Indeed even the most structurally detailed protein engineering method pioneered by Fersht and coworkers [69] was sometimes "visually" interpreted as "high ϕ-value residues belong to the nucleus, while low ϕ-value ones do not". Such reasoning being qualitatively acceptable in some cases, often misleads, e.g. I76 in CI2 shows low ϕ-value in many mutations [42], however a careful double-mutant study attributes it to folding nucleus [70]. In another example, of protein G highest f-values are observed in the turn of second hairpin, while f-values in other locations are markedly low [71]. While this observation points out to importance of the hairpin it is hard to imagine a TSE (i.e. set of conformations with $p_{fold} = 0.5$) where only one hairpin is folded while the rest of the protein is not.

The qualitative character of the "visual" interpretation of protein engineering method was noted by Fersht and Daggett [72] who insightfully pointed out that ϕ-values should be treated as experimental constraints akin to NOESY in NMR determination of protein structure. This idea was further developed by Vendruscolo [73, 74] and coworkers who used ϕ-values to reconstruct *putative* transition state of acylphosphatase – one of the proteins studied by Dobson and coworkers using protein engineering methods. These authors reconstructed putative TSE for the protein using ϕt values as constraints in high-temperature unfolding simulations (using initially reduced C_α model [74] and later all-atom representation of proteins [73]). However they did not test whether the proposed conformations represent true TSE, i.e. set of conformations for which transmission coefficient to the folded state $p_{fold} = 0.5$ [75]. All-atom simulations provide a unique opportunity to address this issue. First we carried out the analysis of TSE for CI2 [76] – perhaps the best characterized protein in terms of ϕ-value analysis [42]. We showed there that ϕ-values correctly specify, in general the TSE as $\langle p_{fold} \rangle$ over putative TSE ensemble appeared to be close to 0.5. The work presented in [76] was like a "proof of principle" both for p_{fold} calculations and ϕ-value analysis. Our subsequent study [77] presented a much more detailed picture of TSE for protein G folding. In particular it clarified a number of key issues related to the ϕ-value analysis: a) What is the minimal number of ϕ-value constraints to enable reliable reconstruction of TSE? b) What is the relation between ϕ-values of residues reported in various mutations and its role in forming the TSE? The analysis of protein G provides answers to these questions for that protein. In particular it was shown that upon gradual addition of ϕ-value constraints the $\langle p_{fold} \rangle$ ($\langle\rangle$ means averaging over many starting conformations generated using constraints) first grows and then saturates reaching limiting values of 0.5. Most importantly, distribution of p_{fold} value over constraint-generated starting conformations is pronouncedly bimodal: many conformations are found with low and high p_{fold} and relatively few in between, with $p_{fold} = 0.5$. This is perhaps not surprising because TSE corresponds to free energy maximum, i.e. it is comprised of least stable conformations (see Fig. 7). However this simple observation clearly indicates that no structural characterization of TSE without p_{fold} analysis is possible. In particular the models of transition states based only on constraints may be highly misleading. For example an unverified model of TSE for SH3 domains (based on constraints only) posits that TSE ensemble for these proteins has native-like topology and is structurally close to native state for all three SH3 domains studied [78]. However careful analysis of SH3 TSE that includes p_{fold} verification presents a completely different picture: of highly polarized TSE with well-defined small nucleus but with significant part of the chain disordered almost as much as in unfolded state [57, 79]. It should be noted that folding nucleus in SH3 domains (as well as in other studied proteins [77, 80]) is "diffuse" *in sequence:* it is comprised of residues that are uniformly distributed throughout the sequence. However the residues belonging to folding nucleus are well packed in *space* in TSE

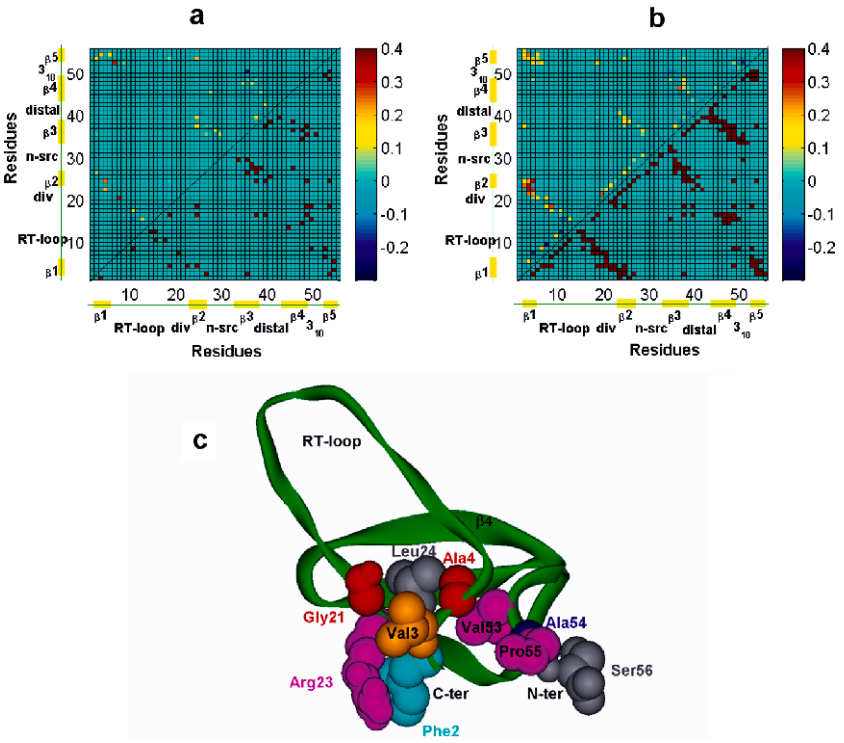


Fig. 7. Differential contact maps between pre-TS ensemble and TSE for src SH3 domain folding [53] (*upper panels*). (**a**) – for contacts between geometric centers of side chains, (**b**) – for contacts between C_β atoms. Lower panels on both contact maps correspond to native structure the SH3 domain. (**c**) – Cartoon diagram of a sample TS structure determined by p_{fold} analysis. Residues with contact probably change from pre-TSE to TSE (as shown on upper panel of (b) greater than 0.1 are shown in space-filling scheme. They constitute a polarized folding nucleus for this domain. The figure and analysis are from [79]

conformations. This very clear from the pfold-based analysis of contact maps in pre-TS ($p_{fold} < 0.5$) conformational ensemble, TSE ($p_{fold} = 0.5$) and post-TS ($p_{fold} > 0.5$) conformational ensemble. Contact maps are constructed to show contacts that are most probable in corresponding ensembles. Of special importance are differential contact maps between TSE and pre-TS ensembles. Apparently such differential contact map shows only contacts that are most important for TSE: without them the TSE is not reached. These are the contacts that are necessary to form in folding TSE, i.e. *nucleation contacts*. The analysis of nucleation in SH3 domains reveals important necessary structural features. It is primarily central β-sheet consisting of strands 2-4. It is a necessary feature because it is always present in all TSE conformations. However it is not sufficient to form this better strand to reach TSE. Indeed

the same b-sheet is formed in pre-TS ensemble. In other words formation of the central β-sheet is very important but it does not guarantee that TSE is reached. What does? The answer to this question comes from the analysis of differential contact map between pre-TS ensemble and TSE that points out to contacts that are key to TSE i.e. that appear only in TSE but not before. The analysis of differential contact maps revealed that a few key contacts are crucial for TSE. These contacts are between residues that are spread all over the sequence but form tight cluster in structure. These residues constitute *folding nucleus* for SH3 folding: its formation is key to reaching the TSE. Also this set of contacts, corresponding to folding nucleus corresponds to common, invariant feature among all TSE conformations.

4 Sequence or Structure: Insight from High-Resolution Simulations

One of the most debated issues in protein folding is what determines folding pathways: final structure or protein sequence. While this question may sound a somewhat scholastic (since sequence always determines final structure) it is not: there are many proteins that have similar structure but very different sequences and the relevant question is whether such proteins have similar or different folding mechanisms. This question has a long history. First observation that structure may be a more robust determinant of the folding mechanism than sequence was made in [40]. This proposal was based on lattice model study which showed that nonhomologous sequences designed to fold into the same native structure of a model lattice protein (see Fig. 2) have similar folding nucleus. Subsequently several authors arrived at similar conclusions [81,82]. The evolutionary implications of the primacy of native topology in defining folding mechanism was noted in [49,83–85]. In particular if folded structure is determinant of the mechanism then same set of residues in structurally aligned proteins with non-homologous sequences may play important (thermodynamic and kinetic) role. Such residues may be under evolutionary pressure due to folding requirement which leads to their *universal conservation*. Universal conservation of some residues means that residues occupying same positions in protein structures appear to be conserved in their respective sequence families though the aminoacid types of these residues may vary from family to family. Such remarkable universal conservation, despite diversity of functions between families, was indeed observed for many populated folds [49,85,86]. The observation that location of folding nucleus in structure is robust, i.e. proteins with same structure but different sequences have the same location of folding nucleus implies that folding nucleus residues should be universally conserved if evolution exerts additional pressure to control folding rates. Experimentally nucleus residues were indeed found to be universally conserved in several cases [54, 87]. In particular, for Ig domains studied by

Clarke and coworkers [87] folding nucleus was predicted, based on universally conserved residue analysis prior to experiments [49]. However in some cases the apparent exceptions from the perceived robustness of folding pathway were found. For example, in small helical protein Im7 mutations changed observed pathway – from apparent two-state to three-state [88] folding mechanism. Similarly Baker and coauthors showed that structurally similar proteins G and L have different distribution of ϕ-values [89] suggesting that these two proteins may have different folding pathways. However a detailed analysis based on simulations of protein G in structure-centric Go model [52] showed that certain features of folding pathway are flexible and certain features are robust. In particular, there may be many pathways leading to nucleus formation passing through various metastable intermediates. This aspect is flexible as mutations can easily shift distribution between different paths and stability of the intermediates. However all these pathways converge to a single nucleation step and the structure of the nucleus is robust in a sense that it is mostly determined by final structure of the protein. Proteins having different sequences but similar structures have very similar folding nuclei. This conclusion is supported by experimental studies, e.g. Radford and coworkers showed that despite the fact that two homologous helical proteins – Im7 and Im9 fold via two and three-state mechanism, TSEs structures of these proteins are very similar [90]. The apparent discrepancy between results for L and G proteins obtained by Baker and coworkers [71, 89] can be attributed to difficulties of derivation of TSE from "visual" inspection of ϕ-values. Indeed, when detailed analysis using p_{fold} was carried out for protein G [77] (using experimental constraints and Go model simulations) its folding nucleus appears to consist of several tightly packed hydrophobic residues (consistent with other proteins such as S6, SH3, CI2 etc.) rather than a β-turn as one would naively expect based on visual inspection of ϕ-values. The locus of the correctly determined nucleus appears invariant between proteins G and L. Similarly, location and composition of folding nucleus is invariant between three SH3 domains (spectrin, src and fyn) as revealed by recent study [57]. Davidson and coauthors [91] suggest that answer to the question "do proteins with similar structures fold via the same pathway?" is ambiguous. However our analysis based on combination of detailed high-resolution computations with experimental data gives less ambiguous answer: that *folding nucleus is a robust feature of protein and its location is determined primarily by proteins final structure.* Other aspects of the folding pathway (e.g. how protein "ascends" to TSE) may be more sensitive to details of sequences and change even upon single mutations.

5 How Dynamically Realistic is MC? – Comparison with Discrete Molecular Dynamics Simulations

An alternative simulation method – Discontinuous Molecular Dynamics – was used in a number of study to explore folding mechanisms in coarse grained

models of folding [50, 92–97]. This method is based on direct propagation of dynamics by solving energy and momentum conservation equations each time protein atoms interact between themselves or with "ghost" solvent particles. Several models were studied within Go model energetic prescription – from generic compact structure [50] to SH3 domain [96, 98] to amyloid-like aggregates [97, 99]. The analysis of these simulations shows that the developed picture of specific nucleation is very robust between models and simulation techniques. Further, a very promising model to study protein aggregation and amyloidosis [97] was developed within the DMD simulations approach. Energetics of this model is based on specific side-chain-like interactions combined with nonspecific backbone hydrogen bonding. It provided an intriguing experimentally testable generic model of amyloid fibril formation. More recently a similar model was used by Wolynes and coauthors for their study of aggregation of SH3 domains with similar conclusions concerning domain swap mechanism [100] of aggregation and very similar structural model of dimers of SH3 molecules – precursors of amyloid fibrils. In a more recent work [101] a sequence-based coarse-grained energetics model (as opposed to structure-based Go model) was developed to fold Trp-Cage miniprotein using DMD simulation technique. The authors of [101] note that success in folding of Trp-Cage miniprotein by this method and by atomistic MD simulations [102, 103] may be attributable to specific features of folding and energetics of this miniprotein and may not necessarily be transferable to other cases.

6 Towards Realistic Transferable Sequence-Based Potentials for Folding Simulations and Protein Design

The all-atom Monte-Carlo algorithm as well as several other efficient all-atom and coarse grained folding dynamics algorithms appear to be valuable tool to study folding dynamics and thermodynamics. However, any folding study has two major components: a) search strategy/dynamic algorithm and b) energy function that should select native structure as global minimum. The energy function used in most of all-atom studies described above is based on Go prescription. This may be a good choice as it indeed guarantees that the native state is global energy minimum. However it requires knowledge of the native structure (or at least NOESY constraints from NMR experiments) and may underestimate energetic contribution and persistence of some non-native contacts. The latter were shown to play a possible role in nucleus formation, as predicted in simulations and bioinfomatics analysis [86] and confirmed in experiment [104].

The next step therefore is to develop atomic sequence-based potentials for all-atom simulations that would not require the knowledge of the native state and that may be transferable between proteins. This task is extremely

challenging as many who work in protein structure prediction and simulations may appreciate. A few avenues can be explored here. Fundamentals of simple knowledge-based approaches using quasichemical approximation of the type suggested by Miyazawa and Jernigan [105] were studied and generalized to atomic level of description [106, 107] by Skolnick and coworkers. In particular these authors addressed a difficult question of what should be considered a *reference state* for such potentials. Reference state refers to the question of statistics of pairwise frequencies in the case when no interactions are present. Another class of approaches are Z-score and related optimization methods [8, 63]. A more recent new approach to design atomic potentials for protein folding was developed in our lab. It is based on selection of atomic potentials to make realistic protein energetics resemble Go-based energetics as much as possible. To this end, in spirit of knowledge-based potentials the interactions often observed in protein structures are deemed more attractive, while non-existent interactions are more repulsive. The form of the new potential (called m-potential) is designed to coinide with Go-potential when derived on one protein and be closest to Go in terms of energetic bias to native state when derived on independent training dataset of protein structures:

$$E_{AB} = \frac{-\mu N_{AB} + (1-\mu)\tilde{N}_{AB}}{\mu N_{AB} + (1-\mu)\tilde{N}_{AB}} \, .$$

where E_{AB} is contact interaction energy between atom types A and B, N_{AB} is the number of AB pairs found in contact and $\tilde{N}_{AB}$ is the number of AB pairs in the database that are not in contact. μ is a parameter that determines the average interaction (repulsion or attraction); it can be chosen to provide uniform and high (10–20%) acceptance rate in Monte Carlo simulation by preventing overly rapid collapse or excessively slow compactisation. A systematic comparison of all methods to derive atomic potentials – quasichemical approximation, μ-potentials and optimization techniques was analyzed in a recent paper [63] based on results of fold recognition in gapless threading and against standard sets of decoys. It appears that all derived potentials show significant degree of consistency in a sense that in all cases the dominant interactions contributing to stabilization of the native fold are the same – interaction between side-chain atoms of aliphatic groups. However in terms of performance (Z-score of the native conformation) μ-potentials perform better than quasichemical potentials and about as good as optimized potentials. This is important given that μ-potentials were derived on an independent dataset of proteins and were not optimized to perform a specific task.

The first application of μ-potential for folding of a small three-helix bundle protein A provided repetitive and systematic folding within 2A RMS from crystal structure [108]. However this was not a fully transferable μ-potential – it was derived using statistics of contacts in native structure of protein A itself. However the potentials derived from different databases seem to correlate well [108] which is an encouraging sign that it may be transferable. A

more stringent test of atomic potentials was done recently in our lab (unpublished results) where potential was developed on a database that did not contain tested proteins or their homologs. The energy function used for this study represented a linear combination of explicit hydrogen bonding potential (well suited to stabilize helical conformation) and μ-potential derived on an independent database of 119 proteins that did not contain tested proteins of their homologs. 84 atom types were considered (same as in [108]). Simulations performed on seven small nonhomologous alpha-helical proteins showed encouraging results providing in 6 out of seven cases folding to less than 4A RMSD structures from the native state [58]. The analysis of simulation results included clustering of structures and observation that largest disjoint cluster – Giant Component contained most native-like conformations. (Fig. 8)

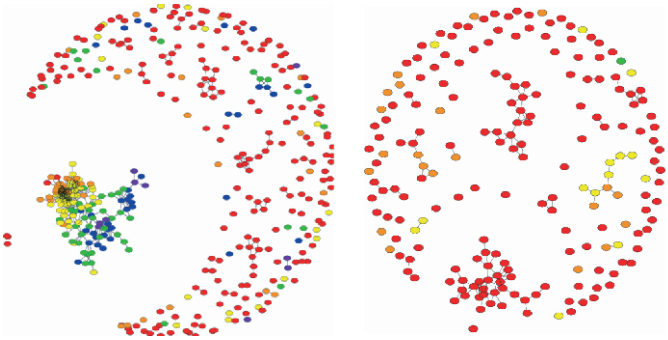


Fig. 8. Clustering of 200 conformations obtained in 200 independent simulation runs of all-atom MC folding algorithm with sequence-based transferable atomic μ-potential for protein A (1BDD). Each node corresponds to the lowest energy conformation obtained in each run and an edge is drawn between any two conformations if RMSD between them is less than 3.5A. Color code indicates RMSD from the native structure: purple: <4A, blue: <5A, green: <6A, yellow < 7A,orange < 8A, red: >8A. The central cluster- giant component – contains all native-like structures, while "peripheral" nodes are mostly misfolds. Figure on the right shows control: clustering of 200 conformations obtained in the same way but for random sequence with the same composition as for 1BDD. Comparison clearly shows that we observe sequence-guided non-trivial folding and that clustering focuses landscape for real sequence towards correct native structure

Various graph-theoretical measures were applied to select "best" prediction and it appeared that most connected conformations – the ones that have most similar conformations – appeared to be statistically closer to the native state. Energy alone was effective but not most effective predictor of the native-like conformations. One possibility, as pointed out by Baker and Shortle [109] the clustering procedure alleviates some inaccuracies that are present with inexact potentials taking advantage of possibly "broader funnel" (whatever it means) surrounding the native structure of the protein rather than infrequent

low-energy decoys. Heteropolymer theory is consistent with that view pointing out that "random decoys" are akin to deep minima in random heteroplymers and represent isolated small sets of conformations on a rugged landscape while native-like structures are less randomly organized [6,110].

Of special interest are control simulations carried out for this study (Hubner et al, unpublished). Simulation of randomized sequence folding resulted in collection of conformations from which native structure of simulated proteins could not be identified by energy or graph-theoretical criteria. However, interestingly some selected conformations were found that exhibited relatively low (4.2 A) RMSD with selected three-helix bundles. This result may reflect on some conclusions from so-called distributed computing approach where many folding simulations are run independently on a grid of computers. There also some conformations were found among many simulations that were close in RMSD to small target protein – villin headpiece [111,112]. However these low RMSD conformations did not appear to be lowest energy ones. A possibility exists therefore that low-RMSD conformations observed in distributed computing simulations are result of random collapse rather than sequence-based energy-guided folding. A similar random control for distributed computing simulations is needed to address this issue.

Another important control concerns the issue of relative importance of pairwise interactions vs. explicit hydrogen bonds in formation of proper protein-like conformations. To this end a number of simulations were performed using energy function in which explicit hydrogen bond term was turned off. Resulting conformations formed almost perfect hydrophobic cores and were as compact as native ones but did not contain any helixes (less than 1% helical content). (Fig. 9)

This result while seems almost obvious is important in light of recent claims that geometrical/topological and generic factors alone (such as excluded volume, topological constraints, compactness) are sufficient to provide protein-like architecture of compact polypeptide globules (modeled as polymers with "finite thickness") [113–115]. These results show that this inference from simple "tube" model is likely to be incorrect. In apparent departure from this view more recently the same authors incorporated explicit hydrogen bond into their model [116] to explain existing protein architectures. However the theory based on the premise that basic physical requirements such as compactness, hydrogen bond formation and loop entropy explain existing protein architectures has been proposed much earlier – in 1987 – by Ptitsyn and Finkelstein [117]. It is not clear therefore what are the novel aspects of the nonspecific "tube" model of protein architecture.

7 Concluding Remarks

Protein folding and design are mature fields where theory and experiment fruitfully interact. While many (but not all) conceptual aspects of protein

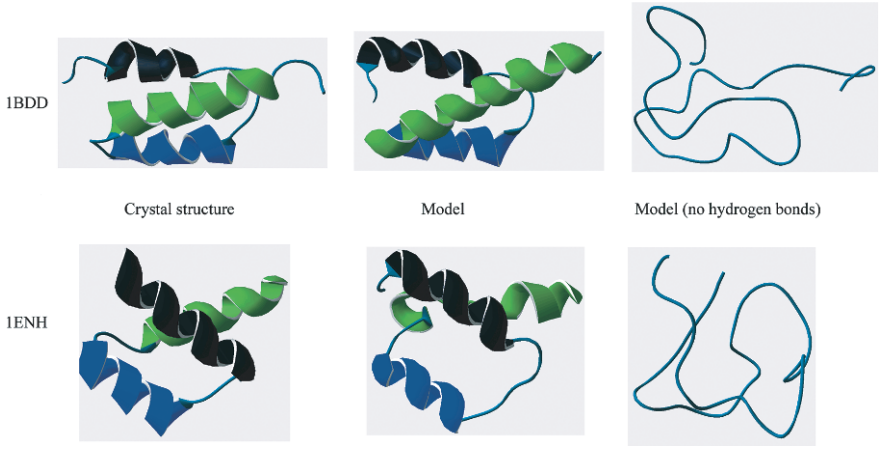


Fig. 9. Protein models from the PDB and representatives from simulation. Model simulations with full energy function (μ-potential pairwise interaction + hydrogen bonding) fold to near native conformations while simulations without hydrogen bonding collapse without helices. Excluded volume and an attractive potential ensure a protein-like hydrophobic core and sidechain packing. However, representation of hydrogen bonding interactions is essential for formation of secondary structure

folding (that used to be centered around so-called "Levinthal paradox") appear well understood and established, there is lot of room for development and further studies. Perhaps in coming years we will see further progress in using predictive atomistic level model combined with effective search strategies such as Monte-Carlo to achieve complete description of folding pathway for several proteins.

Decisive departure from structure-centric (Go) models to sequence based all-atom models that are capable to simulate full folding process from random coil to native ensemble of conformations is an urgent need and an emerging reality. While the consequences of such models for structural genomics are obvious it is equally clear that their study will have significant impact on further understanding of protein folding mechanisms. We are bound to witness decisive successes of theoretical studies of protein folding in coming years.

References

1. L. Mirny and E. Shakhnovich (2001) Protein folding theory: from lattice to all-atom models. *Annu. Rev. Biophys. Biomol. Struct.* **30**, pp. 361–96
2. J. N. Onuchic, Z. Luthey-Schulten, and P. G. Wolynes (1997) Theory of protein folding: the energy landscape perspective. *Annu. Rev. Phys. Chem.* **48**, pp. 545–600

3. S. S. Plotkin and J. N. Onuchic (2002) Understanding protein folding with energy landscape theory. Part II: Quantitative aspects. *Q. Rev. Biophys.* **35**, pp. 205–86
4. J. E. Shea and C. L. Brooks, 3rd. (2001) From folding theories to folding proteins: a review and assessment of simulation studies of protein folding and unfolding. *Annu. Rev. Phys. Chem.* **52**, pp. 499–535
5. E. I. Shakhnovich and A. M. Gutin (1993) Engineering of stable and fast-folding sequences of model proteins. *Proc. Natl. Acad. Sci. USA* **90**, pp. 7195–9
6. S. Ramanathan and E. Shakhnovich (1994) Statistical mechanics of proteins with "evolutionary selected" sequences. *Phys. Rev. E. Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics* **50**, pp. 1303–1312
7. E. I. Shakhnovich (1994) Proteins with selected sequences fold into unique native conformation. *Phys. Rev. Lett.* **72**, pp. 3907–3910
8. R. A. Goldstein, Z. A. Luthey-Schulten, and P. G. Wolynes (1992) Optimal protein-folding codes from spin-glass theory. *Proc. Natl. Acad. Sci. USA* **89**, pp. 4918–22
9. A. Sali, E. Shakhnovich, and M. Karplus (1994) How does a protein fold? *Nature* **369**, pp. 248–51
10. E. I. Shakhnovich (1998) Protein design: a perspective from simple tractable models. *Fold. Des.* **3**, pp. R45–58
11. B. Kuhlman, G. Dantas, G. C. Ireton, G. Varani, B. L. Stoddard, and D. Baker (2003) Design of a novel globular protein fold with atomic-level accuracy. *Science* **302**, pp. 1364–8
12. E. I. Shakhnovich and A. M. Gutin (1993) A new approach to the design of stable proteins. *Protein. Eng.* **6**, pp. 793–800
13. J. U. Bowie, R. Luthy, and D. Eisenberg, (1991) A method to identify protein sequences that fold into a known three-dimensional structure. *Science* **253**, pp. 164–70
14. V. I. Abkevich, A. M. Gutin, and E. I. Shakhnovich (1995) Impact of local and non-local interactions on thermodynamics and kinetics of protein folding. *J. Mol. Biol.* **252**, pp. 460–71
15. N. V. Dokholyan and E. I. Shakhnovich (2001) Understanding hierarchical protein evolution from first principles. *J. Mol. Biol.* **312**, pp. 289–307
16. M. P. Morrissey and E. I. Shakhnovich (1996) Design of proteins with selected thermal properties. *Fold. Des.* **1**, pp. 391–405
17. G. Dantas, B. Kuhlman, D. Callender, M. Wong, and D. Baker (2003) A large scale test of computational protein design: folding and stability of nine completely redesigned globular proteins. *J. Mol. Biol.* **332**, pp. 449–60
18. B. Kuhlman, J. W. O'Neill, D. E. Kim, K. Y. Zhang, and D. Baker (2002) Accurate computer-based design of a new backbone conformation in the second turn of protein L. *J. Mol. Biol.* **315**, pp. 471–7
19. C. A. Voigt, D. B. Gordon, and S. L. Mayo (2000) Trading accuracy for speed: A quantitative comparison of search algorithms in protein sequence design. *J. Mol. Biol.* **299**, pp. 789–803
20. L. Meyerguz, C. Grasso, J. Kleinberg, and R. Elber (2004) Computational analysis of sequence selection mechanisms. *Structure (Camb.)* **12**, pp. 547–57
21. H. Li, R. Helling, C. Tang, and N. Wingreen (1996) Emergence of preferred structures in a simple model of protein folding. *Science* **273**, pp. 666–9
22. H. Li, C. Tang, and N. S. Wingreen (1998) Are protein folds atypical? *Proc. Natl. Acad. Sci. USA* **95**, pp. 4987–90

23. N. E. Buchler and R. A. Goldstein (1999) Effect of alphabet size and foldability requirements on protein structure designability. *Proteins* **34**, pp. 113–24
24. S. Govindarajan and R. A. Goldstein (1996) Why are some proteins structures so common? *Proc. Natl. Acad. Sci. USA* **93**, pp. 3341–5
25. J. L. England and E. I. Shakhnovich (2003) Structural determinant of protein designability. *Phys. Rev. Lett.* **90**, p. 218101
26. E. I. Shakhnovich and A. Gutin (1990) Enumeration of all compact conformations of copolymers with random sequence links. *J. Chem. Phys.* **93**, pp. 5967–5971
27. J. L. England, B. E. Shakhnovich, and E. I. Shakhnovich (2003) Natural selection of more designable folds: a mechanism for thermophilic adaptation. *Proc. Natl. Acad. Sci. USA* **100**, pp. 8727–31
28. A. V. Finkelstein, A. Badretdinov, and A. M. Gutin (1995) Why do protein architectures have Boltzmann-like statistics? *Proteins* **23**, pp. 142–50
29. B. E. Shakhnovich, E. Deeds, C. Delisi, and E. Shakhnovich (2005) Protein structure and evolutionary history determine sequence space topology. *Genome Res.* **15**, pp. 385–92
30. D. M. Taverna and R. A. Goldstein (2000) The distribution of structures in evolving protein populations. *Biopolymers* **53**, pp. 1–8
31. G. Tiana, B. E. Shakhnovich, N. V. Dokholyan, and E. I. Shakhnovich (2004) Imprint of evolution on protein structures. *Proc. Natl. Acad. Sci. USA* **101**, pp. 2846–51
32. N. V. Dokholyan, B. Shakhnovich, and E. I. Shakhnovich (2002) Expanding protein universe and its origin from the biological Big Bang. *Proc. Natl. Acad. Sci. USA* **99**, pp. 14132–6
33. A. V. Finkelstein, A. M. Gutun, and A. Badretdinov (1993) Why are the same protein folds used to perform different functions? *FEBS Lett.* **325**, pp. 23–8
34. R. Melin, H. Li, N.S. Wingreen, and C. Tang (1999) Designability, thermodynamic stability, and dynamics in protein folding: A lattice model study. *J. Chem. Phys.* **110**, pp. 1252
35. E. Shakhnovich and A. Gutin (1990) Enumeration of all compact conformations of copolymers with random sequence of links. *J. Chem. Phys.* **93**, pp. 5967–5971
36. F. Seno, M. Vendruscolo, A. Maritan, and J. R. Banavar (1996) Optimal Protein Design Procedure. *Phys. Rev. Lett.* **77**, pp. 1901–1904
37. C. A. Orengo, D. T. Jones, and J. M. Thornton (1994) Protein superfamilies and domain superfolds. *Nature* **372**, pp. 631–4
38. K. W. Plaxco, K. T. Simons, and D. Baker (1998) Contact order, transition state placement and the refolding rates of single domain proteins. *J. Mol. Biol.* **277**, pp. 985–94
39. I. N. Berezovsky and E. I. Shakhnovich (2005) From mechanism to strategy. Physics and evolution of thermophilic adaptation. *Proc. Natl. Acad. Sci. USA* **102**, pp. 12742–12747
40. V. I. Abkevich, A. M. Gutin, and E. I. Shakhnovich (1994) Specific nucleus as the transition state for protein folding: evidence from the lattice model. *Biochemistry* **33**, pp. 10026–36
41. E. I. Shakhnovich (1997) Theoretical studies of protein-folding thermodynamics and kinetics. *Curr. Opin. Struct. Biol.* **7**, pp. 29–40

42. L. S. Itzhaki, D. E. Otzen, and A. R. Fersht (1995) The structure of the transition state for folding of chymotrypsin inhibitor 2 analysed by protein engineering methods: evidence for a nucleation-condensation mechanism for protein folding. *J. Mol. Biol.* **254**, pp. 260–88
43. A. R. Fersht (1995) Optimization of rates of protein folding: the nucleation-condensation mechanism and its implications. *Proc. Natl. Acad. Sci. USA* **92**, pp. 10869–73
44. A. R. Fersht and V. Daggett (2002) Protein folding and unfolding at atomic resolution. *Cell* **108**, pp. 573–82
45. A. M. Gutin, V. I. Abkevich, and E. I. Shakhnovich (1995) Is burst hydrophobic collapse necessary for protein folding? *Biochemistry* **34**, pp. 3066–76
46. T. R. Sosnick, L. Mayne, R. Hiller, and S. W. Englander (1994) The barriers in protein folding. *Nat. Struct. Biol.* **1**, pp. 149–56
47. P. X. Qi, T. R. Sosnick, and S. W. Englander (1998) The burst phase in ribonuclease: A folding and solvent dependence of the unfolded state. *Nat. Struct. Biol.* **5**, pp. 882–4
48. S. H. Park, M. C. Shastry, and H. Roder (1999) Folding dynamics of the B1 domain of protein G explored by ultrarapid mixing. *Nat. Struct. Biol.* **6**, pp. 943–7
49. L. A. Mirny and E. I. Shakhnovich (1999) Universally conserved positions in protein folds: reading evolutionary signals about stability, folding kinetics and function. *J. Mol. Biol.* **291**, pp. 177–96
50. N. V. Dokholyan, S. V. Buldyrev, H. E. Stanley, and E. I. Shakhnovich, (2000) Identifying the protein folding nucleus using molecular dynamics. *J. Mol. Biol.* **296**, pp. 1183–8
51. J. E. Shea, J. N. Onuchic, and C. L. Brooks, 3rd. (2002) Probing the folding free energy landscape of the src-SH3 protein domain. *Proc. Natl. Acad. Sci. USA* **99**, pp. 16064–8
52. J. Shimada and E. I. Shakhnovich (2002) The ensemble folding kinetics of protein G from an all-atom Monte Carlo simulation. *Proc. Natl. Acad. Sci. USA* **99**, pp. 11175–80
53. F. Ding, W. Guo, N. V. Dokholyan, E. Shakhnovich, and J. E. Shea (2005) Reconstruction of the src-SH3 protein domain transition state ensemble using multiscale molecular dynamics simulations. *J. Mol. Biol.* **350**, pp. 1035–50
54. F. Chiti, N. Taddei, P. M. White, M. Bucciantini, F. Magherini, M. Stefani, and C. M. Dobson (1999) Mutational analysis of acylphosphatase suggests the importance of topology and contact order in protein folding. *Nat. Struct. Biol.* **6**, pp. 1005–9
55. J. C. Martinez, M. T. Pisabarro, and L. Serrano (1998) Obligatory steps in protein folding and the conformational diversity of the transition state. *Nat. Struct. Biol.* **5**, pp. 721–9
56. J. Shimada, E. L. Kussell, and E. I. Shakhnovich (2001) The folding thermodynamics and kinetics of crambin using an all-atom Monte Carlo simulation. *J. Mol. Biol.* **308**, pp. 79–95
57. Hubner (2005) Nucleation and the transition state of the SH3 domain. (under review at JMB)
58. I. A. Hubner, E. J. Deeds, and E. I. Shakhnovich (2005) High resolution protein folding with a transferable potential. *Proc. Natl. Acad. Sci. USA* **102**, pp. 18914–9

59. F. M. Richards and W. A. Lim (1993) An analysis of packing in the protein folding problem. *Q. Rev. Biophys.* **26**, pp. 423–98
60. S. Bromberg and K. A. Dill (1994) Side-chain entropy and packing in proteins. *Protein Sci.* **3**, pp. 997–1009
61. E. Kussell, J. Shimada, and E. I. Shakhnovich (2001) Excluded volume in protein side-chain packing. *J. Mol. Biol.* **311**, pp. 183–93
62. V. N. Maiorov and G. M. Crippen (1992) Contact potential that recognizes the correct folding of globular proteins. *J. Mol. Biol.* **227**, pp. 876–88
63. W. Chen and E. I. Shakhnovich (2005) Lessons from design of new all-atom potential for protein folding. *Protein Sci.* in press
64. M. Vendruscolo, L. A. Mirny, E. I. Shakhnovich, and E. Domany (2000) Comparison of two optimization methods to derive energy parameters for protein folding: perceptron and Z score. *Proteins* **41**, pp. 192–201
65. E. Kussell, J. Shimada, and E. I. Shakhnovich (2003) Side-chain dynamics and protein folding. *Proteins* **52**, pp. 303–21
66. S. Takada, J. J. Portman, and P. G. Wolynes (1997) An elementary mode coupling theory of random heteropolymer dynamics. *Proc. Natl. Acad. Sci. USA* **94**, pp. 2318–21
67. D. Thirumalai, V. V. Ashwin, and J. K. Bhattacharjee (1996) Dynamics of Random Hydrophobic-Hydrophilic Copolymers with Implications for Protein Folding. *Phys. Rev. Lett.* **77**, pp. 5385–5388
68. E. Pitard and E. I. Shakhnovich (2001) Mode-coupling theory for heteropolymers. *Phys. Rev. E. Stat. Nonlin. Soft. Matter Phys.* **63**, p. 041501
69. A. Matouschek, J. T. Kellis, L. Jr. Serrano, and A. R. Fersht (1989) Mapping the transition state and pathway of protein folding by protein engineering. *Nature* **340**, pp. 122–6
70. A. G. Ladurner, L. S. Itzhaki, and A. R. Fersht (1997) Strain in the folding nucleus of chymotrypsin inhibitor 2. *Fold Des.* **2**, pp. 363–8
71. E. L. McCallister, E. Alm, and D. Baker (2000) Critical role of beta-hairpin formation in protein G folding. *Nat. Struct. Biol.* **7**, pp. 669–73
72. V. Daggett, A. Li, L. S. Itzhaki, D. E. Otzen, and A. R. Fersht (1996) Structure of the transition state for folding of a protein derived from experiment and simulation. *J. Mol. Biol.* **257**, pp. 430–40
73. E. Paci, M. Vendruscolo, C. M. Dobson, and M. Karplus (2002) Determination of a transition state at atomic resolution from protein engineering data. *J. Mol. Biol.* **324**, pp. 151–63
74. M. Vendruscolo, E. Paci, C. M. Dobson, and M. Karplus (2001) Three key residues form a critical contact network in a protein folding transition state. *Nature* **409**, pp. 641–5
75. R. Du, V. Pande, A. Grosberg, T. Tanaka, and E. I. Shakhnovich (1998) On the transition coordinate for protein folding. *J. Chem. Phys.* **108**, pp. 334–50
76. L. Li and E. I. Shakhnovich (2001) Constructing, verifying, and dissecting the folding transition state of chymotrypsin inhibitor 2 with all-atom simulations. *Proc. Natl. Acad. Sci. USA* **98**, pp. 13014–8
77. I. A. Hubner, J. Shimada, and E. I. Shakhnovich (2004) Commitment and nucleation in the protein G transition state. *J. Mol. Biol.* **336**, pp. 745–61
78. C. F. Wright, K. Lindorff-Larsen, L. G. Randles, and J. Clarke (2003) Parallel protein-unfolding pathways revealed and mapped. *Nat. Struct. Biol.* **10**, pp. 658–62

79. F. Ding, W. Guo, N. V. Dokholyan, E. Shakhnovich, and J. E. Shea (2005) Reconstruction of the src-SH3 protein domain transition state ensemble using multiscale molecular dynamics simulations. *J. Mol. Biol.* in press
80. I. A. Hubner, M. Oliveberg, and E. I. Shakhnovich (2004) Simulation, experiment, and evolution: understanding nucleation in protein S6 folding. *Proc. Natl. Acad. Sci. USA* **101**, pp. 8354–9
81. D. S. Riddle, V. P. Grantcharova, J. V. Santiago, E. Alm, I. Ruczinski, and D. Baker (1999) Experiment and theory highlight role of native state topology in SH3 folding. *Nat. Struct. Biol.* **6**, pp. 1016–24
82. E. Alm and D. Baker (1999) Prediction of protein-folding mechanisms from free-energy landscapes derived from native structures. *Proc. Natl. Acad. Sci. USA* **96**, pp. 11305–10
83. E. Shakhnovich, V. Abkevich, and O. Ptitsyn (1996) Conserved residues and the mechanism of protein folding. *Nature* **379**, pp. 96–8
84. L. A. Mirny, V. I. Abkevich, and E. I. Shakhnovich (1998) How evolution makes proteins fold quickly? *Proc. Natl. Acad. Sci. USA* **95**, pp. 4976–81
85. O. B. Ptitsyn and K. L. Ting (1999) Non-functional conserved residues in globins and their possible role as a folding nucleus. *J. Mol. Biol.* **291**, pp. 671–82
86. L. Li, L. A. Mirny, and E. I. Shakhnovich (2000) Kinetics, thermodynamics and evolution of non-native interactions in a protein folding nucleus. *Nat. Struct. Biol.* **7**, pp. 336–42
87. J. Clarke, E. Cota, S. B. Fowler, and S. J. Hamill (1999) Folding studies of immunoglobulin-like beta-sandwich proteins suggest that they share a common folding pathway. *Structure Fold. Des.* **7**, pp. 1145–53
88. C. T. Friel, G. S. Beddard, and S. E. Radford (2004) Switching two-state to three-state kinetics in the helical protein Im9 via the optimisation of stabilising non-native interactions by design. *J. Mol. Biol.* **342**, pp. 261–73
89. D. E. Kim, C. Fisher, and D. Baker (2000) A breakdown of symmetry in the folding transition state of protein L. *J. Mol. Biol.* **298**, pp. 971–84
90. C. T. Friel, A. P. Capaldi, and S. E. Radford (2003) Structural analysis of the rate-limiting transition states in the folding of Im7 and Im9: similarities and differences in the folding of homologous proteins. *J. Mol. Biol.* **326**, pp. 293–305
91. A. Zarrine-Afsar, S. M. Larson, and A. R. Davidson (2005) The family feud: Do proteins with similar structures fold via the same pathway? *Curr. Opin. Struct. Biol.* **15**, pp. 42–9
92. Y. Zhou, C. K. Hall, and M. Karplus (1996) First-Order Disorder-to-Order Transition in an Isolated Homopolymer Model. *Phys. Rev. Lett.* **77**, pp. 2822–2825
93. H. Jang, C. K. Hall, and Y. Zhou (2002) Protein folding pathways and kinetics: molecular dynamics simulations of beta-strand motifs. *Biophys. J.* **83**, pp. 819–35
94. Y. Zhou and M. Karplus (1999) Interpreting the folding kinetics of helical proteins. *Nature* **401**, pp. 400–3
95. N. V. Dokholyan, S. V. Buldyrev, H. E. Stanley, and E. I. Shakhnovich (1998) Discrete molecular dynamics studies of the folding of a protein-like model. *Fold. Des.* **3**, pp. 577–87

96. F. Ding, N. V. Dokholyan, S. V. Buldyrev, H. E. Stanley, and E. I. Shakhnovich (2002) Direct molecular dynamics observation of protein folding transition state ensemble. *Biophys. J.* **83**, pp. 3525–32
97. F. Ding, N. V. Dokholyan, S. V. Buldyrev, H. E. Stanley, and E. I. Shakhnovich (2002) Molecular dynamics simulation of the SH3 domain aggregation suggests a generic amyloidogenesis mechanism. *J. Mol. Biol.* **324**, pp. 851–7
98. J. M. Borreguero, F. Ding, S. V. Buldyrev, H. E. Stanley, and N. V. Dokholyan (2004) Multiple folding pathways of the SH3 domain. *Biophys. J.* **87**, pp. 521–33
99. S. Peng, F. Ding, B. Urbanc, S. V. Buldyrev, L. Cruz, H. E. Stanley, and N. V. Dokholyan (2004) Discrete molecular dynamics simulations of peptide aggregation. *Phys. Rev. E. Stat. Nonlin. Soft. Matter. Phys.* **69**, p. 041908
100. S. Yang, S. S. Cho, Y. Levy, M. S. Cheung, H. Levine, P. G. Wolynes, and J. N. Onuchic (2004) Domain swapping is a consequence of minimal frustration. *Proc. Natl. Acad. Sci. USA* **101**, pp. 13786–91
101. F. Ding, S. V. Buldyrev, and N. V. Dokholyan (2005) Folding Trp-cage to NMR resolution native structure using a coarse-grained protein model. *Biophys. J.* **88**, pp. 147–55
102. C. Simmerling, B. Strockbine, and A. E. Roitberg (2002) All-atom structure prediction and folding simulations of a stable protein. *J. Am. Chem. Soc.* **124**, pp. 11258–9
103. L. Qiu, S. A. Pabit, A. E. Roitberg, and S. J. Hagen (2002) Smaller and faster: the 20-residue Trp-cage protein folds in 4 micros. *J. Am. Chem. Soc.* **124**, pp. 12952–3
104. A. R. Viguera, C. Vega, and L. Serrano (2002) Unspecific hydrophobic stabilization of folding transition states. *Proc. Natl. Acad. Sci. USA* **99**, pp. 5349–54
105. S. Miyazawa and R. L. Jernigan (1996) Residue-residue potentials with a favorable contact pair term and an unfavorable high packing density term, for simulation and threading. *J. Mol. Biol.* **256**, pp. 623–44
106. J. Skolnick, L. Jaroszewski, A. Kolinski, and A. Godzik (1997) Derivation and testing of pair potentials for protein folding. When is the quasichemical approximation correct? *Protein. Sci.* **6**, pp. 676–88
107. L. Zhang and J. Skolnick (1998) How do potentials derived from structural databases relate to "true" potentials? *Protein. Sci.* **7**, pp. 112–22
108. E. Kussell, J. Shimada, and E. I. Shakhnovich (2002) A structure-based method for derivation of all-atom potentials for protein folding. *Proc. Natl. Acad. Sci. USA* **99**, pp. 5343–8
109. D. Shortle, K. T. Simons, and D. Baker (1998) Clustering of low-energy conformations near the native structures of small proteins. *Proc. Natl. Acad. Sci. USA* **95**, pp. 11158–62
110. C. D. Sfatos and E. Shakhnovich (1997) Statistical Mechanics of Random Heteropolymers. *Physics Reports* **288**, pp. 77–108
111. B. Zagrovic, C. D. Snow, M. R. Shirts, and V. S. Pande (2002) Simulation of folding of a small alpha-helical protein in atomistic detail using worldwide-distributed computing. *J. Mol. Biol.* **323**, pp. 927–37
112. C. D. Snow, H. Nguyen, V. S. Pande, and M. Gruebele (2002) Absolute comparison of simulated and experimental protein-folding dynamics. *Nature* **420**, pp. 102–6
113. J. R. Banavar and A. Maritan (2003) Colloquium: Geometrical approach to protein folding: a tube picture. *Rev. Mod. Phys.* **75**, p. 23

114. J. R. Banavar, A. Maritan, C. Micheletti, and A. Trovato (2002) Geometry and physics of proteins. *Proteins* **47**, pp. 315–22
115. A. Maritan, C. Micheletti, A. Trovato, and J. R. Banavar (2000) Optimal shapes of compact strings. *Nature* **406**, pp. 287–90
116. T. X. Hoang, A. Trovato, F. Seno, J. R. Banavar, and A. Maritan (2004) Geometry and symmetry presculpt the free-energy landscape of proteins. *Proc. Natl. Acad. Sci. USA* **101**, pp. 7960–4
117. A. V. Finkelstein and O. B. Ptitsyn (1987) Why do globular proteins fit the limited set of folding patterns? *Prog. Biophys. Mol. Biol.* **50**, pp. 171–90

Index

ab initio centroid molecular dynamics 518
ab initio molecular dynamics 509
ab initio path integrals 513
Adam-Gibbs 11
Adam-Gibbs theory 13
adiabatic basis 530
algorithm 533
Amontons' law 74
amphiphiles 217, 224
Amphiphilic membranes 214
Angell-plot 4
approximation 514
AT line 41
atactic polystyrene 427

bead-spring models 315
biomembrane 214
blob size 315
Boltzmann equation 331
Born term 292, 294
Born–Oppenheimer 514
Born–Oppenheimer approximation 510
Boson peak 8, 28
Bridging 105
Brownian dynamics 23
bulk viscosity 89, 333

cage-effect 8, 20
Car 510
Car–Parrinello Lagrangian 511
Car–Parrinello scheme 509
Cauchy stress tensor 290
centroid adiabaticity parameter 522
centroid force 521
centroid molecular dynamics 518
chain entanglements 350, 362, 436
chain form factor 355, 365
chemical gels 11
coarse-graining 213, 325, 422, 426
coarse-graining using pretabulated potentials 431
colloid 141, 203
colloidal gels 11
complex fluid 213
compressible Ising model 129
confined polar fluids 51
connectivity-altering algorithms 423
constraint forces 511
cooperatively rearranging regions 12
correlation function 8
Coulomb's law of friction 72
coupling elements, D^{kl} 532
CReTA (contour reduction topological analysis) 441
critical temperature of mode-coupling theory 5, 17
CRR 12, 13

defect 203, 205
dendrimers 142
dielectric permittivity 47
diffusion constant 349, 362
dipolar fluctuations 51
direct correlation function 16
director 193, 195
discretization errors 328

dissipative particle dynamics 86, 326
double bridging 423
DPD thermostat 243
dynamic exponent 319
dynamic triangulation algorithm 229
dynamical heterogeneities 22, 23, 25, 28

effective force 206
effective interaction 143
Ehrenfest equations 531
Ehrenfest molecular dynamics 529
elastic constants 167, 195, 196, 291
elasticity theory 290
entanglement molecular weight 345, 372
equilibrated melts 354, 355
extended mode-coupling equations 27
extrapolation length 200

FENE potential 357
Fermi–Pasta–Ulam chains 517
fewest switches 533
fictitious masses 510
fictitious normal mode masses 522
film rupture 120
finite elements 82
finite size scaling 39, 168
fluctuation dissipation theorem 24, 321, 328
fluctuation term 292
Fokker–Planck equation 327
Fokker–Planck operator 327
fragile glass-former 5, 9, 10, 13
Frank free energy 195
Frenkel Kontorova (FK) model 77
frequency matrix 14
friction coefficient 335
frustration 3

Galilean invariance 328
Gaussian thermostats 242
generalized Brillouin theorem 527
generalized Langevin equation 15
glass 10
glass transition 3, 13
glass transition temperature 9
glass-forming liquids 5
glass-forming systems 3, 4, 6, 28
glassy liquids 6
glassy systems 27
ground state dominance 515
gyration radius 314

hard disks 165
hard particles 194
helical twisting power 197
hopping processes 27
hydrodynamic interaction 320
Hydrodynamic screening 322
hydrodynamic screening length 322

ideal mode-coupling equations 27
image charges 58
influence functional 514
intermediate scattering function 16, 17, 19–21, 24, 25
invariance 314
Ising model-binary alloy 130
iterative Boltzmann inversion 427

Kauzmann temperature 9, 10, 13
Keating interatomic potential 130
kinematic viscosity 317
kinetic friction 76
Kirkwood's fluctuation formula 49
Kohlrausch-Williams-Watts function 8
(KTHNY) theory 165
Kuhn length 344
KWW 8, 21, 22

Lagrangian strain tensor 290
Lagrangians 510
Landau-Ginsburg-Wilson hamiltonian 129
Langevin equation 316
Langevin thermostat 84, 243
laser induced freezing 175
laser induced melting 175
lattice BGK 332
lattice Boltzmann method 331
Lattice-Boltzmann scheme 249
Lees-Edwards 239
Lees-Edwards periodic boundary conditions 91
Lennard-Jones 7, 18
liquid benzene 433

liquid crystals 235
local permittivity 47
long-time tails 317
lubricant additives 99

Massey parameter 530
MCT 13, 24, 25, 27
mean square displacements 360, 362
mean-squared displacement 19
melt viscosity 349
membrane phases 216
membrane pores 225
memory function 15
mesogen 194
mesoscopic simulation methods 335
mode-coupling equations 17
mode-coupling equations with hopping terms 27
mode-coupling theory 11, 13, 17
multi-particle collision dynamics 330

nanodroplets 105
nanofibers 112
nematic phase 193, 194
nematic-isotropic interface 201
network elasticity 345, 352
non-equilibrium molecular dynamics 239
nonadiabatic coupling elements 530, 533
nonadiabatic coupling *vectors*, $\mathbf{d}_I^{lk}$ 532
nonadiabatic ab initio molecular dynamics 529, 533
nonergodic behavior 517
normal mode frequencies 519
normal mode variables 519
normal modes 519
Nosé Hoover thermostats 242
Nosé–Hoover chain thermostats 517

occupation numbers 530
on-the-fly 509
Onsager free energy 195
order parameter 195
order tensor 194

packing length 372, 439
parallel tempering 37
Parrinello 510
path integral 513
phase coexistence 201
phase diagram of a polymer solution 311
plastic deformation 71
plateau modulus 353, 364, 369, 372
polar fluids 47
polarisation density 52
polycarbonate 366
polymer dynamics 311
polymer melts 343, 422
polymer networks 346, 352
polymer solutions 313
potential of mean force 315
power-law behavior 314
Prandtl-Tomlinson model 70
primitive path analysis, PPA 369, 370
profile-unbiased thermostats 328
projectors 14
pseudo-equilibrium distribution 332

(quasiclassical) time evolution 518

radius of gyration 344
random walk 314
random walks 352
renormalization group 314
reptation 345, 346, 350
restricted open-shell Kohn–Sham 523
restricted open-shell singlet 528
ripple phase 220
ROKS functional 526
Rouse model 315, 318, 348, 349
Rouse time 349

Schmidt number 320
second-order memory function 15
self avoiding walk 344
self consistent field theory 299
self-affine surfaces 81
self-assembly 217
self-avoiding walk 314
shear banding 234
shear stress 233
shear thickening 234
shear thinning 234
shear viscosity 334
silica 3
smectic phase 193

soft matter 141
sol-gel transition 11
specific heat 9, 13
spin glasses 33
spreading 105
star polymers 142
static friction 68
static structure factor 16
stochastic dynamics 24, 326
Stokes equation 317
strain rate 233, 234
strong glass-formers 5
structure factor 6, 7, 314
structure property relations 344, 354
supercooled liquids 1
superlubricity 95
surface anchoring 195, 199, 200
surface tension 202

theory 11
thermodynamic *tension* tensor 290
thermostat 241
time correlation functions 7, 14, 15
time-temperature superposition principle 21
topological analysis of atomistic melt configurations 439
topological costraints 369
Trotter factorization 515
Trotter slices 515
tube diameter 346, 360
tube model 343, 369
Tully molecular dynamics 533

Vesicles 215
viscosity 4, 6
viscosity tensor 333
Vogel temperature 10, 13
Vogel-Fulcher-Tammann-law 7
Vogel-temperature 7

Wetting 105

Zimm model 320
Zwanzig-Mori formalism 15, 16, 27
Zwanzig-Mori projection operator formalism 14

Lecture Notes in Physics

For information about earlier volumes
please contact your bookseller or Springer
LNP Online archive: springerlink.com

Vol.659: E. Bick, F. D. Steffen (Eds.), Topology and Geometry in Physics

Vol.660: A. N. Gorban, I. V. Karlin, Invariant Manifolds for Physical and Chemical Kinetics

Vol.661: N. Akhmediev, A. Ankiewicz (Eds.) Dissipative Solitons

Vol.662: U. Carow-Watamura, Y. Maeda, S. Watamura (Eds.), Quantum Field Theory and Noncommutative Geometry

Vol.663: A. Kalloniatis, D. Leinweber, A. Williams (Eds.), Lattice Hadron Physics

Vol.664: R. Wielebinski, R. Beck (Eds.), Cosmic Magnetic Fields

Vol.665: V. Martinez (Ed.), Data Analysis in Cosmology

Vol.666: D. Britz, Digital Simulation in Electrochemistry

Vol.667: W. D. Heiss (Ed.), Quantum Dots: a Doorway to Nanoscale Physics

Vol.668: H. Ocampo, S. Paycha, A. Vargas (Eds.), Geometric and Topological Methods for Quantum Field Theory

Vol.669: G. Amelino-Camelia, J. Kowalski-Glikman (Eds.), Planck Scale Effects in Astrophysics and Cosmology

Vol.670: A. Dinklage, G. Marx, T. Klinger, L. Schweikhard (Eds.), Plasma Physics

Vol.671: J.-R. Chazottes, B. Fernandez (Eds.), Dynamics of Coupled Map Lattices and of Related Spatially Extended Systems

Vol.672: R. Kh. Zeytounian, Topics in Hyposonic Flow Theory

Vol.673: C. Bona, C. Palenzula-Luque, Elements of Numerical Relativity

Vol.674: A. G. Hunt, Percolation Theory for Flow in Porous Media

Vol.675: M. Kröger, Models for Polymeric and Anisotropic Liquids

Vol.676: I. Galanakis, P. H. Dederichs (Eds.), Halfmetallic Alloys

Vol.677: A. Loiseau, P. Launois, P. Petit, S. Roche, J.-P. Salvetat (Eds.), Understanding Carbon Nanotubes

Vol.678: M. Donath, W. Nolting (Eds.), Local-Moment Ferromagnets

Vol.679: A. Das, B. K. Chakrabarti (Eds.), Quantum Annealing and Related Optimization Methods

Vol.680: G. Cuniberti, G. Fagas, K. Richter (Eds.), Introducing Molecular Electronics

Vol.681: A. Llor, Statistical Hydrodynamic Models for Developed Mixing Instability Flows

Vol.682: J. Souchay (Ed.), Dynamics of Extended Celestial Bodies and Rings

Vol.683: R. Dvorak, F. Freistetter, J. Kurths (Eds.), Chaos and Stability in Planetary Systems

Vol.684: J. Dolinšek, M. Vilfan, S. Žumer (Eds.), Novel NMR and EPR Techniques

Vol.685: C. Klein, O. Richter, Ernst Equation and Riemann Surfaces

Vol.686: A. D. Yaghjian, Relativistic Dynamics of a Charged Sphere

Vol.687: J. W. LaBelle, R. A. Treumann (Eds.), Geospace Electromagnetic Waves and Radiation

Vol.688: M. C. Miguel, J. M. Rubi (Eds.), Jamming, Yielding, and Irreversible Deformation in Condensed Matter

Vol.689: W. Pötz, J. Fabian, U. Hohenester (Eds.), Quantum Coherence

Vol.690: J. Asch, A. Joye (Eds.), Mathematical Physics of Quantum Mechanics

Vol.691: S. S. Abdullaev, Construction of Mappings for Hamiltonian Systems and Their Applications

Vol.692: J. Frauendiener, D. J. W. Giulini, V. Perlick (Eds.), Analytical and Numerical Approaches to Mathematical Relativity

Vol.693: D. Alloin, R. Johnson, P. Lira (Eds.), Physics of Active Galactic Nuclei at all Scales

Vol.694: H. Schwoerer, J. Magill, B. Beleites (Eds.), Lasers and Nuclei

Vol.695: J. Dereziński, H. Siedentop (Eds.), Large Coulomb Systems

Vol.696: K.-S. Choi, J. E. Kim, Quarks and Leptons From Orbifolded Superstring

Vol.697: E. Beaurepaire, H. Bulou, F. Scheurer, J.-P. Kappler (Eds.), Magnetism: A Synchrotron Radiation Approach

Vol.698: S. Bellucci (Ed.), Supersymmetric Mechanics – Vol. 1

Vol.699: J.-P. Rozelot (Ed.), Solar and Heliospheric Origins of Space Weather Phenomena

Vol.700: J. Al-Khalili, E. Roeckl (Eds.), The Euroschool Lectures on Physics with Exotic Beams, Vol. II

Vol.701: S. Bellucci, S. Ferrara, A. Marrani, Supersymmetric Mechanics – Vol. 2

Vol.702: J. Ehlers, C. Lämmerzahl, Special Relativity

Vol.703: M. Ferrario, G. Ciccotti, K. Binder (Eds.), Computer Simulations in Condensed Matter Systems: From Materials to Chemical Biology Volume 1

Vol.704: M. Ferrario, G. Ciccotti, K. Binder (Eds.), Computer Simulations in Condensed Matter Systems: From Materials to Chemical Biology Volume 2