

John von Neumann Institute for Computing (NIC)

Computational Soft Matter: From Synthetic Polymers to Proteins

edited by

Norbert Attig

Kurt Binder

Helmut Grubmüller

Kurt Kremer

Winter School, 29 February - 6 March 2004

Gustav-Stresemann-Institut, Bonn, Germany

Lecture Notes

organized by

John von Neumann Institute for Computing

Johannes Gutenberg University of Mainz

Max Planck Institute for Biophysical Chemistry, Göttingen

Max Planck Institute for Polymer Research, Mainz

NIC Series

Volume 23

ISBN 3-00-012641-4

Die Deutsche Bibliothek – CIP-Cataloguing-in-Publication-Data
A catalogue record for this publication is available from Die Deutsche
Bibliothek.

Publisher: NIC-Directors
Distributor: NIC-Secretariat
Research Centre Jülich
52425 Jülich
Germany
Internet: www.fz-juelich.de/nic

Printer: Graphische Betriebe, Forschungszentrum Jülich

© 2004 by John von Neumann Institute for Computing
Permission to make digital or hard copies of portions of this work
for personal or classroom use is granted provided that the copies
are not made or distributed for profit or commercial advantage and
that copies bear this notice and the full citation on the first page. To
copy otherwise requires prior specific permission by the publisher
mentioned above.

NIC Series Volume 23
ISBN 3-00-012641-4

Preface

The John von Neumann-Institut für Computing (NIC) - a joint institute of Deutsches Elektronen-Synchrotron (DESY) and Forschungszentrum Jülich - supports a large number of research projects in computational science, mainly through the Zentralinstitut für Angewandte Mathematik (ZAM) in Jülich and its supercomputing facilities. Furthermore, NIC also plays an active role in the education of young researchers in the various areas of computational science. It may already be called a tradition that every second year in February/March NIC offers a Winter School about a topic of outstanding methodological importance to the NIC user community.

This year the focus of the Winter School is on Computational Soft Matter which has become a very active field of research. Characteristic features of soft matter simulations are the nontrivial geometric structures that occur from the atomistic to the mesoscopic scales, the importance of entropic effects, and the cooperative complex dynamics. Different application fields profit from the recent progress of simulation methods.

The Winter School covers the following research topics:

- Introduction to Simulation Methods
- Polymer Solutions and Melts
- Liquid Crystals, Surfactants and Membranes
- Biopolymers

The introductory overview lectures on simulation methods are written by leading experts who are internationally very well known also because of their excellent books: Mike Allen, coauthor of the book “Computer simulation of liquids”, surveys the Molecular Dynamics methods; Daan Frenkel, coauthor of the book “Understanding Molecular Simulation: From Algorithms to Applications”, deals with Monte Carlo methods; Burkhard Dünweg, coeditor of the book “Computer simulation of surfaces and interfaces”, introduces mesoscopic simulation methods, the Lattice Boltzmann technique and Dissipative Particle Dynamics in particular.

The following group of lecture notes describes the state of the art with respect to the simulation of polymer solutions of melts. Again these contributions are given by worldwide recognized specialists. Jörg Baschnagel addresses Monte Carlo simulations of coarse-grained models, while Kurt Kremer focuses on the problem of entangled polymers in melts and rubbers, applying mostly Molecular Dynamics methods. The latter are also used in the simulations described by Wolfgang Paul, who reviews work on chemically realistic models of melts of neutral polymers. In contrast, Christian Holm deals with polyelectrolyte solutions, where special problems occur since one must account for the long range Coulomb interactions. Still a different level of complexity is treated in the contribution by Marcus Müller on macrophase and microphase separation in polymer mixtures and block copolymer melts, respectively. The last lecture note in this group by Cameron Abrams describes polymers at polymer-metal interfaces and the adaption of coarse-graining methods to this problem.

All these papers restrict attention to flexible polymers where orientational order is not an issue. However, orientational order is an important aspect in the second group of lecture notes that are about the simulation of liquid crystals, surfactants and membranes, again written by very well known researchers whose work is on the forefront of their respective fields. Mike Allen describes simulations for models of nematics and other types of liquid crystalline systems, while Alan Mark deals with the self-organisation in peptide and lipid systems, and Friederike Schmid addresses structure and phase behavior of surfactants at interfaces. Surfactants and lipids are amphiphilic (macro-)molecules, e.g. a hydrophilic head group is covalently bonded to a hydrophobic tail, typically a short alkane chain, allowing liquid-crystalline order in the surfactant or liquid layer due to the chain stiffness. While biomembranes have a particularly complex chemical structure, Ole Mouritsen's article shows that nevertheless progress in the understanding of biomembranes can be obtained by simulations of very simplified lattice models of these systems.

The last group of lecture notes, once more coming from established leaders of the field, describe some aspects of the simulation of biopolymers. Alexander Grosberg addresses the statistical mechanics approach to protein folding, while Helmut Grubmüller focuses on proteins as molecular machines, and Klaus Schulten on protein mechanics. Thus very important steps on the way towards the simulation of complex biological systems by models in the framework of statistical mechanics and physical chemistry are taken here.

The target group of the School are again young scientists, especially graduate students and postdocs, in the areas of theoretical/computational physics, chemistry, and biophysics who have a basic knowledge of classical, statistical, and quantum mechanics. They benefit from this School by learning about recent methodological advances within and outside their field of specialization.

The Winter School is organized by the John von Neumann-Institut für Computing (Forschungszentrum Jülich), Institut für Physik (Johannes Gutenberg-Universität Mainz), Max-Planck-Institut für Polymerforschung (Mainz), and Max-Planck-Institut für biophysikalische Chemie (Göttingen).

Financial support mainly came from the Forschungszentrum Jülich. Despite the lack of additional funding either from the EU or the German Ministry of Education and Research (BMBF), we succeeded to organize this high level international School on a very important field of computational physics/chemistry. The number of participants is restricted to about 80 participants compared to about 150 in previous times. This allows for a much closer contact between the lecturers and the students and fosters direct discussions between all participants throughout the School. Applicants for the School were selected on the basis of scientific background and excellence. In spite of this barrier, we received a wide-spread national and international resonance of applications, in most cases together with the submission of a poster abstract. This reflects the attractiveness of the programme and demonstrates the expectation of the participants to play an active role in this high-level scientific School. We are sure that the School is stimulating for both sides, the students as well as the lecturers.

This preface also offers an opportunity to thank all the individuals and institutions that significantly contributed to the success of the School. First of all we wish to thank all speakers for their written contributions. Without their efforts to generate the extended lecture notes at hand, in spite of the heavy work load they all have to carry, such an excellent reference to the rapidly evolving field of computational soft matter would not have been possible. We also wish to thank the Forschungszentrum Jülich, which this year was the main sponsor of the School. For their most valuable help with the local arrangements we are greatly indebted to several staff members of the Forschungszentrum Jülich, namely Rüdiger Esser (finance), Rene Gail (conference service), and last but not least the School's secretaries Anke Reinartz and Yasmin Abdel-Fattah. Special thanks go to Anke Visser for her commitment concerning the composition and realization of this Book of Lecture Notes.

Jülich, Mainz, and Göttingen
February 2004

Norbert Attig
Kurt Binder
Helmut Grubmüller
Kurt Kremer

Contents

Introduction to Simulation Methods

Introduction to Molecular Dynamics Simulation

<i>Michael P. Allen</i>	1
1 The Aims of Molecular Dynamics	1
2 Molecular Interactions	2
3 The MD Algorithm	7
4 Time Dependence	12
5 Rigid Molecule Rotation	17
6 Molecular Dynamics in Different Ensembles	20
7 How Long? How Large?	20
8 Conclusions	22

Introduction to Monte Carlo Methods

<i>Daan Frenkel</i>	29
1 Introduction	29
2 Canonical Ensemble	38
3 Grand-Canonical Ensemble	40
4 Beyond Metropolis	47
5 Relative Merits	55

Advanced Simulations for Hydrodynamic Problems: Lattice Boltzmann and Dissipative Particle Dynamics

<i>Burkhard Dünweg</i>	61
1 Polymer Dynamics	61
2 Simulation Methods	71
3 A Final Remark	80

Polymer Solutions and Melts

Monte Carlo Simulation of Polymers: Coarse-Grained Models

<i>Jörg Baschnagel, Joachim P. Wittmer, Hendrik Meyer</i>	83
1 Introduction	83
2 A Primer to Polymer Physics	85
3 Monte Carlo Methods: A Brief Overview	93
4 Some Coarse-Grained Simulation Models	97
5 Monte Carlo Methods for Polymers: From Local to Non-Local Moves	102
6 Monte Carlo Methods for Polymers: Rosenbluth Sampling and its Modern Variants	115
7 Conclusions	129

Entangled Polymers: From Universal Aspects to Structure Property Relations		
<i>Kurt Kremer</i>		141
1	Introduction	141
2	Polymer Dynamics and Network Elasticity	142
3	Theoretical Concepts	145
4	Summary	166
Chemically Realistic Computer Simulations of Polymer Melts: Equilibration Issues and Study of Relaxation Processes		
<i>Wolfgang Paul</i>		169
1	A Simple Chemically Realistic Force Field for PB	170
2	Equilibration of Polymer Melts	172
3	Local Relaxation in Polymer Melts	177
4	Large Scale Relaxation as Observable in Neutron Scattering Experiments	188
5	Summary	191
Efficient Methods for Long Range Interactions in Periodic Geometries Plus One Application		
<i>Christian Holm</i>		195
1	Introduction	195
2	Methods for the Coulomb Sum in 3D	196
3	Two Periodic Dimensions	206
4	The Coulomb Sum in One Periodic Dimension	212
5	The Dipolar Ewald Summation in 3D	213
6	The ESPResSo Project	220
7	One Application: Polyelectrolytes in Poor Solvent	225
8	Concluding Remarks	231
Phase Behavior and Chain Conformations in Polymer Blends: Monte Carlo Simulation vs Mean Field Theory		
<i>Marcus Müller</i>		237
1	Introduction	237
2	Background	239
3	Models and Techniques	242
4	Results	247
Inhomogeneous Coarse-Graining of Polymers and Polymer/Metal Interfaces		
<i>Cameron F. Abrams</i>		275
1	Introduction	275
2	The “Molecule-Fixed” Dual-Resolution Coarse-Grained Model of BPA-PC	277
3	Dual-Resolution BPA-PC: Results and Discussion	281
4	On the Development of Domain-Fixed Inhomogeneous Coarse-Graining Schemes: Initial Attempts with Liquid Methane	283
5	Conclusions	285

Liquid Crystals, Surfactants, and Membranes

Liquid Crystal Systems

Michael P. Allen

289

1	Introduction	289
2	Molecular Models	289
3	Theoretical Approaches	290
4	Liquid Crystal Phase Diagrams	291
5	Dynamics of Liquid Crystals	297
6	Elastic Properties	299
7	Nematic-Isotropic Interface	302
8	Colloid Particles in Liquid Crystals	307
9	Conclusions	314

Self-Organization of Biomolecular Systems: Simulating the Folding and Aggregation of Peptides, Proteins and Lipids.

Alan E. Mark, Siewert-Jan Marrink

321

Amphiphiles at Interfaces: Simulation of Structure and Phase Behavior

Friederike Schmid, Dominik Dücks, Olaf Lenz, Claire Loison

323

1	Introduction	323
2	Models and Levels of Coarse Graining	326
3	Applications to Amphiphilic Layers	328
4	Outlook	342

Random-Lattice Simulations of Biomembranes

Ole G. Mouritsen, Myer Bloom, John H. Ipsen, Ling Miao, Morten Nielsen, James Polson, Jenifer Thewalt, Ilpo Vattulainen, Hong Zhu, Martin J. Zuckermann

347

1	Introduction	347
2	Off-Lattice Models	352
3	Simulation Methods	356
4	Results	363
5	The Final Words	371

Biopolymers

Statistical Mechanics of Protein Folding: Some Outstanding Problems

Alexander Grosberg

375

1	Introduction	375
2	Sequences	380
3	Conformations, Conformation Space, and the Levinthal Paradox	382
4	Compact Conformations: Lattice Model	395
5	Conclusion	396

Proteins as Molecular Machines: Force Probe Simulations

Helmut Grubmüller

401

1	Introduction	401
2	Protein Structure and Dynamics	401
3	Protein Dynamics Simulations	403
4	Force Fields for Protein Dynamics	405
5	Force Probe Simulations	407
6	Simple Non-Equilibrium Statistical Mechanics of Enforced Dissociation	410
7	Primary Mechanical Energy Transfer Steps in the “Molecular Nano-Machine” F ₁ ATP Synthase	413

Mechanical Functions of Proteins

Klaus Schulten

423

Introduction to Molecular Dynamics Simulation

Michael P. Allen

Centre for Scientific Computing and Department of Physics,
University of Warwick, Coventry CV4 7AL, United Kingdom
E-mail: m.p.allen@warwick.ac.uk

In this chapter a summary is given of the key ingredients necessary to carry out a molecular dynamics simulation, with particular emphasis on macromolecular systems. We discuss the form of the intermolecular potential for molecules composed of atoms, and of non-spherical sub-units, giving examples of how to compute the forces and torques. We also describe some of the MD algorithms in current use. Finally, we briefly refer to the factors that influence the size of systems, and length of runs, that are needed to calculate statistical properties.

1 The Aims of Molecular Dynamics

We carry out computer simulations in the hope of understanding the properties of assemblies of molecules in terms of their structure and the microscopic interactions between them. This serves as a complement to conventional experiments, enabling us to learn something new, something that cannot be found out in other ways. The two main families of simulation technique are molecular dynamics (MD) and Monte Carlo (MC); additionally, there is a whole range of hybrid techniques which combine features from both. In this lecture we shall concentrate on MD. The obvious advantage of MD over MC is that it gives a route to dynamical properties of the system: transport coefficients, time-dependent responses to perturbations, rheological properties and spectra.

Computer simulations act as a bridge (see Fig. 1) between microscopic length and time scales and the macroscopic world of the laboratory: we provide a guess at the interactions between molecules, and obtain ‘exact’ predictions of bulk properties. The predictions are ‘exact’ in the sense that they can be made as accurate as we like, subject to the limitations imposed by our computer budget. At the same time, the hidden detail behind bulk measurements can be revealed. An example is the link between the diffusion coefficient and velocity autocorrelation function (the former easy to measure experimentally, the latter much harder). Simulations act as a bridge in another sense: between theory and experiment. We may test a theory by conducting a simulation using the same model. We may test the model by comparing with experimental results. We may also carry out simulations on the computer that are difficult or impossible in the laboratory (for example, working at extremes of temperature or pressure).

Ultimately we may want to make direct comparisons with experimental measurements made on specific materials, in which case a good model of molecular interactions is essential. The aim of so-called *ab initio* molecular dynamics is to reduce the amount of fitting and guesswork in this process to a minimum. On the other hand, we may be interested in phenomena of a rather generic nature, or we may simply want to discriminate between good and bad theories. When it comes to aims of this kind, it is not necessary to have a perfectly realistic molecular model; one that contains the essential physics may be quite suitable.

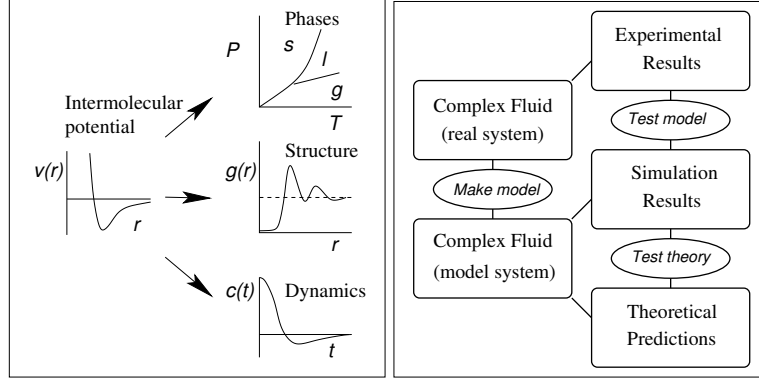


Figure 1. Simulations as a bridge between (a) microscopic and macroscopic; (b) theory and experiment.

2 Molecular Interactions

Molecular dynamics simulation consists of the numerical, step-by-step, solution of the classical equations of motion, which for a simple atomic system may be written

$$m_i \ddot{\mathbf{r}}_i = \mathbf{f}_i \quad \mathbf{f}_i = -\frac{\partial}{\partial \mathbf{r}_i} \mathcal{U} \quad (1)$$

For this purpose we need to be able to calculate the forces $\mathbf{f}_i$ acting on the atoms, and these are usually derived from a potential energy $\mathcal{U}(\mathbf{r}^N)$, where $\mathbf{r}^N = (\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N)$ represents the complete set of $3N$ atomic coordinates. In this section we focus on this function $\mathcal{U}(\mathbf{r}^N)$, restricting ourselves to an atomic description for simplicity. (In simulating soft condensed matter systems, we sometimes wish to consider non-spherical rigid units which have rotational degrees of freedom: rotational equations of motion and interaction potentials will be considered in section 5).

2.1 Non-bonded Interactions

The part of the potential energy $\mathcal{U}_{\text{non-bonded}}$ representing non-bonded interactions between atoms is traditionally split into 1-body, 2-body, 3-body ... terms:

$$\mathcal{U}_{\text{non-bonded}}(\mathbf{r}^N) = \sum_i u(\mathbf{r}_i) + \sum_i \sum_{j>i} v(\mathbf{r}_i, \mathbf{r}_j) + \dots \quad (2)$$

The $u(\mathbf{r})$ term represents an externally applied potential field or the effects of the container walls; it is usually dropped for fully periodic simulations of bulk systems. Also, it is usual to concentrate on the pair potential $v(\mathbf{r}_i, \mathbf{r}_j) = v(r_{ij})$ and neglect three-body (and higher order) interactions. There is an extensive literature on the way these potentials are determined experimentally, or modelled theoretically¹⁻⁴.

In some simulations of complex fluids, it is sufficient to use the simplest models that faithfully represent the essential physics. In this chapter we shall concentrate on continuous, differentiable pair-potentials (although discontinuous potentials such as hard spheres

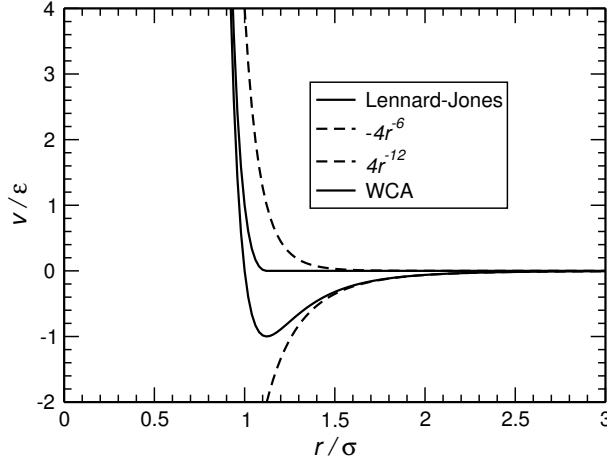


Figure 2. Lennard-Jones pair potential showing the r^{-12} and r^{-6} contributions. Also shown is the WCA shifted repulsive part of the potential.

and spheroids have also played a role, see e.g. 5). The Lennard-Jones potential is the most commonly used form:

$$v^{\text{LJ}}(r) = 4\epsilon \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^6 \right]. \quad (3)$$

with two parameters: σ , the diameter, and ϵ , the well depth. This potential was used, for instance, in the earliest studies of the properties of liquid argon^{6,7} and is illustrated in Fig. 2. For applications in which attractive interactions are of less concern than the excluded volume effects which dictate molecular packing, the potential may be truncated at the position of its minimum, and shifted upwards to give what is usually termed the WCA model⁸. If electrostatic charges are present, we add the appropriate Coulomb potentials

$$v^{\text{Coulomb}}(r) = \frac{Q_1 Q_2}{4\pi\epsilon_0 r}, \quad (4)$$

where Q_1, Q_2 are the charges and ϵ_0 is the permittivity of free space. The correct handling of long-range forces in a simulation is an essential aspect of polyelectrolyte simulations, which will be the subject of the later chapter of Holm⁹.

2.2 Bonding Potentials

For molecular systems, we simply build the molecules out of site-site potentials of the form of Eq. (3) or similar. Typically, a single-molecule quantum-chemical calculation may be used to estimate the electron density throughout the molecule, which may then be modelled by a distribution of partial charges via Eq. (4), or more accurately by a distribution of electrostatic multipoles^{4,10}. For molecules we must also consider the intramolecular bonding interactions. The simplest molecular model will include terms of the following kind:

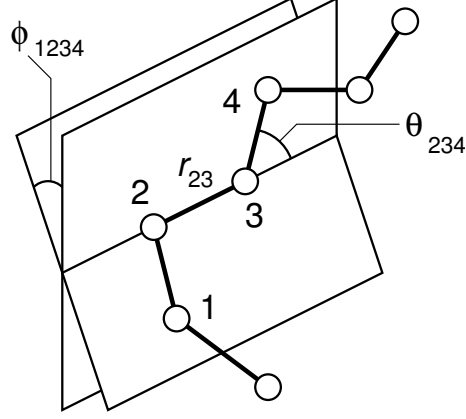


Figure 3. Geometry of a simple chain molecule, illustrating the definition of interatomic distance r_{23} , bend angle θ_{234} , and torsion angle ϕ_{1234} .

$$\mathcal{U}_{\text{intramolecular}} = \frac{1}{2} \sum_{\text{bonds}} k_{ij}^r (r_{ij} - r_{\text{eq}})^2 \quad (5a)$$

$$+ \frac{1}{2} \sum_{\text{bend angles}} k_{ijk}^\theta (\theta_{ijk} - \theta_{\text{eq}})^2 \quad (5b)$$

$$+ \frac{1}{2} \sum_{\text{torsion angles}} \sum_m k_{ijkl}^{\phi, m} (1 + \cos(m\phi_{ijkl} - \gamma_m)) \quad (5c)$$

The geometry is illustrated in Fig. 3. The “bonds” will typically involve the separation $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$ between adjacent pairs of atoms in a molecular framework, and we assume in Eq. (5a) a harmonic form with specified equilibrium separation, although this is not the only possibility. The “bend angles” θ_{ijk} are between successive bond vectors such as $\mathbf{r}_i - \mathbf{r}_j$ and $\mathbf{r}_j - \mathbf{r}_k$, and therefore involve three atom coordinates:

$$\cos \theta_{ijk} = \hat{\mathbf{r}}_{ij} \cdot \hat{\mathbf{r}}_{jk} = (\mathbf{r}_{ij} \cdot \mathbf{r}_{ij})^{-1/2} (\mathbf{r}_{jk} \cdot \mathbf{r}_{jk})^{-1/2} (\mathbf{r}_{ij} \cdot \mathbf{r}_{jk})$$

where $\hat{\mathbf{r}} = \mathbf{r}/r$. Usually this bending term is taken to be quadratic in the angular displacement from the equilibrium value, as in Eq. (5b), although periodic functions are also used. The “torsion angles” ϕ_{ijkl} are defined in terms of three connected bonds, hence four atomic coordinates:

$$\cos \phi_{ijkl} = -\hat{\mathbf{n}}_{ijk} \cdot \hat{\mathbf{n}}_{jkl}, \quad \text{where} \quad \mathbf{n}_{ijk} = \mathbf{r}_{ij} \times \mathbf{r}_{jk}, \quad \mathbf{n}_{jkl} = \mathbf{r}_{jk} \times \mathbf{r}_{kl},$$

and $\hat{\mathbf{n}} = \mathbf{n}/n$, the unit normal to the plane defined by each pair of bonds. Usually the torsional potential involves an expansion in periodic functions of order $m = 1, 2, \dots$, Eq. (5c).

A simulation package force-field will specify the precise form of Eq. (5), and the various strength parameters k and other constants therein. Actually, Eq. (5) is a considerable oversimplification. Molecular mechanics force-fields, aimed at accurately predicting

structures and properties, will include many cross-terms (e.g. stretch-bend): MM3^{11–13} and MM4^{14–16} are examples. Quantum mechanical calculations may give a guide to the “best” molecular force-field; also comparison of simulation results with thermophysical properties and vibration frequencies is invaluable in force-field development and refinement. A separate family of force fields, such as AMBER^{17,18}, CHARMM¹⁹ and OPLS²⁰ are geared more to larger molecules (proteins, polymers) in condensed phases; their functional form is simpler, closer to that of Eq. (5), and their parameters are typically determined by quantum chemical calculations combined with thermophysical and phase coexistence data. This field is too broad to be reviewed here; several molecular modelling texts^{21–23} (albeit targeted at biological applications) should be consulted by the interested reader. The modelling

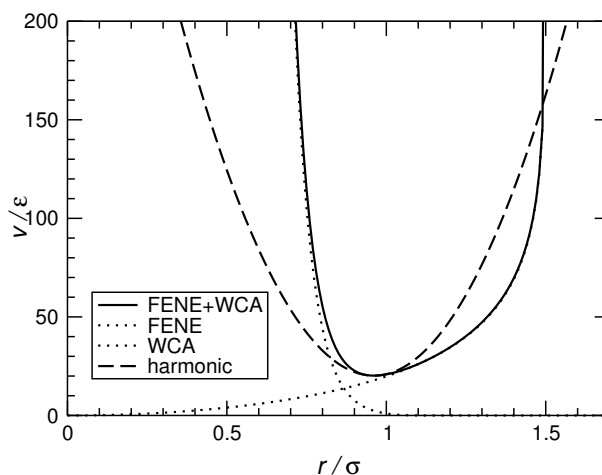


Figure 4. The FENE+WCA potential, with its separate FENE (attractive) and WCA (repulsive) components, between bonded atoms in a coarse-grained polymer chain. Also shown is the equivalent harmonic potential. Unlike the harmonic spring, the FENE potential cannot be extended beyond a specified limit, here $R_0 = 1.5\sigma$. For more details see Ref. 24.

of long chain molecules will be of particular interest to us, especially as an illustration of the scope for progressively simplifying and “coarse-graining” the potential model. Various explicit-atom potentials have been devised for the *n*-alkanes²⁵. More approximate potentials have also been constructed^{26–28} in which the CH₂ and CH₃ units are represented by single “united atoms”. These potentials are typically less accurate and less transferable than the explicit-atom potentials, but significantly less expensive; comparisons have been made between the two approaches²⁹. For more complicated molecules this approach may need to be modified. In the liquid crystal field, for instance, a compromise has been suggested³⁰: use the united-atom approach for hydrocarbon chains, but model phenyl ring hydrogens explicitly.

In polymer simulations, there is frequently a need to economize further and coarse-grain the interactions more dramatically: significant progress has been made in recent years in approaching this problem systematically^{31,32}. Finally, the most fundamental properties, such as the entanglement length in a polymer melt³³, may be investigated using a

simple chain of pseudo-atoms or beads (modelled using the WCA potential of Fig. 2, and each representing several monomers), joined by an attractive finitely-extensible non-linear elastic (FENE) potential²⁴ which is illustrated in Fig. 4.

$$v^{\text{FENE}}(r) = \begin{cases} -\frac{1}{2}kR_0^2 \ln(1 - (r/R_0)^2) & r < R_0 \\ \infty & r \geq R_0 \end{cases} \quad (6)$$

The key feature of this potential is that it cannot be extended beyond $r = R_0$, ensuring (for suitable choices of the parameters k and R_0) that polymer chains cannot move through one another.

2.3 Force Calculation

Having specified the potential energy function $\mathcal{U}(\mathbf{r}^N)$, the next step is to calculate the atomic forces

$$\mathbf{f}_i = -\frac{\partial}{\partial \mathbf{r}_i} \mathcal{U}(\mathbf{r}^N)$$

For site-site potentials this is a simple exercise. For the intramolecular part of the potential, it is a little more involved, but still a relatively straightforward application of the chain rule. Examples of how to do it are given in appendix C of Ref. 34. As a simple illustration, consider one of the bending potential terms for the polymer of Fig. 3, supposing that it may be written

$$v = -k \cos \theta_{234} = -k(\mathbf{r}_{23} \cdot \mathbf{r}_{23})^{-1/2}(\mathbf{r}_{34} \cdot \mathbf{r}_{34})^{-1/2}(\mathbf{r}_{23} \cdot \mathbf{r}_{34})$$

This will contribute to the forces on all three atoms. To calculate these, we need:

$$\begin{aligned} \frac{\partial}{\partial \mathbf{r}_2}(\mathbf{r}_{23} \cdot \mathbf{r}_{34}) &= \mathbf{r}_{34} & \frac{\partial}{\partial \mathbf{r}_3}(\mathbf{r}_{23} \cdot \mathbf{r}_{34}) &= \mathbf{r}_{23} - \mathbf{r}_{34} & \frac{\partial}{\partial \mathbf{r}_4}(\mathbf{r}_{23} \cdot \mathbf{r}_{34}) &= -\mathbf{r}_{23} \\ \frac{\partial}{\partial \mathbf{r}_2}(\mathbf{r}_{23} \cdot \mathbf{r}_{23}) &= 2\mathbf{r}_{23} & \frac{\partial}{\partial \mathbf{r}_3}(\mathbf{r}_{23} \cdot \mathbf{r}_{23}) &= -2\mathbf{r}_{23} & \frac{\partial}{\partial \mathbf{r}_4}(\mathbf{r}_{23} \cdot \mathbf{r}_{23}) &= 0 \\ \frac{\partial}{\partial \mathbf{r}_2}(\mathbf{r}_{34} \cdot \mathbf{r}_{34}) &= 0 & \frac{\partial}{\partial \mathbf{r}_3}(\mathbf{r}_{34} \cdot \mathbf{r}_{34}) &= 2\mathbf{r}_{34} & \frac{\partial}{\partial \mathbf{r}_4}(\mathbf{r}_{34} \cdot \mathbf{r}_{34}) &= -2\mathbf{r}_{34} \end{aligned}$$

and hence

$$\begin{aligned} \frac{\partial}{\partial \mathbf{r}_2} \cos \theta_{234} &= r_{23}^{-1} r_{34}^{-1} \left(\mathbf{r}_{34} - \frac{\mathbf{r}_{23} \cdot \mathbf{r}_{34}}{r_{23}^2} \mathbf{r}_{23} \right) \\ \frac{\partial}{\partial \mathbf{r}_3} \cos \theta_{234} &= r_{23}^{-1} r_{34}^{-1} \left(\frac{\mathbf{r}_{23} \cdot \mathbf{r}_{34}}{r_{23}^2} \mathbf{r}_{23} - \frac{\mathbf{r}_{23} \cdot \mathbf{r}_{34}}{r_{34}^2} \mathbf{r}_{34} + \mathbf{r}_{23} - \mathbf{r}_{34} \right) \\ \frac{\partial}{\partial \mathbf{r}_4} \cos \theta_{234} &= r_{23}^{-1} r_{34}^{-1} \left(\frac{\mathbf{r}_{23} \cdot \mathbf{r}_{34}}{r_{34}^2} \mathbf{r}_{34} - \mathbf{r}_{23} \right) \end{aligned}$$

A similar approach applied to the torsional potential gives the forces on all four involved atoms.

3 The MD Algorithm

Solving Newton’s equations of motion does not immediately suggest activity at the cutting edge of research. The molecular dynamics algorithm in most common use today may even have been known to Newton³⁵. Nonetheless, the last decade has seen a rapid development in our understanding of numerical algorithms; a forthcoming review³⁶ and a book³⁷ summarize the present state of the field.

Continuing to discuss, for simplicity, a system composed of atoms with coordinates $\mathbf{r}^N = (\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N)$ and potential energy $\mathcal{U}(\mathbf{r}^N)$, we introduce the atomic momenta $\mathbf{p}^N = (\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_N)$, in terms of which the kinetic energy may be written $\mathcal{K}(\mathbf{p}^N) = \sum_{i=1}^N |\mathbf{p}_i|^2 / 2m_i$. Then the energy, or hamiltonian, may be written as a sum of kinetic and potential terms $\mathcal{H} = \mathcal{K} + \mathcal{U}$. Write the classical equations of motion as

$$\dot{\mathbf{r}}_i = \mathbf{p}_i / m_i \quad \text{and} \quad \dot{\mathbf{p}}_i = \mathbf{f}_i \quad (7)$$

This is a system of coupled ordinary differential equations. Many methods exist to perform step-by-step numerical integration of them. Characteristics of these equations are: (a) they are ‘stiff’, i.e. there may be short and long timescales, and the algorithm must cope with both; (b) calculating the forces is expensive, typically involving a sum over pairs of atoms, and should be performed as infrequently as possible. Also we must bear in mind that the advancement of the coordinates fulfils two functions: (i) accurate calculation of dynamical properties, especially over times as long as typical correlation times τ_a of properties a of interest (we shall define this later); (ii) accurately staying on the constant-energy hypersurface, for much longer times $\tau_{\text{run}} \gg \tau_a$, in order to sample the correct ensemble.

To ensure rapid sampling of phase space, we wish to make the timestep as large as possible consistent with these requirements. For these reasons, simulation algorithms have tended to be of *low order* (i.e. they do not involve storing high derivatives of positions, velocities etc.): this allows the time step to be increased as much as possible without jeopardizing energy conservation. It is unrealistic to expect the numerical method to accurately follow the true trajectory for very long times τ_{run} . The ‘ergodic’ and ‘mixing’ properties of classical trajectories, i.e. the fact that nearby trajectories diverge from each other exponentially quickly, make this impossible to achieve.

All these observations tend to favour the Verlet algorithm in one form or another, and we look closely at this in the following section. For historical reasons only, we mention the more general class of predictor-corrector methods which have been optimized for classical mechanical equations^{38,39}; further details are available elsewhere^{34,40,41}.

3.1 The Verlet Algorithm

There are various, essentially equivalent, versions of the Verlet algorithm, including the original method^{7,42} and a ‘leapfrog’ form⁴³. Here we concentrate on the ‘velocity Verlet’ algorithm⁴⁴, which may be written

$$\mathbf{p}_i(t + \tfrac{1}{2}\delta t) = \mathbf{p}_i(t) + \tfrac{1}{2}\delta t \mathbf{f}_i(t) \quad (8a)$$

$$\mathbf{r}_i(t + \delta t) = \mathbf{r}_i(t) + \delta t \mathbf{p}_i(t + \tfrac{1}{2}\delta t) / m_i \quad (8b)$$

$$\mathbf{p}_i(t + \delta t) = \mathbf{p}_i(t + \tfrac{1}{2}\delta t) + \tfrac{1}{2}\delta t \mathbf{f}_i(t + \delta t) \quad (8c)$$

After step (8b), a force evaluation is carried out, to give $\mathbf{f}_i(t + \delta t)$ for step (8c). This scheme advances the coordinates and momenta over a timestep δt . A piece of pseudo-code illustrates how this works:

```
do step = 1, nstep
  p = p + 0.5*dt*f
  r = r + dt*p/m
  f = force(r)
  p = p + 0.5*dt*f
enddo
```

As we shall see shortly there is an interesting theoretical derivation of this version of the algorithm. Important features of the Verlet algorithm are: (a) it is *exactly* time reversible; (b) it is *symplectic* (to be discussed shortly); (c) it is *low* order in time, hence permitting long timesteps; (d) it requires just one (expensive) force evaluation per step; (e) it is easy to program.

3.2 Constraints

It is quite common practice in classical computer simulations *not* to attempt to represent intramolecular bonds by terms in the potential energy function, because these bonds have very high vibration frequencies (and arguably should be treated in a quantum mechanical way rather than in the classical approximation). Instead, the bonds are treated as being constrained to have fixed length. In classical mechanics, constraints are introduced through the Lagrangian⁴⁵ or Hamiltonian⁴⁶ formalisms. Given an algebraic relation between two atomic coordinates, for example a fixed bond length b between atoms 1 and 2, one may write a constraint equation, plus an equation for the time derivative of the constraint

$$\chi(\mathbf{r}_1, \mathbf{r}_2) = (\mathbf{r}_1 - \mathbf{r}_2) \cdot (\mathbf{r}_1 - \mathbf{r}_2) - b^2 = 0 \quad (9a)$$

$$\dot{\chi}(\mathbf{r}_1, \mathbf{r}_2) = 2(\mathbf{v}_1 - \mathbf{v}_2) \cdot (\mathbf{r}_1 - \mathbf{r}_2) = 0. \quad (9b)$$

In the Lagrangian formulation, the constraint forces acting on the atoms will enter thus:

$$m_i \ddot{\mathbf{r}}_i = \mathbf{f}_i + \Lambda \mathbf{g}_i$$

where Λ is the undetermined multiplier and

$$\mathbf{g}_1 = -\frac{\partial \chi}{\partial \mathbf{r}_1} = -2(\mathbf{r}_1 - \mathbf{r}_2) \quad \mathbf{g}_2 = -\frac{\partial \chi}{\partial \mathbf{r}_2} = 2(\mathbf{r}_1 - \mathbf{r}_2)$$

It is easy to derive an exact expression for the multiplier Λ from the above equations; if several constraints are imposed, a system of equations (one per constraint) is obtained. However, this exact solution is not what we want: in practice, since the equations of motion are only solved approximately, in discrete time steps, the constraints will be increasingly violated as the simulation proceeds. The breakthrough in this area came with the proposal to determine the constraint forces in such a way that the constraints are satisfied *exactly* at the end of each time step⁴⁷⁻⁴⁹. For the original Verlet algorithm, this scheme is called SHAKE. The appropriate version of this scheme for the velocity Verlet algorithm is called RATTLE⁵⁰. Formally, we wish to solve the following scheme, in which we combine

$(\mathbf{r}_1, \mathbf{r}_2)$ into $\mathbf{r}$, $(\mathbf{p}_1, \mathbf{p}_2)$ into $\mathbf{p}$, etc. for simplicity:

$$\begin{aligned}\mathbf{p}(t + \tfrac{1}{2}\delta t) &= \mathbf{p}(t) + \tfrac{1}{2}\delta t \mathbf{f}(t) + \lambda \mathbf{g}(t) \\ \mathbf{r}(t + \delta t) &= \mathbf{r}(t) + \delta t \mathbf{p}(t + \tfrac{1}{2}\delta t)/m \\ \text{choosing } \lambda \text{ such that: } 0 &= \chi(\mathbf{r}(t + \delta t))\end{aligned}\tag{10a}$$

$$\begin{aligned}\mathbf{p}(t + \delta t) &= \mathbf{p}(t + \tfrac{1}{2}\delta t) + \tfrac{1}{2}\delta t \mathbf{f}(t + \delta t) + \mu \mathbf{g}(t + \delta t) \\ \text{choosing } \mu \text{ such that: } 0 &= \dot{\chi}(\mathbf{r}(t + \delta t), \mathbf{p}(t + \delta t))\end{aligned}\tag{10b}$$

Step (10a) may be implemented by defining *unconstrained* variables

$$\bar{\mathbf{p}}(t + \tfrac{1}{2}\delta t) = \mathbf{p}(t) + \tfrac{1}{2}\delta t \mathbf{f}(t), \quad \bar{\mathbf{r}}(t + \delta t) = \mathbf{r}(t) + \delta t \bar{\mathbf{p}}(t + \tfrac{1}{2}\delta t)/m$$

then solving the nonlinear equation for λ

$$\chi(t + \delta t) = \chi(\bar{\mathbf{r}}(t + \delta t) + \lambda \delta t \mathbf{g}(t)/m) = 0$$

and substituting back

$$\mathbf{p}(t + \tfrac{1}{2}\delta t) = \bar{\mathbf{p}}(t + \tfrac{1}{2}\delta t) + \lambda \mathbf{g}(t), \quad \mathbf{r}(t + \delta t) = \bar{\mathbf{r}}(t + \delta t) + \delta t \lambda \mathbf{g}(t)/m$$

Step (10b) may be handled by defining

$$\bar{\mathbf{p}}(t + \delta t) = \mathbf{p}(t + \tfrac{1}{2}\delta t) + \tfrac{1}{2}\delta t \mathbf{f}(t + \delta t)$$

solving the equation for the second Lagrange multiplier μ

$$\dot{\chi}(t + \delta t) = \dot{\chi}(\mathbf{r}(t + \delta t), \bar{\mathbf{p}}(t + \delta t) + \mu \mathbf{g}(t + \delta t)) = 0$$

(which is actually linear, since $\dot{\chi}(\mathbf{r}, \mathbf{p}) = -\mathbf{g}(\mathbf{r}) \cdot \mathbf{p}/m$) and substituting back

$$\mathbf{p}(t + \delta t) = \bar{\mathbf{p}}(\delta t) + \mu \mathbf{g}(t + \delta t)$$

In pseudo-code this scheme may be written

```
do step = 1, nstep
  p = p + (dt/2)*f
  r = r + dt*p/m
  lambda_g = shake(r)
  p = p + lambda_g
  r = r + dt*lambda_g/m
  f = force(r)
  p = p + (dt/2)*f
  mu_g = rattle(r, p)
  p = p + mu_g
enddo
```

The routine called `shake` here calculates the constraint forces $\lambda \mathbf{g}_i$ necessary to ensure that the end-of-step positions $\mathbf{r}_i$ satisfy Eq. (9a). For a system of many constraints, this calculation is usually performed in an iterative fashion, so as to satisfy each constraint in turn until convergence; the original SHAKE algorithm was framed in this way. These constraint forces are incorporated into both the end-of-step positions and the mid-step momenta. The routine called `rattle` calculates a new set of constraint forces $\mu \mathbf{g}_i$ to ensure that the end-of-step momenta satisfy Eq. (9b). This also may be carried out iteratively.

It is important to realize that a simulation of a system with rigidly constrained bond lengths, is not equivalent to a simulation with, for example, harmonic springs representing the bonds, even in the limit of very strong springs. A subtle, but crucial, difference lies in the distribution function for the other coordinates. If we obtain the configurational distribution function by integrating over the momenta, the difference arises because in one case a set of momenta is set to zero, and not integrated, while in the other an integration is performed, which may lead to an extra term depending on particle coordinates. This is frequently called the ‘metric tensor problem’; it is explained in more detail in the references^{34,51}, and there are well-established ways of determining when the difference is likely to be significant⁵² and how to handle it, if necessary⁵³. Constraints also find an application in the study of rare events⁵¹, or for convenience when it is desired to fix, for example, the director in a liquid crystal simulation⁵⁴.

An alternative to constraints, is to retain the intramolecular bond potentials and use a multiple time step approach to handle the fast degrees of freedom. We discuss this shortly.

3.3 Periodic Boundary Conditions

Small sample size means that, unless surface effects are of particular interest, periodic boundary conditions need to be used. Consider 1000 atoms arranged in a $10 \times 10 \times 10$ cube. Nearly half the atoms are on the outer faces, and these will have a large effect on the measured properties. Even for $10^6 = 100^3$ atoms, the surface atoms amount to 6% of the total, which is still nontrivial. Surrounding the cube with replicas of itself takes care of this problem. Provided the potential range is not too long, we can adopt the *minimum image convention* that each atom interacts with the nearest atom or image in the periodic array. In the course of the simulation, if an atom leaves the basic simulation box, attention can be switched to the incoming image. This is shown in Figure 5. Of course, it is important to bear in mind the imposed artificial periodicity when considering properties which are influenced by long-range correlations. Special attention must be paid to the case where the potential range is not short: for example for charged and dipolar systems.

3.4 Neighbour Lists

Computing the non-bonded contribution to the interatomic forces in an MD simulation involves, in principle, a large number of pairwise calculations: we consider each atom i and loop over all other atoms j to calculate the minimum image separations r_{ij} . Let us assume that the interaction potentials are of short range, $v(r_{ij}) = 0$ if $r_{ij} > r_{\text{cut}}$, the potential cutoff. In this case, the program skips the force calculation, avoiding expensive calculations, and considers the next candidate j . Nonetheless, the time to examine all pair separations is proportional to the number of distinct pairs, $\frac{1}{2}N(N-1)$ in an N -atom system, and for every pair one must compute at least r_{ij}^2 ; this still consumes a lot of time.

Some economies result from the use of lists of nearby pairs of atoms. Verlet⁷ suggested such a technique for improving the speed of a program. The potential cutoff sphere, of radius r_{cut} , around a particular atom is surrounded by a ‘skin’, to give a larger sphere of radius r_{list} as shown in Figure 6. At the first step in a simulation, a list is constructed of all the neighbours of each atom, for which the pair separation is within r_{list} . Over the next few MD time steps, only pairs appearing in the list are checked in the force routine. From

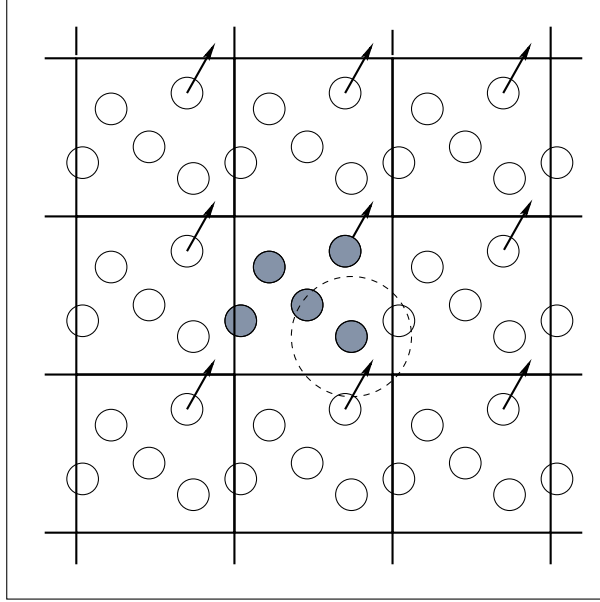


Figure 5. Periodic boundary conditions. As a particle moves out of the simulation box, an image particle moves in to replace it. In calculating particle interactions within the cutoff range, both real and image neighbours are included.

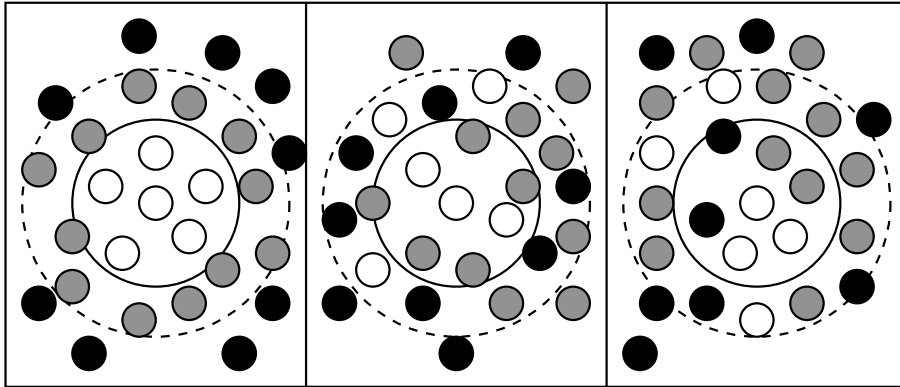


Figure 6. The Verlet list on its construction, later, and too late. The potential cutoff range (solid circle), and the list range (dashed circle), are indicated. The list must be reconstructed before particles originally outside the list range (black) have penetrated the potential cutoff sphere.

time to time the list is reconstructed: it is important to do this before any unlisted pairs have crossed the safety zone and come within interaction range. It is possible to trigger the list reconstruction automatically, if a record is kept of the distance travelled by each atom since the last update. The choice of list cutoff distance r_{list} is a compromise: larger lists

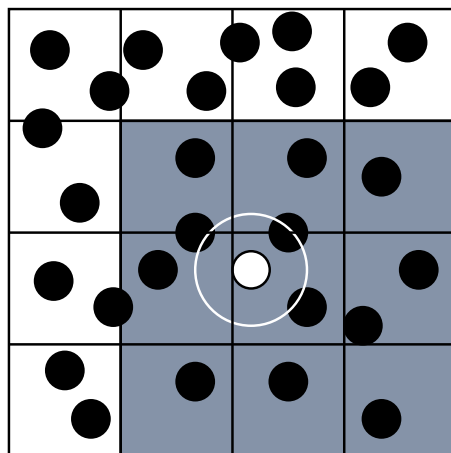


Figure 7. The cell structure. The potential cutoff range is indicated. In searching for neighbours of an atom, it is only necessary to examine the atom's own cell, and its nearest-neighbour cells (shaded).

will need to be reconstructed less frequently, but will not give as much of a saving on cpu time as smaller lists. This choice can easily be made by experimentation.

For larger systems ($N \geq 1000$ or so, depending on the potential range) another technique becomes preferable. The cubic simulation box (extension to non-cubic cases is possible) is divided into a regular lattice of $n_{\text{cell}} \times n_{\text{cell}} \times n_{\text{cell}}$ cells; see Figure 7. These cells are chosen so that the side of the cell $\ell_{\text{cell}} = L/n_{\text{cell}}$ is greater than the potential cutoff distance r_{cut} . If there is a separate list of atoms in each of those cells, then searching through the neighbours is a rapid process: it is only necessary to look at atoms in the same cell as the atom of interest, and in nearest neighbour cells. The cell structure may be set up and used by the method of linked lists^{55,43}. The first part of the method involves sorting all the atoms into their appropriate cells. This sorting is rapid, and may be performed every step. Then, within the force routine, pointers are used to scan through the contents of cells, and calculate pair forces. This approach is very efficient for large systems with short-range forces. A certain amount of unnecessary work is done because the search region is cubic, not (as for the Verlet list) spherical.

4 Time Dependence

A knowledge of time-dependent statistical mechanics is important in three general areas of simulation. Firstly, in recent years there have been significant advances in the understanding of molecular dynamics algorithms, which have arisen out of an appreciation of the formal operator approach to classical mechanics. Second, an understanding of equilibrium time correlation functions, their link with dynamical properties, and especially their connection with transport coefficients, is essential in making contact with experiment. Third, the last decade has seen a rapid development of the use of nonequilibrium molecular dynamics, with a better understanding of the formal aspects, particularly the link between the dynamical algorithm, dissipation, chaos, and fractal geometry. Space does not permit a full

description of all these topics here: the interested reader should consult Refs 56,57,51 and references therein.

The Liouville equation dictates how the classical statistical mechanical distribution function $\varrho(\mathbf{r}^N, \mathbf{p}^N, t)$ evolves in time. From considerations of standard, Hamiltonian, mechanics⁴⁵ and the flow of representative systems in an ensemble through a particular region of phase space, it is easy to derive the Liouville equation

$$\frac{\partial \varrho}{\partial t} = - \left\{ \sum_i \dot{\mathbf{r}}_i \cdot \frac{\partial}{\partial \mathbf{r}_i} + \dot{\mathbf{p}}_i \cdot \frac{\partial}{\partial \mathbf{p}_i} \right\} \varrho \equiv -i\mathcal{L}\varrho, \quad (11)$$

defining the Liouville operator $i\mathcal{L}$ as the quantity in braces. Contrast this equation for ϱ with the time evolution equation for a dynamical variable $A(\mathbf{r}^N, \mathbf{p}^N)$, which comes directly from the chain rule applied to Hamilton's equations

$$\dot{A} = \sum_i \dot{\mathbf{r}}_i \cdot \frac{\partial A}{\partial \mathbf{r}_i} + \dot{\mathbf{p}}_i \cdot \frac{\partial A}{\partial \mathbf{p}_i} \equiv i\mathcal{L}A. \quad (12)$$

The formal solutions of the time evolution equations are

$$\varrho(t) = e^{-i\mathcal{L}t} \varrho(0) \quad \text{and} \quad A(t) = e^{i\mathcal{L}t} A(0) \quad (13)$$

where, in either case, the exponential operator is called the *propagator*. A number of manipulations are possible, once this formalism has been established. There are useful analogies both with the Eulerian and Lagrangian pictures of incompressible fluid flow, and with the Heisenberg and Schrödinger pictures of quantum mechanics (see e.g. Ref. 58, Chap. 7, and Ref. 59, Chap. 11). These analogies are particularly useful in formulating the equations of classical response theory⁶⁰, linking transport coefficients with both equilibrium and nonequilibrium simulations⁵⁶.

The Liouville equation applies to any ensemble, equilibrium or not. Equilibrium means that ϱ should be *stationary*, i.e. that $\partial \varrho / \partial t = 0$. In other words, if we look at any phase-space volume element, the rate of incoming state points should equal the rate of outflow. This requires that ϱ be a function of the constants of the motion, and especially $\varrho = \varrho(\mathcal{H})$. Equilibrium also implies $d\langle A \rangle / dt = 0$ for any A . The extension of the above equations to nonequilibrium ensembles requires a consideration of entropy production, the method of controlling energy dissipation (thermostatting) and the consequent non-Liouville nature of the time evolution⁵⁶.

4.1 Propagators and the Verlet Algorithm

The velocity Verlet algorithm may be derived by considering a standard approximate decomposition of the Liouville operator which preserves reversibility and is *symplectic* (which implies that volume in phase space is conserved). This approach⁶¹ has had several beneficial consequences.

The Liouville operator of equation (13) may be written⁶²

$$e^{i\mathcal{L}t} = \left(e^{i\mathcal{L}\delta t} \right)_{\text{approx}}^{n_{\text{step}}} + \mathcal{O}(n_{\text{step}}\delta t^3)$$

where $\delta t = t/n_{\text{step}}$ and an approximate propagator, correct at short timesteps $\delta t \rightarrow 0$, appears in the parentheses. This is a formal way of stating what we do in molecular dynamics, when we split a long time period t into a large number n_{step} of small timesteps δt ,

using an *approximation* to the true equations of motion over each timestep. It turns out that useful approximations arise from splitting iL into two parts⁶¹

$$iL = iL_p + iL_r . \quad (14)$$

where

$$iL_p = \sum_i \dot{\mathbf{p}}_i \cdot \frac{\partial}{\partial \mathbf{p}_i} = \sum_i \mathbf{f}_i \cdot \frac{\partial}{\partial \mathbf{p}_i} \quad (15a)$$

$$iL_r = \sum_i \dot{\mathbf{r}}_i \cdot \frac{\partial}{\partial \mathbf{r}_i} = \sum_i m_i^{-1} \mathbf{p}_i \cdot \frac{\partial}{\partial \mathbf{r}_i} . \quad (15b)$$

The following approximation

$$e^{iL\delta t} = e^{(iL_p + iL_r)\delta t} \approx e^{iL_p\delta t/2} e^{iL_r\delta t} e^{iL_p\delta t/2} \quad (16)$$

is asymptotically exact in the limit $\delta t \rightarrow 0$. For nonzero δt this is an approximation to $e^{iL\delta t}$ because in general iL_p and iL_r do not commute, but it is still exactly time reversible and symplectic.

Effectively, we are propagating the equation of motion in steps which ignore, in turn, the kinetic part and the potential part of the hamiltonian. A straightforward derivation⁵¹ shows that the effect of each operator on a dynamical variable $A(\mathbf{r}^N, \mathbf{p}^N)$ is to advance, respectively, the coordinates and the momenta separately:

$$e^{iL_r\delta t} A(\mathbf{r}, \mathbf{p}) = A(\mathbf{r} + m^{-1}\mathbf{p}\delta t, \mathbf{p}) \quad (17a)$$

$$e^{iL_p\delta t} A(\mathbf{r}, \mathbf{p}) = A(\mathbf{r}, \mathbf{p} + \mathbf{f}\delta t) \quad (17b)$$

where, to avoid clutter, we have written $\mathbf{r}, \mathbf{p}$ for $\mathbf{r}^N, \mathbf{p}^N$. It is then easy to see that the three successive steps embodied in equation (16), with the above choice of operators, generate the velocity Verlet algorithm.

Such a formal approach may seem somewhat abstract, but has been invaluable in understanding the excellent performance of Verlet-like algorithms in molecular dynamics, and in extending the range of algorithms available to us. It may be shown that, although the trajectories generated by the above scheme are approximate, and will not conserve the true energy $\mathcal{H}$, nonetheless, they do exactly conserve a “pseudo-hamiltonian” or “shadow hamiltonian” $\overline{\mathcal{H}}$ which differs from the true one by a small amount (vanishing as $\delta t \rightarrow 0$). This means that no drift in the energy will occur: the system will remain on a hypersurface in phase space which is “close” (in the above sense) to the true constant-energy hypersurface. Such a stability property is extremely useful in molecular dynamics, since we wish to sample constant-energy states.

An example may make things clearer. Consider a simple harmonic oscillator⁶³, of natural frequency ω , representing perhaps an interatomic bond in a diatomic molecule. The equations of motion are

$$\dot{x} = p/m \quad \dot{p} = -m\omega^2 x$$

For these equations, a few lines of algebra shows that the following shadow hamiltonian

$$\overline{\mathcal{H}}(x, p) = \frac{p^2/2m}{1 - (\omega\delta t/2)^2} + \frac{1}{2}m\omega^2 x^2$$

is *exactly* conserved by the velocity Verlet algorithm. In a phase portrait, the simulated system remains on a constant- $\bar{\mathcal{H}}$ ellipse which differs only slightly (for small $\omega\delta t$) from the true constant- $\mathcal{H}$ ellipse. This is illustrated in Fig. 8, where we deliberately choose a high step size $\delta t = 0.7/\omega$ to accentuate the differences. Note that, even for this value of δt , the energy conservation is very good (the deviation is $\mathcal{O}(\delta t^2)$). At the same time, the frequency of oscillation is shifted from the true value, so the trajectories steadily diverge from each other.

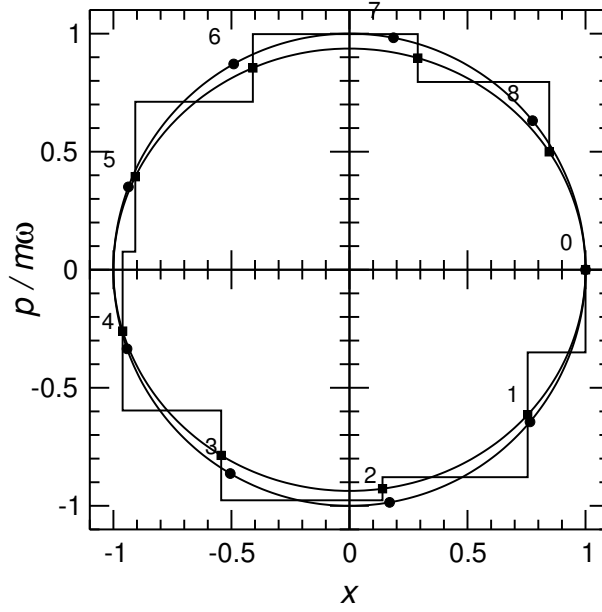


Figure 8. Velocity Verlet algorithm for simple harmonic oscillator with initial conditions $x(0) = 1, p(0) = 0$. The outer circle shows the exact trajectory, conserving the true hamiltonian $\mathcal{H}$; the inner ellipse is a contour of constant shadow hamiltonian $\bar{\mathcal{H}}$ for a (relatively large) timestep $\delta t = 0.7/\omega$. The circles show the exact solutions at regular intervals δt ; the squares show the corresponding velocity Verlet phase points, connected by straight sections representing the intermediate steps in the algorithm.

4.2 Multiple Timesteps

An important extension of the MD method allows it to tackle systems with multiple time scales: for example, molecules which have very strong internal springs representing the bonds, while interacting externally through softer potentials; molecules having strongly-varying short-range interactions but more smoothly-varying long-range interactions; or perhaps molecules consisting of both heavy and light atoms. A simple MD algorithm will have to adopt a timestep short enough to handle the fastest variables.

An attractive improvement is to handle the fast motions with a shorter timestep^{64, 61, 65}. A time-reversible Verlet-like multiple-timestep algorithm may be generated using the Liouville operator formalism described above. Here we suppose that there are two types of

force in the system: “slow” F_i , and “fast” f_i . The momentum satisfies $\dot{\mathbf{p}}_i = \mathbf{f}_i + \mathbf{F}_i$. Then we break up the Liouville operator $i\mathbf{L} = i\mathbf{L}_p + i\mathbf{L}_p + i\mathbf{L}_r$:

$$i\mathbf{L}_p = \sum_i \mathbf{F}_i \cdot \frac{\partial}{\partial \mathbf{p}_i} \quad (18a)$$

$$i\mathbf{L}_p = \sum_i \mathbf{f}_i \cdot \frac{\partial}{\partial \mathbf{p}_i} \quad (18b)$$

$$i\mathbf{L}_r = \sum_i m^{-1} \mathbf{p}_i \cdot \frac{\partial}{\partial \mathbf{r}_i} \quad (18c)$$

The propagator approximately factorizes

$$e^{i\mathbf{L}\Delta t} \approx e^{i\mathbf{L}_p\Delta t/2} e^{i(\mathbf{L}_p+\mathbf{L}_r)\Delta t} e^{i\mathbf{L}_p\Delta t/2}$$

where Δt represents a long time step. The middle part is then split again, using the conventional separation, and iterating over small time steps $\delta t = \Delta t/n_{\text{step}}$:

$$e^{i(\mathbf{L}_p+\mathbf{L}_r)\Delta t} \approx \left[e^{i\mathbf{L}_p\delta t/2} e^{i\mathbf{L}_r\delta t} e^{i\mathbf{L}_p\delta t/2} \right]^{n_{\text{step}}}.$$

So the fast-varying forces must be computed many times at short intervals; the slow-varying forces are used just before and just after this stage, and they only need be calculated once per long timestep.

This translates into a fairly simple algorithm, based closely on the standard velocity Verlet method. Written in a Fortran-like pseudo-code, it is as follows. At the start of the run we calculate both rapidly-varying ($\mathbf{f}$) and slowly-varying ($\mathbf{F}$) forces, then, in the main loop:

```

do STEP = 1, NSTEP
  p = p + (DT/2)*F
  do step = 1, nstep
    p = p + (dt/2)*f
    r = r + dt*p/m
    f = force(r)
    p = p + (dt/2)*f
  enddo
  F = FORCE(r)
  p = p + (DT/2)*F
enddo

```

The entire simulation run consists of NSTEP long steps; each step consists of nstep shorter sub-steps. DT and dt are the corresponding timesteps, $DT = nstep*dt$.

A particularly fruitful application, which has been incorporated into computer codes such as ORAC⁶⁶, is to split the interatomic force law into a succession of components covering different ranges: the short-range forces change rapidly with time and require a short time step, but advantage can be taken of the much slower time variation of the long-range forces, by using a longer time step and less frequent evaluation for these. Having said this, multiple-time-step algorithms are still under active study⁶⁷, and there is some concern that resonances may occur between the natural frequencies of the system and the

various timesteps used in schemes of this kind^{68,69}. Significant efforts have been made in recent years to overcome these problems and achieve significant increases in step size by alternative methods^{70–72}. The area remains one of active research^{73,36}.

5 Rigid Molecule Rotation

In certain applications, particularly in the simulation of liquid crystals, colloidal systems, and polymers, it is advantageous to include non-spherical rigid bodies in the molecular model. This means that we must calculate intermolecular torques as well as forces, and implement the classical dynamical equations for rotational motion.

If the intermolecular forces are expressed as sums of site-site (or atom-atom) terms, the conversion of these into centre-of-mass forces, and torques about the centre of mass, is easily performed. Consider two molecules A and B, centre-of-mass position vectors $\mathbf{R}_A$, $\mathbf{R}_B$. Define the intermolecular vector $\mathbf{R}_{AB} = \mathbf{R}_A - \mathbf{R}_B$, and suppose that the interaction potential may be expressed

$$v_{AB} = \sum_{i \in A} \sum_{j \in B} v(r_{ij})$$

where i and j are atomic sites in the respective molecules. Then we may compute

$$\text{force on } A \text{ due to } B: \quad \mathbf{F}_{AB} = \sum_{i \in A} \sum_{j \in B} \mathbf{f}_{ij} = -\mathbf{F}_{BA}$$

$$\text{torque on } A \text{ due to } B: \quad \mathbf{N}_{AB} = \sum_{i \in A} \sum_{j \in B} \mathbf{r}_{iA} \times \mathbf{f}_{ij}$$

$$\text{torque on } B \text{ due to } A: \quad \mathbf{N}_{BA} = \sum_{i \in A} \sum_{j \in B} \mathbf{r}_{jB} \times \mathbf{f}_{ji}$$

where $\mathbf{r}_{iA} = \mathbf{r}_i - \mathbf{R}_A$ is the position of site i relative to the centre of molecule A (and likewise for $\mathbf{r}_{jB}$). Note that $\mathbf{N}_{AB} \neq -\mathbf{N}_{BA}$ (a common misconception), but that the above equations give directly

$$\mathbf{N}_{AB} + \mathbf{N}_{BA} + \mathbf{R}_{AB} \times \mathbf{F}_{AB} = 0 \quad (19)$$

provided the forces satisfy $\mathbf{f}_{ij} = -\mathbf{f}_{ji}$ and act along the site-site vectors $\mathbf{r}_{ij}$. This is the expression of local angular momentum conservation, which follows directly from the invariance of the potential energy v_{AB} to a rotation of the coordinate system. (Note, in passing, that, in periodic boundaries, angular momentum is not globally conserved).

There is also a trend to use rigid-body potentials which are defined explicitly in terms of centre-of-mass positions and molecular orientations. An example is the Gay-Berne potential⁷⁴

$$v_{AB}^{\text{GB}}(\mathbf{R}, \hat{\mathbf{a}}, \hat{\mathbf{b}}) = 4\varepsilon(\hat{\mathbf{R}}, \hat{\mathbf{a}}, \hat{\mathbf{b}}) [\varrho^{-12} - \varrho^{-6}] \quad (20a)$$

$$\text{with} \quad \varrho = \frac{R - \sigma(\hat{\mathbf{R}}, \hat{\mathbf{a}}, \hat{\mathbf{b}}) + \sigma_0}{\sigma_0} \quad (20b)$$

which depends upon the molecular axis vectors $\hat{\mathbf{a}}$ and $\hat{\mathbf{b}}$, and on the direction $\hat{\mathbf{R}}$ and magnitude R of the centre-centre vector $\mathbf{R}_{AB}$, which we write $\mathbf{R}$ here and henceforth. The

parameter σ_0 determines the smallest molecular diameter and there are two orientation-dependent quantities in the above shifted-Lennard-Jones form: a diameter $\sigma(\hat{\mathbf{R}}, \hat{\mathbf{a}}, \hat{\mathbf{b}})$ and an energy $\varepsilon(\hat{\mathbf{R}}, \hat{\mathbf{a}}, \hat{\mathbf{b}})$. Each quantity depends in a complicated way (not given here) on parameters characterizing molecular shape and structure. This potential has been extensively used in the study of molecular liquids and liquid crystals^{75,76,54,77-79} and will be discussed further in a later chapter⁸⁰. Generalizations to non-uniaxial rigid bodies⁸¹⁻⁸³ have also been studied: here, the diameter and energy parameters depend on the full orthogonal orientation matrices $\hat{\mathbf{a}}, \hat{\mathbf{b}}$, which convert from space-fixed (xyz) to molecule-fixed (123) coordinates, and whose rows are the three molecule-fixed orthonormal principal axis vectors $\hat{\mathbf{a}}_\alpha, \hat{\mathbf{b}}_\beta$, $\alpha, \beta = 1, 2, 3$.

We go through the following derivation in some detail, as it is seldom presented. A very common case is when the pair potential may be written in the form⁸⁴

$$v_{AB} = v_{AB}(R, \{\hat{\mathbf{a}}_\alpha \cdot \hat{\mathbf{R}}\}, \{\hat{\mathbf{b}}_\beta \cdot \hat{\mathbf{R}}\}, \{\hat{\mathbf{a}}_\alpha \cdot \hat{\mathbf{b}}_\beta\}) \quad (21)$$

i.e. a function of the centre-centre separation R , and all possible scalar products of the unit vectors $\hat{\mathbf{R}}, \hat{\mathbf{a}}_\alpha$ and $\hat{\mathbf{b}}_\beta$. Using the chain rule, we may write the force on A:

$$\begin{aligned} \mathbf{F}_{AB} &= -\frac{\partial v_{AB}}{\partial \mathbf{R}} = -\frac{\partial v_{AB}}{\partial R} \frac{\partial R}{\partial \mathbf{R}} - \sum_{\hat{\mathbf{e}}=\hat{\mathbf{a}}, \hat{\mathbf{b}}} \frac{\partial v_{AB}}{\partial(\hat{\mathbf{e}} \cdot \hat{\mathbf{R}})} \frac{\partial(\hat{\mathbf{e}} \cdot \hat{\mathbf{R}})}{\partial \mathbf{R}} \\ &= -\frac{\partial v_{AB}}{\partial R} \hat{\mathbf{R}} - \sum_{\hat{\mathbf{e}}=\hat{\mathbf{a}}, \hat{\mathbf{b}}} \frac{\partial v_{AB}}{\partial(\hat{\mathbf{e}} \cdot \hat{\mathbf{R}})} \frac{\hat{\mathbf{e}} - (\hat{\mathbf{e}} \cdot \hat{\mathbf{R}}) \hat{\mathbf{R}}}{R}. \end{aligned} \quad (22)$$

The sum ranges over all the orientation vectors on both molecules, $\hat{\mathbf{e}} = \{\hat{\mathbf{a}}_\alpha, \hat{\mathbf{b}}_\beta\}$. The derivatives of the potential are easily evaluated, assuming that it has the general form of Eq. (21). To calculate the torques, we follow the general approach of Ref. 84. Consider the derivative of v_{AB} with respect to rotation of molecule A through an angle ψ about any axis $\hat{\mathbf{n}}$. By definition, this gives:

$$\hat{\mathbf{n}} \cdot \mathbf{N}_{AB} = -\frac{\partial v_{AB}}{\partial \psi} = -\sum_{\alpha} \sum_{\hat{\mathbf{e}}=\hat{\mathbf{R}}, \hat{\mathbf{b}}} \frac{\partial v_{AB}}{\partial(\hat{\mathbf{e}} \cdot \hat{\mathbf{a}}_\alpha)} \frac{\partial(\hat{\mathbf{e}} \cdot \hat{\mathbf{a}}_\alpha)}{\partial \psi}. \quad (23)$$

The sum is over all combinations of unit vectors $(\hat{\mathbf{a}}_\alpha, \hat{\mathbf{e}})$ for which one, $\hat{\mathbf{a}}_\alpha$, rotates with the molecule while the other, $\hat{\mathbf{e}} = \hat{\mathbf{R}}$ or $\hat{\mathbf{b}}_\beta$, remains stationary. This has the effect⁴⁵

$$\frac{\partial \hat{\mathbf{a}}_\alpha}{\partial \psi} = \hat{\mathbf{n}} \times \hat{\mathbf{a}}_\alpha \quad \Rightarrow \quad \frac{\partial(\hat{\mathbf{e}} \cdot \hat{\mathbf{a}}_\alpha)}{\partial \psi} = \hat{\mathbf{e}} \cdot \hat{\mathbf{n}} \times \hat{\mathbf{a}}_\alpha = -\hat{\mathbf{n}} \cdot \hat{\mathbf{e}} \times \hat{\mathbf{a}}_\alpha.$$

Then Eq. (23) gives

$$\hat{\mathbf{n}} \cdot \mathbf{N}_{AB} = \hat{\mathbf{n}} \cdot \sum_{\alpha} \sum_{\hat{\mathbf{e}}=\hat{\mathbf{R}}, \hat{\mathbf{b}}} \frac{\partial v_{AB}}{\partial(\hat{\mathbf{e}} \cdot \hat{\mathbf{a}}_\alpha)} \hat{\mathbf{e}} \times \hat{\mathbf{a}}_\alpha. \quad (24)$$

Choosing $\hat{\mathbf{n}}$ to be each of the coordinate directions in turn allows us to identify every component of the torque:

$$\mathbf{N}_{AB} = \sum_{\alpha} \sum_{\hat{\mathbf{e}}=\hat{\mathbf{R}}, \hat{\mathbf{b}}} \frac{\partial v_{AB}}{\partial(\hat{\mathbf{e}} \cdot \hat{\mathbf{a}}_\alpha)} \hat{\mathbf{e}} \times \hat{\mathbf{a}}_\alpha. \quad (25)$$

Writing this out explicitly for molecules A and B:

$$\mathbf{N}_{AB} = \sum_{\alpha} \frac{\partial v_{AB}}{\partial(\hat{\mathbf{a}}_{\alpha} \cdot \hat{\mathbf{R}})} \hat{\mathbf{R}} \times \hat{\mathbf{a}}_{\alpha} - \sum_{\alpha\beta} \frac{\partial v_{AB}}{\partial(\hat{\mathbf{a}}_{\alpha} \cdot \hat{\mathbf{b}}_{\beta})} \hat{\mathbf{a}}_{\alpha} \times \hat{\mathbf{b}}_{\beta} \quad (26a)$$

$$\mathbf{N}_{BA} = \sum_{\beta} \frac{\partial v_{AB}}{\partial(\hat{\mathbf{b}}_{\beta} \cdot \hat{\mathbf{R}})} \hat{\mathbf{R}} \times \hat{\mathbf{b}}_{\beta} + \sum_{\alpha\beta} \frac{\partial v_{AB}}{\partial(\hat{\mathbf{a}}_{\alpha} \cdot \hat{\mathbf{b}}_{\beta})} \hat{\mathbf{a}}_{\alpha} \times \hat{\mathbf{b}}_{\beta}. \quad (26b)$$

Note that eqns (22), (26a) and (26b) give $\mathbf{N}_{AB} + \mathbf{N}_{BA} + \mathbf{R} \times \mathbf{f}_{AB} = 0$ as before.

If the potential is not (easily) expressible in terms of scalar products, a similar derivation gives the expressions

$$\mathbf{f}_{AB} = -\frac{\partial v_{AB}}{\partial \mathbf{R}} = -\frac{\partial v_{AB}}{\partial R} \hat{\mathbf{R}} - \frac{\partial v_{AB}}{\partial \hat{\mathbf{R}}} \left[\frac{\mathbf{1} - \hat{\mathbf{R}}\hat{\mathbf{R}}}{R} \right] \quad (27a)$$

$$\mathbf{N}_{AB} = -\sum_{\alpha} \hat{\mathbf{a}}_{\alpha} \times \frac{\partial v_{AB}}{\partial \hat{\mathbf{a}}_{\alpha}} \quad (27b)$$

$$\mathbf{N}_{BA} = -\sum_{\beta} \hat{\mathbf{b}}_{\beta} \times \frac{\partial v_{AB}}{\partial \hat{\mathbf{b}}_{\beta}} \quad (27c)$$

which may be more convenient. Specific examples of the use of these equations in the context of liquid crystal simulations are given elsewhere⁸⁵.

Methods for integrating the rotational equations of motion, using a symplectic splitting method akin to that of section 4.1 have been described elsewhere. These tend to fall into two categories: those based on the rotation matrix^{86–88}, and those based on quaternion parameters⁸⁹. Here, we present briefly the former approach. Consider molecule *A*, and drop its identifying suffix. Assuming that the inertia tensor $\mathbf{I}$ is diagonal in the frame defined by the $\hat{\mathbf{a}}_{\alpha}$ vectors, the body-fixed angular momentum vector is defined by $\boldsymbol{\pi} = \mathbf{I} \cdot \boldsymbol{\omega}$, and the rotation matrix satisfies the equation

$$\frac{d\hat{\mathbf{a}}}{dt} = \hat{\mathbf{a}} \cdot \boldsymbol{\Omega} \quad \text{where} \quad \boldsymbol{\Omega} = \begin{pmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{pmatrix}$$

The angular momentum satisfies

$$\frac{d\boldsymbol{\pi}}{dt} = \boldsymbol{\pi} \times \mathbf{I}^{-1} \boldsymbol{\pi} + \mathbf{N}$$

where the torque $\mathbf{N}$ is expressed in body-fixed coordinates. Then, the essential idea is to split the rotational propagator

$$e^{i\mathbf{L}_{\pi}\delta t/2} e^{i\mathbf{L}_{\text{free}}\delta t} e^{i\mathbf{L}_{\pi}\delta t/2}$$

where the first and last components advance the angular momentum, using the computed torque, and the central part corresponds to free rotation governed by the kinetic energy part of the hamiltonian. In some special cases, the free rotation problem can be solved exactly, but in the general case it is split further into three separate rotations, each corresponding to a single element of the diagonal inertia tensor.

6 Molecular Dynamics in Different Ensembles

In this section we briefly discuss molecular dynamics methods in the constant- NVT ensemble; the reader should be aware that analogous approaches exist for other ensembles, particularly to simulate at constant pressure or stress.

There are three general approaches to conducting molecular dynamics at constant temperature rather than constant energy. One method, simple to implement and reliable, is to periodically reselect atomic velocities at random from the Maxwell-Boltzmann distribution⁹⁰. This is rather like an occasional random coupling with a thermal bath. The resampling may be done to individual atoms, or to the entire system; some guidance on the reselection frequency may be found in Ref. 90.

A second approach^{91,92}, is to introduce an extra ‘thermal reservoir’ variable into the dynamical equations:

$$\dot{\mathbf{r}}_i = \mathbf{p}_i/m \quad (28a)$$

$$\dot{\mathbf{p}}_i = \mathbf{f}_i - \zeta \mathbf{p}_i \quad (28b)$$

$$\dot{\zeta} = \frac{\sum_{i\alpha} p_{i\alpha}^2/m - gk_B T}{Q} \equiv \nu_T^2 \left[\frac{\sum_{i\alpha} p_{i\alpha}^2/m}{gk_B T} - 1 \right] = \nu_T^2 \left[\frac{\mathcal{T}}{T} - 1 \right]. \quad (28c)$$

Here ζ is a friction coefficient which is allowed to vary in time; Q is a thermal inertia parameter, which may be replaced by ν_T , a relaxation rate for thermal fluctuations; $g \approx 3N$ is the number of degrees of freedom. $\mathcal{T}$ stands for the instantaneous ‘mechanical’ temperature. It may be shown that the distribution function for the ensemble is proportional to $\exp\{-\beta\mathcal{W}\}$ where $\mathcal{W} = \mathcal{H} + \frac{1}{2}3Nk_B T \zeta^2/\nu_T^2$. These equations lead to the following time variation of the system energy $\mathcal{H} = \sum_{i\alpha} p_{i\alpha}^2/2m + \mathcal{U}$, and for the variable $\mathcal{W}$:

$$\begin{aligned} \dot{\mathcal{H}} &= \sum_{i\alpha} p_{i\alpha} \dot{p}_{i\alpha}/m - \sum_{i\alpha} f_{i\alpha} \dot{\mathbf{r}}_{i\alpha} = -\zeta \sum_{i\alpha} p_{i\alpha}^2/m \\ \dot{\mathcal{W}} &= -3Nk_B T \dot{\zeta}. \end{aligned}$$

If $\mathcal{T} > T$, i.e. the system is too hot, then the ‘friction coefficient’ ζ will tend to increase; when it is positive the system will begin to cool down. If the system is too cold, the reverse happens, and the friction coefficient may become negative, tending to heat the system up again. In some circumstances, this approach generates non-ergodic behaviour, but this may be ameliorated by the use of chains of thermostat variables⁹³. Ref. 94 gives an example of the use of this scheme in a biomolecular simulation.

It is also possible to extend the Liouville operator-splitting approach to generate algorithms for molecular dynamics in these ensembles⁶⁵. Some care needs to be taken, because eqns (28) are not hamiltonian, but it turns out to be possible to correct this using a suitable Poincaré transformation⁹⁵ and to implement the resulting symplectic method in an elegant fashion^{96,36}.

7 How Long? How Large?

Molecular dynamics evolves a finite-sized molecular configuration forward in time, in a step-by-step fashion. There are limits on the typical time scales and length scales that can

be investigated and the consequences must be considered in analyzing the results. Simulation runs are typically short: typically $t \sim 10^3\text{--}10^6$ MD steps, corresponding to perhaps a few nanoseconds of real time, and in special cases extending to the microsecond regime⁹⁷. This means that we need to test whether or not a simulation has reached equilibrium before we can trust the averages calculated in it. Moreover, there is a clear need to subject the simulation averages to a statistical analysis, to make a realistic estimate of the errors. How long should we run? This depends on the system and the physical properties of interest.

Suppose one is interested in a variable a , defined such that $\langle a \rangle = 0$. The time correlation function $\langle a(t_0)a(t_0 + t) \rangle$ relates values calculated at times t apart; assuming that the system is in equilibrium, this function is independent of the choice of time origin and may be written $\langle a(0)a(t) \rangle$. It will decay from an initial value $\langle a(0)a(0) \rangle \equiv \langle a^2 \rangle$ to a long-time limiting value

$$\lim_{t \rightarrow \infty} \langle a(0)a(t) \rangle = \lim_{t \rightarrow \infty} \langle a(0) \rangle \langle a(t) \rangle = 0$$

as the variables $a(0)$ and $a(t)$ become uncorrelated; this decay occurs over a characteristic time τ_a . Formally we may define a correlation time

$$\tau_a = \int_0^\infty dt \langle a(0)a(t) \rangle / \langle a^2 \rangle.$$

Alternatively, if time correlations decay exponentially at long time, τ_a may be identified approximately from the limiting form

$$\langle a(0)a(t) \rangle \propto \exp\{-t/\tau_a\}.$$

Similarly, define a spatial correlation function $\langle a(0)a(r) \rangle$ relating values computed at different points r apart. Spatial isotropy allows us to write this as a function of the distance between the points, r , rather than the vector $\mathbf{r}$: note that this symmetry is broken in a liquid crystal. Spatial homogeneity, which applies to simple liquids (but not to solids or smectic liquid crystals) allows us to omit any reference to an absolute origin of coordinates. This function decays from a short-range nonzero value to zero over a characteristic distance ξ_a , the correlation length.

It is almost essential for simulation box sizes L to be large compared with ξ_a , and for simulation run lengths τ to be large compared with τ_a , for all properties of interest a . Only then can we guarantee that reliably-sampled statistical properties are obtained. Roughly speaking, the statistical error in a property calculated as an average over a simulation run of length τ is proportional to $\sqrt{\tau_a/\tau}$: the time average is essentially a sum of $\sim \tau/\tau_a$ independent quantities, each an average over time τ_a . Within the time periods τ_a , values of a are highly correlated. A similar statement can be made about properties which are effectively spatial averages over the simulation box volume L^3 : root-mean-square variations of such averages are proportional to $\sqrt{(\xi_a/L)^3}$. This means that collective, system-wide properties deviate by only a relatively small amount from their thermodynamic, large-system, limiting values; the deviation becomes smaller as the averaging volume increases, and is also determined by the correlation length.

Near critical points, special care must be taken, in that these inequalities will almost certainly not be satisfied, and indeed one may see the onset of non-exponential decay of the correlation functions. In these circumstances a quantitative investigation of finite size effects and correlation times, with some consideration of the appropriate scaling laws, must

be undertaken. Phase diagrams of soft-matter systems often include continuous phase transitions, or weakly first-order transitions exhibiting significant pre-transitional fluctuations. One of the most encouraging developments of recent years has been the establishment of reliable and systematic methods of studying critical phenomena by simulation, although typically the Monte Carlo method is more useful for this type of study^{98–100, 34, 101, 51}.

8 Conclusions

In this introduction, I have tried to focus on points that seem to me both topical and essential to the mainstream of complex fluid dynamics. Others might have chosen a different perspective, or focused on different aspects. Exciting areas which I have had to omit include the use of molecular dynamics to study rare events, the development of mesoscale modelling techniques such as dissipative particle dynamics, the incorporation of electronic degrees of freedom through ab initio molecular dynamics, and the efficient implementation of simulation algorithms on parallel computers. One could argue that these are advanced topics, but they are progressively entering the mainstream, and ready-written packages are steadily removing the need to explain some of the lower-level technical issues which I have included here. Nonetheless, I am a firm believer in understanding what is happening “under the hood”, even if one does not intend to become a “mechanic”, and hopefully the foregoing material will help towards that end.

Acknowledgements

I have learnt much about the development of intermolecular potentials from Mark Wilson and members of his group. Guido Germano helped me understand several aspects of molecular motion and program design. Conversations with Sebastian Reich, Ben Leimkuhler, and others on the subject of algorithms have convinced me that this is still very much a live area. Finally, I wish to acknowledge the ongoing work of the CCP5 community in the UK.

References

1. G. C. Maitland, M. Rigby, E. B. Smith, and W. A. Wakeham. *Intermolecular forces: their origin and determination*. Clarendon Press, Oxford, 1981.
2. C. G. Gray and K. E. Gubbins. *Theory of molecular fluids. 1. Fundamentals*. Clarendon Press, Oxford, 1984.
3. M. Sprik. Effective pair potentials and beyond. In Michael P. Allen and Dominic J. Tildesley, editors, *Computer simulation in chemical physics*, volume 397 of *NATO ASI Series C*, pages 211–259, Dordrecht, 1993. Kluwer Academic Publishers. Proceedings of the NATO Advanced Study Institute on ‘New Perspectives on Computer Simulation in Chemical Physics’, Alghero, Sardinia, Italy, September 14–24, 1992.
4. A. J. Stone. *The Theory of Intermolecular Forces*. Clarendon Press, Oxford, 1996.
5. Michael P. Allen, Glenn T. Evans, Daan Frenkel, and Bela Mulder. Hard convex body fluids. *Adv. Chem. Phys.*, 86:1–166, 1993.

6. A. Rahman. Correlations in the motion of atoms in liquid argon. *Phys. Rev. A*, 136:405–411, 1964.
7. L. Verlet. Computer experiments on classical fluids. i. thermodynamical properties of Lennard-Jones molecules. *Phys. Rev.*, 159:98–103, 1967.
8. J. Weeks, D. Chandler, and H. C. Andersen. Role of repulsive forces in determining the equilibrium structure of simple liquids. *J. Chem. Phys.*, 54:5237–5247, 1971.
9. C. Holm. Dealing with long range interactions: Polyelectrolytes. In *Computational Soft Matter: From Synthetic Polymers to Proteins*, 2004. this volume.
10. S. L. Price. Toward more accurate model intermolecular potentials for organic molecules. *Rev. Comput. Chem.*, 14:225–289, 2000.
11. N. L. Allinger, Y. H. Yuh, and J.-H. Lii. Molecular mechanics - the MM3 force-field for hydrocarbons. 1. *J. Am. Chem. Soc.*, 111:8551–8566, 1989.
12. J.-H. Lii and N. L. Allinger. Molecular mechanics - the MM3 force-field for hydrocarbons. 2. Vibrational frequencies and thermodynamics. *J. Am. Chem. Soc.*, 111:8566–8575, 1989.
13. J.-H. Lii and N. L. Allinger. Molecular mechanics - the MM3 force-field for hydrocarbons. 3. the Van der Waals potentials and crystal data for aliphatic and aromatic hydrocarbons. *J. Am. Chem. Soc.*, 111:8576–8582, 1989.
14. N. L. Allinger, K. S. Chen, and J.-H. Lii. An improved force field (MM4) for saturated hydrocarbons. *J. Comput. Chem.*, 17:642–668, 1996.
15. N. Nevins, K. S. Chen, and N. L. Allinger. Molecular mechanics (MM4) calculations on alkenes. *J. Comput. Chem.*, 17:669–694, 1996.
16. N. Nevins, J.-H. Lii, and N. L. Allinger. Molecular mechanics (MM4) calculations on conjugated hydrocarbons. *J. Comput. Chem.*, 17:695–729, 1996.
17. S. J. Weiner, P. A. Kollman, D. A. Case, U. C. Singh, C. Ghio, G. Alagona, S. Profeta, and P. Weiner. A new force-field for molecular mechanical simulation of nucleic acids and proteins. *J. Am. Chem. Soc.*, 106:765–784, 1984.
18. W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, and P. A. Kollman. A 2nd generation force-field for the simulation of proteins, nucleic-acids, and organic molecules. *J. Am. Chem. Soc.*, 117:5179–5197, 1995.
19. B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J States, S. Swaminathan, and M. Karplus. CHARMM - A program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.*, 4:187–217, 1983.
20. W. L. Jorgensen, D. S. Maxwell, and J. TiradoRives. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J. Am. Chem. Soc.*, 118:11225–11236, 1996.
21. M. J. Field. *A Practical Introduction to the Simulation of Molecular Systems*. Cambridge University Press, Cambridge, 1999.
22. A. R. Leach. *Molecular Modelling: Principles and Applications*. Prentice Hall, 2nd edition, 2001.
23. H.-D. Holtje, G. Folkers, W. Sippl, and D. Rognan. *Molecular Modeling: Basic Principles and Applications*. Wiley-VCH, 2003.
24. K. Kremer and G. S. Grest. Dynamics of entangled linear polymer melts - a molecular dynamics simulation. *J. Chem. Phys.*, 92:5057–5086, 1990.
25. B. Chen, M. G. Martin, and J. I. Siepmann. Thermodynamic properties of the

- williams, opls-aa, and mmff94 all-atom force fields for normal alkanes. *J. Phys. Chem. B*, 102:2578–2586, 1998.
26. S. K. Nath, F. A. Escobedo, and J. J. de Pablo. On the simulation of vapor-liquid equilibria for alkanes. *J. Chem. Phys.*, 108:9905–9911, 1998.
 27. M. G. Martin and J. I. Siepmann. Transferable potentials for phase equilibria. 1. united-atom description of n-alkanes. *J. Phys. Chem. B*, 102:2569–2577, 1998.
 28. J. R. Errington and A. Z. Panagiotopoulos. A new intermolecular potential model for the n-alkane homologous series. *J. Phys. Chem. B*, 103:6314–6322, 1999.
 29. M. Tsige, J. G. Curro, G. S. Grest, and J. D. McCoy. Molecular dynamics simulations and integral equation theory of alkane chains: comparison of explicit and united atom models. *Macromolecules*, 36:2158–2164, 2003.
 30. E. Garcia, M. A. Glaser, N. A. Clark, and D. M. Walba. HFF: a force field for liquid crystal molecules. *J. Molec. Struct. - THEOCHEM*, 464:39–48, 1999.
 31. D. Reith, H. Meyer, and F. Müller-Plathe. CG-OPT: A software package for automatic force field design. *Comput. Phys. Commun.*, 148:299–313, 2002.
 32. D. Reith, M. Putz, and F. Müller-Plathe. Deriving effective mesoscale potentials from atomistic simulations. *J. Comput. Chem.*, 24:1624–1636, 2003.
 33. M. Pütz, K. Kremer, and G. S. Grest. What is the entanglement length in a polymer melt? *Europhys. Lett.*, 49:735–741, 2000.
 34. Michael P. Allen and Dominic J. Tildesley. *Computer simulation of liquids*. Clarendon Press, Oxford, hardback, 385pp edition, 1987.
 35. E. Hairer, C. Lubich, and Gerhard Wanner. Geometric numerical integration illustrated by the Störmer-Verlet method. *Acta Numerica.*, 12:399–450, 2003.
 36. C. J. Cotter and S. Reich. Time stepping algorithms for classical molecular dynamics. In M. Rieth and W. Schommers, editors, *Computational Nanotechnology*. American Scientific Publishers. to appear.
 37. B. Leimkuhler and S. Reich. *Geometric Numerical methods for Hamiltonian Mechanics*. Cambridge University Press, Cambridge, 2004. in press.
 38. C. W. Gear. The numerical integration of ordinary differential equations of various orders. ANL 7126, Argonne National Laboratory, 1966.
 39. C. W. Gear. *Numerical initial value problems in ordinary differential equations*. Prentice-Hall, Englewood Cliffs, NJ, 1971.
 40. J. M. Haile. *Molecular dynamics simulation: elementary methods*. Wiley, New York, 1992.
 41. D. C. Rapaport. *The art of molecular dynamics simulation*. Cambridge University Press, 1995.
 42. L. Verlet. Computer experiments on classical fluids. ii. equilibrium correlation functions. *Phys. Rev.*, 165:201–214, 1968.
 43. R. W. Hockney and J. W. Eastwood. *Computer simulations using particles*. Adam Hilger, Bristol, 1988.
 44. W. C. Swope, H. C. Andersen, P. H. Berens, and K. R. Wilson. A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules: application to small water clusters. *J. Chem. Phys.*, 76:637–649, 1982.
 45. H. Goldstein. *Classical Mechanics*. Addison Wesley, Reading, Massachusetts, second edition, 1980.

46. S. W. Deleeuw, J. W. Perram, and H. G. Petersen. Hamilton equations for constrained dynamic systems. *J. Stat. Phys.*, 61:1203–1222, 1990.
47. J-P. Ryckaert, G. Ciccotti, and H. J. C. Berendsen. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.*, 23:327–341, 1977.
48. G. Ciccotti, M. Ferrario, and J.-P. Ryckaert. Molecular dynamics of rigid systems in cartesian coordinates. a general formulation. *Molec. Phys.*, 47:1253–1264, 1982.
49. G. Ciccotti and J. P. Ryckaert. Molecular dynamics simulation of rigid molecules. *Comput. Phys. Rep.*, 4:345–392, 1986.
50. H. C. Andersen. Rattle: a “velocity” version of the shake algorithm for molecular dynamics calculations. *J. Comput. Phys.*, 52:24–34, 1983.
51. D. Frenkel and B. Smit. *Understanding molecular simulation : from algorithms to applications*. Academic Press, San Diego, 2nd edition, 2002.
52. W. F. van Gunsteren. Constrained dynamics of flexible molecules. *Molec. Phys.*, 40:1015–1019, 1980.
53. M. Fixman. Classical statistical mechanics of constraints: a theorem and application to polymers. *Proc. Nat. Acad. Sci.*, 71:3050–3053, 1974.
54. Michael P. Allen, Mark A. Warren, Mark. R. Wilson, Alain Sauron, and William Smith. Molecular dynamics calculation of elastic constants in Gay-Berne nematic liquid crystals. *J. Chem. Phys.*, 105:2850–2858, 1996.
55. D. Knuth. *The art of computer programming*. Addison-Wesley, Reading MA, 2nd edition, 1973.
56. D. J. Evans and G. P. Morriss. *Statistical Mechanics of Nonequilibrium Liquids*. Academic Press, London, 1990.
57. B. L. Holian. The character of the nonequilibrium steady state: beautiful formalism meets ugly reality. In K. Binder and G. Ciccotti, editors, *Monte Carlo and molecular dynamics of condensed matter systems*, volume 49, pages 791–822. Italian Physical Society, Bologna, 1996. Proceedings of the Euroconference on ‘Monte Carlo and molecular dynamics of condensed matter systems’, Como, Italy, July 3–28, 1995.
58. L. E. Reichl. *A modern course in Statistical Physics*. University of Texas Press, Austin, 1980.
59. H. L. Friedman. *A course in statistical mechanics*. Prentice-Hall, Englewood Cliffs, NJ, 1985.
60. B. L. Holian and D. J. Evans. Classical response theory in the Heisenberg picture. *J. Chem. Phys.*, 83:3560–3566, 1985.
61. M. Tuckerman, B. J. Berne, and G. J. Martyna. Reversible multiple time scale molecular dynamics. *J. Chem. Phys.*, 97:1990–2001, 1992.
62. H. F. Trotter. *Proc. Amer. Math. Soc.*, 10:545, 1959.
63. S. Toxvaerd. Hamiltonians for discrete dynamics. *Phys. Rev. E*, 50:2271–2274, 1994.
64. H. Grubmüller, H. Heller, A. Windemuth, and K. Schulten. Generalized Verlet algorithm for efficient molecular dynamics simulations with long-range interactions. *Molec. Simul.*, 6:121, 1991.
65. G. J. Martyna, M. E. Tuckerman, D. J. Tobias, and M. L. Klein. Explicit reversible integrators for extended systems dynamics. *Molec. Phys.*, 87:1117–1157, 1996.
66. P. Procacci, T. A. Darden, E. Paci, and M. Marchi. ORAC: A molecular dynamics program to simulate complex molecular systems with realistic electrostatic interac-

- tions. *J. Comput. Chem.*, 18:1848–1862, 1997.
67. P. Deuffhard, J. Hermans, B. Leimkuhler, A. E. Mark, S. Reich, and R. D. Skeel, editors. *Computational Molecular Dynamics: Challenges, Methods, Ideas*, volume 4 of *Lect. Notes Comp. Sci. Eng.* Springer, Berlin, 1998.
 68. T. Schlick, M. Mandziuk, R. D. Skeel, and K. Srinivas. Nonlinear resonance artifacts in molecular dynamics simulations. *J. Comput. Phys.*, 140:1–29, 1998.
 69. Q. Ma, J. A. Izaguirre, and R. D. Skeel. Verlet-I/r-RESPA/Impulse is limited by nonlinear instabilities. *S.I.A.M. J. Sci. Comput.*, 24:1951–1973, 2003.
 70. B. Garcia-Archilla, J. M. Sanz-Serna, and R. D. Skeel. Long-time-step methods for oscillatory differential equations. *S.I.A.M. J. Sci. Comput.*, 20:930–963, 1998.
 71. J. A. Izaguirre, S. Reich, and R. D. Skeel. Longer time steps for molecular dynamics. *J. Chem. Phys.*, 110:9853–9864, 1999.
 72. J. A. Izaguirre, D. P. Catarello, J. M. Wozniak, and R. D. Skeel. Langevin stabilization of molecular dynamics. *J. Chem. Phys.*, 114:2090–2098, 2001.
 73. T. Schlick, R. D. Skeel, A. T. Brunger, L. V. Kale, J. A. Board, J. Hermans, and K. Schulten. Algorithmic challenges in computational molecular biophysics. *J. Comput. Phys.*, 151:9–48, 1999.
 74. J. G. Gay and B. J. Berne. Modification of the overlap potential to mimic a linear site-site potential. *J. Chem. Phys.*, 74:3316–3319, 1981.
 75. E. de Miguel, L. F. Rull, M. K. Chalam, and K. E. Gubbins. Liquid crystal phase diagram of the Gay-Berne fluid. *Molec. Phys.*, 74:405–424, 1991.
 76. R. Berardi, A. P. J. Emerson, and C. Zannoni. Monte Carlo investigations of a Gay-Berne liquid crystal. *J. Chem. Soc. Faraday Trans.*, 89:4069–4078, 1993.
 77. M. A. Bates and G. R. Luckhurst. Computer simulation studies of anisotropic systems. 26. Monte Carlo investigations of a Gay-Berne discotic at constant pressure. *J. Chem. Phys.*, 104:6696–6709, 1996.
 78. M. R. Wilson. Molecular dynamics simulations of flexible liquid crystal molecules using a Gay-Berne / Lennard-Jones model. *J. Chem. Phys.*, 107:8654–8663, 1997.
 79. J. L. Billeter and R. A. Pelcovits. Simulations of liquid crystals. *Comput. Phys.*, 12:440–448, 1998.
 80. M. P. Allen. Liquid crystal systems. In *Computational Soft Matter: From Synthetic Polymers to Proteins*, 2004. this volume.
 81. R. Berardi, C. Fava, and C. Zannoni. A generalized Gay-Berne intermolecular potential for biaxial particles. *Chem. Phys. Lett.*, 236:462–468, 1995.
 82. D. J. Cleaver, C. M. Care, M. P. Allen, and M. P. Neal. Extension and generalization of the Gay-Berne potential. *Phys. Rev. E*, 54:559–567, 1996.
 83. R. Berardi, C. Fava, and C. Zannoni. A Gay-Berne potential for dissimilar biaxial particles. *Chem. Phys. Lett.*, 297:8–14, 1998.
 84. S. L. Price, A. J. Stone, and M. Alderton. Explicit formulae for the electrostatic energy, forces and torques between a pair of molecules of arbitrary symmetry. *Molec. Phys.*, 52:987–1001, 1984.
 85. M. P. Allen and G. Germano. Rigid body potentials, forces and torques. Technical report, University of Warwick, 2002.
<http://eprints.csc.warwick.ac.uk/archive/000000064/>.
 86. A. Dullweber, B. Leimkuhler, and R. McLachlan. Symplectic splitting methods for rigid body molecular dynamics. *J. Chem. Phys.*, 107:5840–5851, 1997.

87. A. Kol, B. B. Laird, and B. J. Leimkuhler. A symplectic method for rigid-body molecular simulation. *J. Chem. Phys.*, 107:2580–2588, 1997.
88. T. Holder, B. Leimkuhler, and S. Reich. Explicit variable step-size and time-reversible integration. *Appl. Numer. Math.*, 39:367–377, 2001.
89. T. F. Miller, M. Eleftheriou, P. Pattnaik, A. Ndirango, D. Newns, and G. J. Martyna. Symplectic quaternion scheme for biophysical molecular dynamics. *J. Chem. Phys.*, 116:8649–8659, 2002.
90. H. C. Andersen. Molecular dynamics simulations at constant pressure and/or temperature. *J. Chem. Phys.*, 72:2384–2393, 1980.
91. S. Nosé. A molecular dynamics method for simulations in the canonical ensemble. *Molec. Phys.*, 52:255–268, 1984.
92. W. G. Hoover. Canonical dynamics - equilibrium phase-space distributions. *Phys. Rev. A*, 31:1695–1697, 1985.
93. G. J. Martyna, M. L. Klein, and M. Tuckerman. Nose-Hoover chains: the canonical ensemble via continuous dynamics. *J. Chem. Phys.*, 97:2635–2643, 1992.
94. D. J. Tobias, G. J. Martyna, and M. L. Klein. Molecular dynamics simulations of a protein in the canonical ensemble. *J. Phys. Chem.*, 97:12959–12966, 1993.
95. S. D. Bond, B. J. Leimkuhler, and B. B. Laird. The nose-poincare method for constant temperature molecular dynamics. *J. Comput. Phys.*, 151:114–134, 1999.
96. S. Nose. An improved symplectic integrator for nose-poincare thermostat. *J. Phys. Soc. Japan*, 70:75–77, 2001.
97. Y. Duan and P. A. Kollman. Pathways to a protein folding intermediate observed in a 1-microsecond simulation in aqueous solution. *Science*, 282:740–744, 1998.
98. H. Müller-Krumbhaar and K. Binder. Dynamic properties of the Monte-Carlo method in statistical mechanics. *J. Stat. Phys.*, 8:1–24, 1973.
99. A. M. Ferrenberg, D. P. Landau, and K. Binder. Statistical and systematic errors in Monte-Carlo sampling. *J. Stat. Phys.*, 63:867–882, 1991.
100. K. Binder, editor. *The Monte Carlo Method in Condensed Matter Physics*, volume 71 of *Topics Appl. Phys.* Springer, Berlin, 2nd edition, 1995.
101. D. P. Landau and K. Binder. *A Guide to Monte Carlo Simulations in Statistical Physics*. Cambridge University Press, Cambridge, 2000.

Introduction to Monte Carlo Methods

Daan Frenkel

FOM Institute for Atomic and Molecular Physics,
Kruislaan 407, 1098 SJ Amsterdam, The Netherlands
E-mail: frenkel@amolf.nl

These give an introduction to Monte Carlo simulations. After a general introduction of the approach and practical implementation, special attention is paid to the use of biased sampling methods in the context of polymer simulations

1 Introduction

The Monte Carlo techniques that are described in this chapter can be used to compute the equilibrium properties of classical many-body systems. In this context, the word ‘classical’ means that the core motion of the constituent particles obeys the laws of classical mechanics. This is an excellent approximation for a wide range of materials. Only when we consider the translational or rotational motion of light atoms or molecules (He, H₂, D₂,) or vibrational motion with a frequency such that $h\nu > k_B T$, should we worry about quantum effects.

These lecture notes provide a somewhat selective introduction to the Monte Carlo (MC) method. The selection reflects my own personal bias. It is largely (but not completely) based on the more complete description given in ref.¹

Before embarking on a description of the MC method, I should first briefly explain the role of computer simulations in general. This topic is best discussed by considering the situation that prevailed before the advent of electronic computers. At that time, there was only one way to predict the outcome of an experiment, namely by making use of a theory that provided an approximate description of the system under consideration. The reason why an approximate theory was almost always used is that there are very few model systems for which the equilibrium properties can be computed exactly (examples are the ideal gas, the harmonic crystal and a number of lattice models, such as the two-dimensional Ising model for ferro-magnets), and even fewer model systems for which the transport properties can be computed exactly. As a result, most properties of real materials were predicted on the basis of approximate theories (examples are the van der Waals equation for dense gases, the Debye-Hückel theory for electrolytes, or the Boltzmann equation to describe the transport properties of dilute gases). Given sufficient information about the intermolecular interactions, these theories will provide us with an estimate of the properties of interest. Unfortunately, our knowledge of the intermolecular interactions of all but the simplest molecules is quite limited. This leads to a problem if we wish to test the validity of a particular theory by comparing directly to experiment. If we find that theory and experiment disagree, it may mean that our theory is wrong, or that we have an incorrect estimate of the intermolecular interactions, or both... Clearly, it would be very nice if we could obtain essentially exact results for a given model system, without having to rely on approximate theories. Computer simulations allow us to do precisely that. On the one hand, we can

now compare the calculated properties of a model system with those of an experimental system: if the two disagree, our model is inadequate, *i.e.* we have to improve on our estimate of the intermolecular interactions. On the other hand, we can compare the result of a simulation of a given model system with the predictions of an approximate analytical theory applied to the same model. If we now find that theory and simulation disagree, we know that the *theory* is flawed. So, in this case, the computer simulation plays the role of the ‘experiment’ that is designed to test the theory. This method to ‘screen’ theories before we apply them to the real world, is called a ‘computer experiment’. Computer experiments have become standard practice, to the extent that they now provide the first (and often the last) test of a new theoretical result.

In fact, the early history of computer simulation (see *e.g.* ref.²) illustrates this role of computer simulation. In some areas of physics there appeared to be little need for simulation because very good analytical theories were available (*e.g.* to predict the properties of dilute gases or of nearly harmonic crystalline solids). However, in other areas, few if any ‘exact’ theoretical results were known, and progress was much hindered by the fact that there were no unambiguous tests to assess the quality of approximate theories. A case in point was the theory of dense liquids. Before the advent of computer simulations, the only way to model liquids was by mechanical simulation^{3–5} of large assemblies of macroscopic spheres (*e.g.* ball bearings). But such mechanical models are rather crude, as they ignore the effect of thermal motion. Moreover, the analysis of the structures generated by mechanical simulation was very laborious and, in the end, had to be performed by computer anyway.

This may explain why, when in 1953 electronic computers were, for the first time, made available for unclassified research, numerical simulation of dense liquids was one of the first problems that was tackled. In fact, the first simulation of a ‘liquid’ was carried out by Metropolis et al.⁶ at Los Alamos, using (or, more properly, introducing) the Monte Carlo method. Almost simultaneously, Fermi, Pasta and Ulam⁷ performed a very famous numerical study of the dynamics of an anharmonic, one-dimensional crystal. However, the first proper Molecular Dynamics simulations were reported in 1956 by Alder and Wainwright⁸ at Livermore, who studied the dynamics of an assembly of hard spheres. The first MD simulation of a model for a ‘real’ material was reported in 1959 (and published in 1960) by the group of Vineyard at Brookhaven⁹, who simulated radiation damage in crystalline Cu (for a historical account, see¹⁰). The first MD simulation of a real liquid (argon) was reported in 1964 by Rahman at Argonne¹¹. After that, computers were increasingly becoming available to scientists outside the US government labs, and the practice of simulation started spreading to other continents^{12–14}. Much of the methodology of MD simulations has been developed since then, although it is fair to say that the basic algorithms for MC and MD have hardly changed since the fifties.

1.1 Statistical mechanics

In the present lecture, we describe the basic principles of the Monte Carlo method and molecular dynamics. In particular, we focus on simulations of systems of a fixed number of particles (N) in a given volume (V) at a temperature (T).

Our aim is to indicate where the Monte Carlo method comes in. We start from the

classical expression for the partition function Q :

$$Q = c \int d\mathbf{p}^N d\mathbf{r}^N \exp[-\mathcal{H}(\mathbf{r}^N, \mathbf{p}^N)/k_B T], \quad (1)$$

where $\mathbf{r}^N$ stands for the coordinates of all N particles, and $\mathbf{p}^N$ for the corresponding momenta. The function $\mathcal{H}(\mathbf{r}^N, \mathbf{p}^N)$ is the Hamiltonian of the system. It expresses the total energy of an isolated system as a function of the coordinates and momenta of the constituent particles: $\mathcal{H} = \mathcal{K} + \mathcal{U}$, where $\mathcal{K}$ is the kinetic energy of the system and $\mathcal{U}$ is the potential energy. Finally, c is a constant of proportionality, chosen such that the sum over quantum states approaches the classical partition function in the limit $\hbar \rightarrow 0$. For instance, for a system of N identical atoms, $c = 1/(h^{3N} N!)$. The classical equation is

$$\langle A \rangle = \frac{\int d\mathbf{p}^N d\mathbf{r}^N A(\mathbf{p}^N, \mathbf{r}^N) \exp[-\beta \mathcal{H}(\mathbf{p}^N, \mathbf{r}^N)]}{\int d\mathbf{p}^N d\mathbf{r}^N \exp[-\beta \mathcal{H}(\mathbf{p}^N, \mathbf{r}^N)]}, \quad (2)$$

where $\beta = 1/k_B T$. In this equation, the observable A has been expressed as a function of coordinates and momenta. As $\mathcal{K}$ is a quadratic function of the momenta the integration over momenta can be carried out analytically. Hence, averages of functions that depend on momenta only are usually easy to evaluate. The difficult problem is the computation of averages of functions $A(\mathbf{r}^N)$. Only in a few exceptional cases can the multidimensional integral over particle coordinates be computed analytically, in all other cases numerical techniques must be used.

Having thus defined the nature of the numerical problem that we must solve, let us next look at possible solutions. It might appear that the most straightforward approach would be to evaluate $\langle A \rangle$ in equation 2 by numerical quadrature, for instance using Simpson's rule. It is easy to see, however, that such a method is completely useless even if the number of independent coordinates DN (D is the dimensionality of the system) is still very small $\mathcal{O}(100)$. Suppose that we plan to carry out the quadrature by evaluating the integrand on a mesh of points in the DN -dimensional configuration space. Let us assume that we take m equidistant points along each coordinate axis. The total number of points at which the integrand must be evaluated is then equal to m^{DN} . For all but the smallest systems this number becomes astronomically large, even for small values of m . For instance, if we take 100 particles in three dimensions, and $m = 5$, then we would have to evaluate the integrand at 10^{210} points! Computations of such magnitude cannot be performed in the known universe. And this is fortunate, because the answer that would be obtained would have been subject to a large statistical error. After all, numerical quadratures work best on functions that are smooth over distances corresponding to the mesh size. But for most intermolecular potentials, the Boltzmann factor in equation 2 is a rapidly varying function of the particle coordinates. Hence an accurate quadrature requires a small mesh spacing (*i.e.* a large value of m). Moreover, when evaluating the integrand for a dense liquid (say), we would find that for the overwhelming majority of points this Boltzmann factor is vanishingly small. For instance, for a fluid of 100 hard spheres at the freezing point, the Boltzmann factor would be nonzero for 1 out of every 10^{260} configurations!

The closing lines of the previous section suggest that it is in general not possible to evaluate an integral, such as $\int d\mathbf{r}^N \exp[-\beta \mathcal{U}(\mathbf{r}^N)]$, by direct Monte Carlo sampling. However, in many cases, we are not interested in the configurational part of the partition func-

tion itself but in averages of the type

$$\langle A \rangle = \frac{\int d\mathbf{r}^N \exp[-\beta\mathcal{U}(\mathbf{r}^N)] A(\mathbf{r}^N)}{\int d\mathbf{r}^N \exp[-\beta\mathcal{U}(\mathbf{r}^N)]}. \quad (3)$$

Hence, we wish to know the *ratio* of two integrals. What Metropolis *et al.*⁶ showed is that it is possible to devise an efficient Monte Carlo scheme to sample such a ratio.^a To understand the Metropolis method, let us first look more closely at the structure of equation 3. In what follows we denote the configurational part of the partition function by Z :

$$Z \equiv \int d\mathbf{r}^N \exp[-\beta\mathcal{U}(\mathbf{r}^N)]. \quad (4)$$

Note that the ratio $\exp(-\beta\mathcal{U})/Z$ in equation 3 is the probability density to find the system in a configuration around $\mathbf{r}^N$. Let us denote this probability density by

$$\mathcal{N}(\mathbf{r}^N) \equiv \frac{\exp[-\beta\mathcal{U}(\mathbf{r}^N)]}{Z}.$$

Clearly, $\mathcal{N}(\mathbf{r}^N)$ is > 0 .

Suppose now that we are somehow able to randomly generate points in configuration space according to this probability distribution $\mathcal{N}(\mathbf{r}^N)$. This means that, on average, the number of points n_i generated per unit volume around a point $\mathbf{r}^N$ is equal to $L\mathcal{N}(\mathbf{r}^N)$, where L is the total number of points that we have generated. In other words;

$$\langle A \rangle \approx \frac{1}{L} \sum_{i=1}^L n_i A(\mathbf{r}_i^N). \quad (5)$$

By now the reader is almost certainly confused, so let us therefore try to clarify this method with the help of a simple example (see Figure 1). In this figure, we compare two ways to measure the depth of the river Nile, by conventional quadrature (left) and by Metropolis sampling; that is, the construction of an importance-weighted random walk (right). In the conventional quadrature scheme, the value of the integrand is measured at a predetermined set of points. As the choice of these points does not depend on the value of the integrand, many points may be located in regions where the integrand vanishes. In contrast, in the Metropolis scheme, a random walk is constructed through that region of space where the integrand is nonnegligible (*i.e.* through the Nile itself). In this random walk, a trial move is rejected if it takes you out of the water, and is accepted otherwise. After *every* trial move (accepted or not), the depth of the water is measured. The (unweighted) average of all these measurements yields an estimate of the average depth of the Nile. This, then, is the essence of the Metropolis method. In principle, the conventional quadrature scheme would also give results for the *total* area of the Nile. In the importance sampling scheme, however, information on the total area cannot be obtained directly, since this quantity is similar to Z .

Let us next consider how to generate points in configuration space with a relative probability proportional to the Boltzmann factor. The general approach is first to prepare the

^a An interesting account of the early history of the Metropolis method may be found in H.L. Anderson, *J. Stat. Phys.* 43:731, 1986; and Wood¹⁵ p. 3.

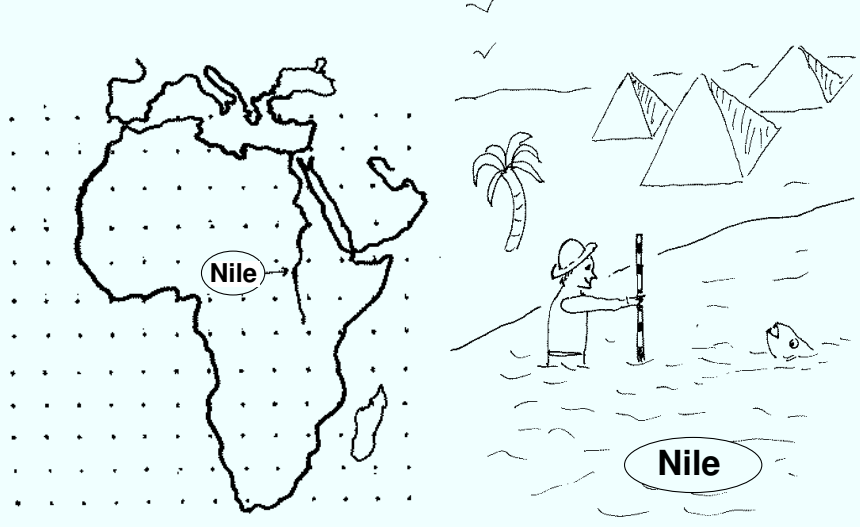


Figure 1. Measuring the depth of the Nile: a comparison of conventional quadrature (left), with the Metropolis scheme (right).

system in a configuration $\mathbf{r}^N$, which we denote by o (old), that has a nonvanishing Boltzmann factor $\exp[-\beta\mathcal{U}(o)]$. This configuration, for example, may correspond to a regular crystalline lattice with no hard-core overlaps. Next, we generate a new trial configuration $\mathbf{r}'^N$, which we denote by n (new), by adding a small random displacement Δ to o . The Boltzmann factor of this trial configuration is $\exp[-\beta\mathcal{U}(n)]$. We must now decide whether we will accept or reject the trial configuration. Many rules for making this decision satisfy the constraint that on average the probability of finding the system in a configuration n is proportional to $\mathcal{N}(n)$. Here we discuss only the Metropolis scheme, because it is simple and generally applicable.

Let us now “derive” the Metropolis scheme to determine the transition probability $\pi(o \rightarrow n)$ to go from configuration o to n . It is convenient to start with a thought experiment (actually a thought simulation). We carry out a very large number (say M) Monte Carlo simulations in parallel, where M is much larger than the total number of accessible configurations. We denote the number of points in any configuration o by $m(o)$. We wish that, on average, $m(o)$ is proportional to $\mathcal{N}(o)$. The matrix elements $\pi(o \rightarrow n)$ must satisfy one obvious condition: they do not destroy such an equilibrium distribution once it is reached. This means that, in equilibrium, the average number of accepted trial moves that result in the system leaving state o must be exactly equal to the number of accepted trial moves from all other states n to state o . It is convenient to impose a much stronger condition; namely, that in equilibrium the average number of accepted moves from o to any other state n is exactly canceled by the number of reverse moves. This detailed balance condition implies the following:

$$\mathcal{N}(o)\pi(o \rightarrow n) = \mathcal{N}(n)\pi(n \rightarrow o). \quad (6)$$

Many possible forms of the transition matrix $\pi(o \rightarrow n)$ satisfy equation (6). Let us look how $\pi(o \rightarrow n)$ is constructed in practice. We recall that a Monte Carlo move consists of

two stages. First, we perform a trial move from state o to state n . We denote the transition matrix that determines the probability to perform a trial move from o to n by $\alpha(o \rightarrow n)$; where α is usually referred to as the underlying matrix of Markov chain¹⁶. The next stage is the decision to either accept or reject this trial move. Let us denote the probability of accepting a trial move from o to n by $\text{acc}(o \rightarrow n)$. Clearly,

$$\pi(o \rightarrow n) = \alpha(o \rightarrow n) \times \text{acc}(o \rightarrow n). \quad (7)$$

In the original Metropolis scheme, α is chosen to be a symmetric matrix ($\text{acc}(o \rightarrow n) = \text{acc}(n \rightarrow o)$). However, in later sections we shall see several examples where α is *not* symmetric. If α is symmetric, we can rewrite equation 6 in terms of the $\text{acc}(o \rightarrow n)$:

$$\mathcal{N}(o) \times \text{acc}(o \rightarrow n) = \mathcal{N}(n) \times \text{acc}(n \rightarrow o). \quad (8)$$

From Eqn. 8 follows

$$\frac{\text{acc}(o \rightarrow n)}{\text{acc}(n \rightarrow o)} = \frac{\mathcal{N}(n)}{\mathcal{N}(o)} = \exp\{-\beta[\mathcal{U}(n) - \mathcal{U}(o)]\}. \quad (9)$$

Again, many choices for $\text{acc}(o \rightarrow n)$ satisfy this condition (and the obvious condition that the probability $\text{acc}(o \rightarrow n)$ cannot exceed 1). The choice of Metropolis *et al.* is

$$\begin{aligned} \text{acc}(o \rightarrow n) &= \mathcal{N}(n)/\mathcal{N}(o) \text{ if } \mathcal{N}(n) < \mathcal{N}(o) \\ &= 1 \text{ if } \mathcal{N}(n) \geq \mathcal{N}(o). \end{aligned} \quad (10)$$

Other choices for $\text{acc}(o \rightarrow n)$ are possible (for a discussion, see for instance¹⁷), but the original choice of Metropolis *et al.* appears to result in a more efficient sampling of configuration space than most other strategies that have been proposed.

1.2 A Basic Monte Carlo Algorithm

It is difficult to talk about Monte Carlo or Molecular Dynamics programs in abstract terms. The best way to explain how such programs work is to write them down. This will be done in the present section.

Most Monte Carlo or Molecular Dynamics programs are only a few hundred to several thousand lines long. This is very short compared to, for instance, a typical quantum-chemistry code. For this reason, it is not uncommon that a simulator will write many different programs that are tailor-made for specific applications. The result is that there is no such thing as a standard Monte Carlo or Molecular Dynamics program. However, the cores of most MD/MC programs are, if not identical, at least very similar. Next, we shall construct such a core. It will be very rudimentary, and efficiency has been traded for clarity. But it will serve to demonstrate how the Monte Carlo method work.

1.3 The Algorithm

The prime purpose of the kind of Monte Carlo or Molecular Dynamics program that we shall be discussing is to compute equilibrium properties of classical many-body systems. From now on, we shall refer to such programs simply as MC or MD programs, although it should be remembered that there exist many other applications of the Monte Carlo method

(and, to a lesser extent, of the Molecular Dynamics method). Let us now look at a simple Monte Carlo program.

In the previous section, the Metropolis method was introduced as a Markov process in which a random walk is constructed in such a way that the probability of visiting a particular point $\mathbf{r}^N$ is proportional to the Boltzmann factor $\exp[-\beta\mathcal{U}(\mathbf{r}^N)]$. There are many ways to construct such a random walk. In the approach introduced by Metropolis *et al.*⁶, the following scheme is proposed:

1. Select a particle at random, and calculate its energy $\mathcal{U}(\mathbf{r}^N)$.
2. Give the particle a random displacement; $\mathbf{r}' = \mathbf{r} + \Delta$, and calculate its new energy $\mathcal{U}(\mathbf{r}'^N)$.
3. Accept the move from $\mathbf{r}^N$ to $\mathbf{r}'^N$ with probability

$$\text{acc}(o \rightarrow n) = \min \left(1, \exp\{-\beta[\mathcal{U}(\mathbf{r}'^N) - \mathcal{U}(\mathbf{r}^N)]\} \right). \quad (11)$$

An implementation of this basic Metropolis scheme is shown in Algorithms 1 and 2.

Algorithm 1

[Basic Metropolis Algorithm]

PROGRAM mc	basic Metropolis algorithm
do icycl=1,ncycl	perform ncycl MC cycles
call mcmove	displace a particle
if (mod(icycl,nsamp).eq.0)	
+ call sample	sample averages
enddo	
end	

Comments:

1. Subroutine `mcmove` attempts to displace a randomly selected particle (see Algorithm 2).
2. Subroutine `sample` samples quantities every `nsamp` cycle.

1.4 Trial Moves

Now that we have specified the general structure of the Metropolis algorithm, we should consider its implementation. We shall not go into the problem of selecting intermolecular potentials for the model system under study. Rather, we shall simply assume that we have an atomic or molecular model system in a suitable starting configuration and that we have specified all intermolecular interactions. We must now set up the underlying Markov

Algorithm 2
[Attempt to Displace a Particle]

SUBROUTINE mcmove	attempts to displace a particle
o=int(ranf()*npart)+1	select a particle at random
call ener(x(o), eno)	energy old configuration
xn=x(o)+(ranf()-0.5)*delx	give particle random displacement
call ener(xn, enn)	energy new configuration
if (ranf().lt.exp(-beta	acceptance rule (11)
+ *(enn-eno)) x(o)=xn	accepted: replace x(o) by xn
return	
end	

Comments:

1. Subroutine `ener` calculates the energy of a particle at the given position.
2. Note that, if a configuration is rejected, the old configuration is retained.
3. The `ranf()` is a random number uniform in $[0, 1]$.

chain; that is, the matrix α . In more down-to-earth terms: we must decide how we are going to generate trial moves. We should distinguish between trial moves that involve only the molecular centers of mass, and those that change the orientation or possibly even the conformation of a molecule.

We start our discussion with trial moves of the molecular centers of mass. A perfectly acceptable method to create a trial displacement is to add random numbers between $-\Delta/2$ and $+\Delta/2$ to the x , y , and z coordinates of the molecular center of mass:

$$\begin{aligned}x'_i &\rightarrow x_i + \Delta (\text{Ranf} - 0.5) \\y'_i &\rightarrow y_i + \Delta (\text{Ranf} - 0.5) \\z'_i &\rightarrow z_i + \Delta (\text{Ranf} - 0.5),\end{aligned}\tag{12}$$

where `Ranf` are random numbers uniformly distributed between 0 and 1. Clearly, the reverse trial move is equally probable (hence, α is symmetric).

We are now faced with two questions: how large should we choose Δ ?, and should we attempt to move all particles simultaneously or one at a time? In the latter case we should pick the molecule that is to be moved at random to ensure that the underlying Markov chain remains symmetric. All other things being equal, we should choose the most efficient sampling procedure. But, to this end, we must first define what we mean by *efficient sampling*. Broadly speaking, sampling is efficient if it gives you good value for money. Good value in a simulation corresponds to high statistical accuracy, and “money” is simply *money*: the money that buys your computer time and even your own time. For the sake of the argument, we assume the average scientific programmer is poorly paid. In

that case we have to worry only about your computer budget.^b Then we could use the following definition of an optimal sampling scheme: a Monte Carlo sampling scheme can be considered optimal if it yields the lowest statistical error in the quantity to be computed for a given expenditure of computing budget. Usually, computing budget is equivalent to CPU time.

From this definition it is clear that, in principle, a sampling scheme may be optimal for one quantity but not for another. Actually, the preceding definition is all but useless in practice (as are most definitions). For instance, it is just not worth the effort to measure the error estimate in the pressure for a number of different Monte Carlo sampling schemes in a series of runs of fixed length. However, it is reasonable to assume that the mean-square error in the observables is inversely proportional to the number of uncorrelated configurations visited in a given amount of CPU time. And the number of independent configurations visited is a measure for the distance covered in configuration space. This suggests a more manageable, albeit rather ad hoc, criterion to estimate the efficiency of a Monte Carlo sampling scheme: the sum of the squares of all accepted trial displacements divided by computing time. This quantity should be distinguished from the mean-square displacement per unit of computing time, because the latter quantity goes to zero in the absence of diffusion (*e.g.* in a solid or a glass), whereas the former does not.

Next, consider the choice of the parameter Δ which determines the size of the trial move. How large should Δ be? If it is very large, it is likely that the resulting configuration will have a high energy and the trial move will probably be rejected. If it is very small, the change in potential energy is probably small and most moves will be accepted. In the literature, one often finds the mysterious statement that an acceptance of approximately 50% should be optimal. This statement is not necessarily true. The optimum acceptance ratio is the one that leads to the most efficient sampling of configuration space. If we express efficiency as mean-square displacement per CPU time, it is easy to see that different Monte Carlo codes will have different optimal acceptance ratios. The reason is that it makes a crucial difference if the amount of computing required to test whether a trial move is accepted depends on the magnitude of the move (see Figure 2).

In the conventional Metropolis scheme, all continuous interactions have to be computed before a move can be accepted or rejected. Hence, for continuous potentials, the amount of computation does not depend on the size of a trial move. In contrast, for simulations of molecules with hard repulsive cores, a move can be rejected as soon as overlap with any neighbor is detected. In that case, a rejected move is cheaper than an accepted one, and hence the average computing time per trial move goes down as the step size is increased. As a result, the optimal acceptance ratio for hard-core systems is appreciably lower than for systems with continuous interactions. Exactly how much depends on the nature of the program, in particular on whether it is a scalar or a vector code (in the latter case, hard-core systems are treated much like continuous systems), how the information about neighbor lists is stored, and even on the computational “cost” of random numbers and exponentiation. The consensus seems to be that for hard-core systems the optimum acceptance ratio is closer to 20 than to 50%, but this is just another rule of thumb that should be checked.

^bStill, we should stress that it is not worthwhile spending a lot of time developing a fancy computational scheme that will be only marginally better than existing, simpler schemes, unless your program will run very often and speed is crucial.

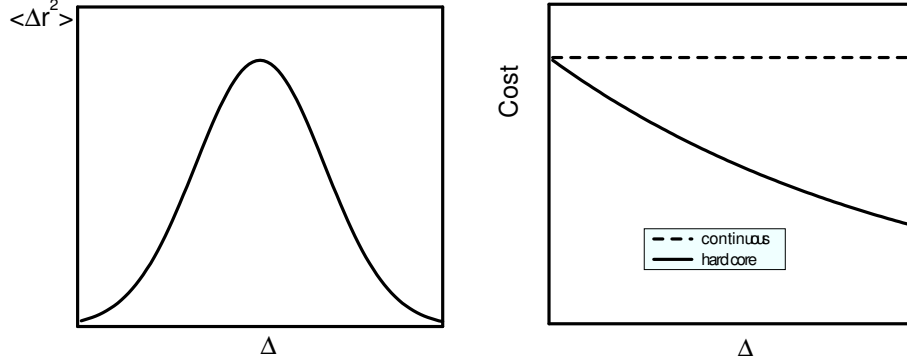


Figure 2. (left) Typical dependence of the mean-square displacement of a particle on the average size Δ of the trial move. (right) Typical dependence of the computational cost of a trial move on the step-size Δ . For continuous potentials, the cost is constant, while for hard-core potentials it decreases rapidly with the size of the trial move.

2 Canonical Ensemble

In a conventional Molecular Dynamics simulation, the total energy E and the total linear momentum P are constants of motion. Hence, Molecular Dynamics simulations measure (time) averages in an ensemble that is very similar to the microcanonical (see¹⁸); namely, the constant- NVE - P ensemble. In contrast, a conventional Monte Carlo simulation probes the canonical (*i.e.* constant- NVT) ensemble. The fact that these ensembles are different leads to observable differences in the statistical averages computed in Molecular Dynamics and Monte Carlo simulations. Most of these differences disappear in the thermodynamic limit and are already relatively small for systems of a few hundred particles. However, the choice of ensemble does make a difference when computing the mean-square value of fluctuations in thermodynamic quantities. Fortunately, techniques exist to relate fluctuations in different ensembles¹⁹. Moreover, nowadays it is common practice to carry out Molecular Dynamics simulations in ensembles other than the microcanonical. In particular, it is possible to do Molecular Dynamics at constant pressure, at constant stress, and at constant temperature. The choice of ensembles for Monte Carlo simulations is even wider: isobaric-isothermal, constant-stress-isothermal, grand-canonical (*i.e.* constant- μVT), and even microcanonical^{20–25}. A more recent addition to this list is a Monte Carlo method that employs the Gibbs-ensemble technique²⁶, which was developed to study phase coexistence in moderately dense (multi component) fluids.

2.1 General Approach

In the following sections, we will use the following procedure to demonstrate the validity of our Monte Carlo algorithms:

1. Decide which distribution we want to sample. This distribution, denoted $\mathcal{N}$, will depend on the details of the ensemble.

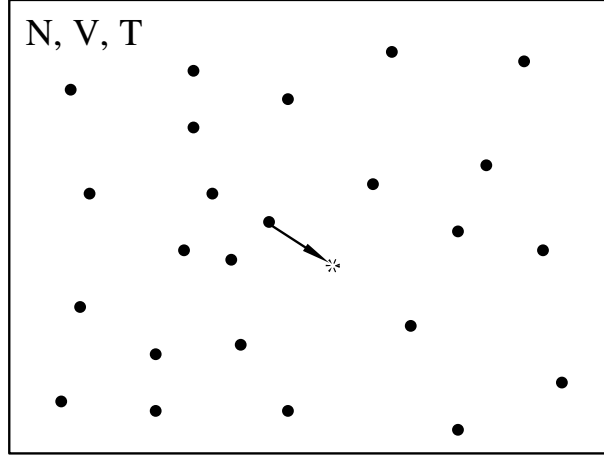


Figure 3. Canonical ensemble. The number of particles, volume, and temperature are constant. Shown is a Monte Carlo move in which a particle is displaced).

2. Impose the condition of detailed balance,

$$K(o \rightarrow n) = K(n \rightarrow o), \quad (13)$$

where $K(o \rightarrow n)$ is the flow of configuration o to n . This flow is given by the product of the probability of being in configuration o , the probability of generating configuration n , and the probability of accepting this move,

$$K(o \rightarrow n) = \mathcal{N}(o) \times \alpha(o \rightarrow n) \times \text{acc}(o \rightarrow n). \quad (14)$$

3. Determine the probabilities of generating a particular configuration.
4. Derive the condition which needs to be fulfilled by the acceptance rules.

It is instructive to apply the preceding recipe to the ordinary Metropolis scheme. In the canonical ensemble, the number of particles, temperature, and volume are constant (see Figure 3).

The partition function is

$$Q(N, V, T) \equiv \frac{1}{\Lambda^{3N} N!} \int d\mathbf{r}^N \exp[-\beta \mathcal{U}(\mathbf{r}^N)], \quad (15)$$

where $\Lambda = \sqrt{h^2/(2\pi m k_B T)}$ is the thermal de Broglie wavelength. From the partition function it follows that the probability of finding configuration $\mathbf{r}^N$ is given by distribution is

$$\mathcal{N}(\mathbf{r}^N) \propto \exp[-\beta \mathcal{U}(\mathbf{r}^N)]. \quad (16)$$

equations 15 and 16 are the basic equations for a simulation in the canonical ensemble.

2.2 Monte Carlo Simulations

In the canonical ensemble, we have to sample distribution (16). This can be done using the following scheme:

1. Select a particle at random and calculate the energy of this configuration $\mathcal{U}(o)$.
2. Give this particle a random displacement (see Figure 3),

$$\mathbf{r}(o) \rightarrow \mathbf{r}(o) + \Delta(\text{Ranf} - 0.5),$$

where $\Delta/2$ is the maximum displacement. The value of Δ should be chosen such that the sampling scheme is optimal (see section 1.4). The new configuration is denoted n and its energy $\mathcal{U}(n)$.

3. The move is accepted with a probability (see equation 10)

$$\text{acc}(o \rightarrow n) = \min(1, \exp\{-\beta[\mathcal{U}(n) - \mathcal{U}(o)]\}). \quad (17)$$

If rejected, the old configuration is kept.

An implementation of this basic Metropolis scheme is shown in Section 1.2 (Algorithms 1 and 2).

2.3 Justification of the Algorithm

The probability of generating a particular configuration is constant and independent of the conformation of the system

$$\alpha(o \rightarrow n) = \alpha(n \rightarrow o) = \alpha.$$

Substitution of this equation in the condition of detailed balance (13) and substitution of the desired distribution (16) gives as condition for the acceptance rules

$$\frac{\text{acc}(o \rightarrow n)}{\text{acc}(n \rightarrow o)} = \exp\{-\beta[\mathcal{U}(n) - \mathcal{U}(o)]\}. \quad (18)$$

It is straightforward to demonstrate that acceptance rule (17) obeys this condition.

3 Grand-Canonical Ensemble

The ensembles we have discussed so far have the total number of particles imposed. For some systems, however, one would like to obtain information on the average number of particles in a system as a function of the external conditions. For example, in adsorption studies one would like to know the amount of material adsorbed as a function of the pressure and temperature of the reservoir with which the material is in contact. A naive but theoretically valid approach would be to use the Molecular Dynamics technique (micro-canonical ensemble) and simulate the experimental situation; an adsorbent in contact with a gas (see Figure 4).

Such a simulation is possible only for very simple systems. In real experiments, equilibration may take minutes or even several hours depending on the type of gas molecules.

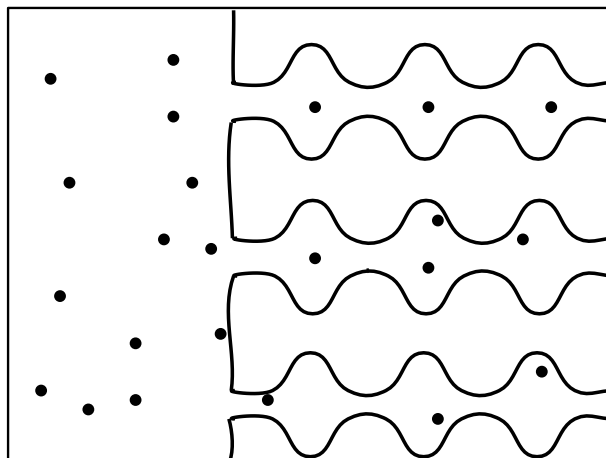


Figure 4. Adsorbent (for example a zeolite) in direct contact with a gas).

These equilibration times would be reflected in a Molecular Dynamics simulation, the difference being that a minute of experimental time takes of the order of 10^9 seconds on a computer. Furthermore, in most cases, we are not interested in the properties of the gas phase, yet a significant amount of CPU time will be spent on the simulation of this phase. Finally, in such a simulation, there is an interface between the gas phase and the adsorbent. In the interfacial region the properties of the system are different from the bulk properties in which we are interested. Since in a simulation the system is relatively small, we have to simulate a very large system to minimize the influence of this interfacial region.^c

Most of these problems can be solved by a careful choice of ensembles. For adsorption studies, a natural ensemble to use is the grand-canonical ensemble (or μ, V, T ensemble). In this ensemble, the temperature, volume, and chemical potential are fixed. In the experimental setup, the adsorbed gas is in equilibrium with the gas in the reservoir. The equilibrium conditions are that the temperature and chemical potential of the gas inside and outside the adsorbent must be equal.^d The gas that is in contact with the adsorbent can be considered as a reservoir that imposes a temperature and chemical potential on the adsorbed gas (see Figure 5).

We therefore have to know only the temperature and chemical potential of this reservoir to determine the equilibrium concentration inside the adsorbent. This is exactly what is mimicked in the grand-canonical ensemble: the temperature and chemical potential are imposed and the number of particles is allowed to fluctuate during the simulation. This makes these simulations different from the conventional ensembles, where the number of molecules is fixed.

^cSuch a simulation, of course, would be appropriate if the interest is in just this region.

^dNote that the pressure is *not* defined inside the zeolite; therefore, the pressure cannot be an equilibrium quantity. However, the pressure is related to the chemical potential via an equation of state, and it is always possible to calculate the pressure of the gas that corresponds to a given chemical potential and vice versa.

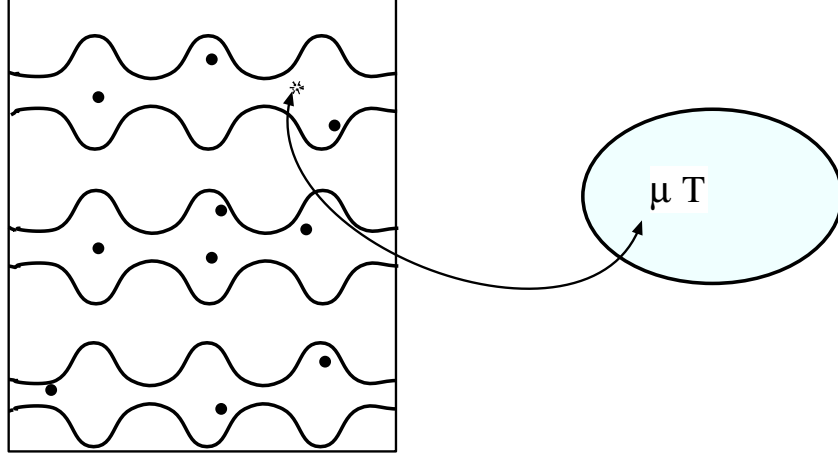


Figure 5. Adsorbent in contact with a reservoir that imposes constant chemical potential and temperature by exchanging particles and energy).

3.1 Statistical Mechanical Basis

The Metropolis sampling scheme as a method to compute thermal averages of functions $A(\mathbf{r}^N)$ that depend explicitly on the coordinates of the molecules in the N -body system under study. Examples of such mechanical properties are the potential energy or the virial contribution to the pressure. However, the Metropolis method could not be used to determine the integral $\int d\mathbf{r}^N \exp[-\beta\mathcal{U}(\mathbf{r}^N)]$ itself. The latter quantity measures the effective volume in configuration space that is accessible to the system. Hence, the original Metropolis scheme could not be used to determine those thermodynamic properties of a system that depend explicitly on the configurational integral. Examples of such thermal properties are the Helmholtz free energy F , the entropy S , and the Gibbs free energy G . However, although the Metropolis method cannot be used to measure, for instance, free energies directly, it can be used to measure the difference in free energy between two possible states of an N -body system. This fact is exploited in the grand-canonical Monte Carlo method first implemented for classical fluids by Norman and Filinov²³, and later extended and improved by a number of other groups^{24,27-34}. The basic idea of the grand-canonical Monte Carlo method is explained next.

To understand the statistical mechanical basis for the grand-canonical Monte Carlo technique, let us return to consider the partition function of a combined system of N interacting particles in volume V and $M - N$ ideal gas molecules in volume $V_0 - V$:

$$Q(N, M, V, V_0, T) = \frac{V^N (V_0 - V)^{M-N}}{\Lambda^{3M} N! (M - N)!} \int d\mathbf{s}^{M-N} \int d\mathbf{s}^N \exp[-\beta\mathcal{U}(\mathbf{s}^N)].$$

Now, instead of allowing the two systems to exchange volume, let us see what happens if the systems can also exchange particles (see Figure 6).

To be more precise, we assume that the molecules in the two subvolumes are actually identical particles. The only difference is that when they find themselves in volume V , they interact and, when they are in volume $V_0 - V$, they do not. If we transfer a molecule

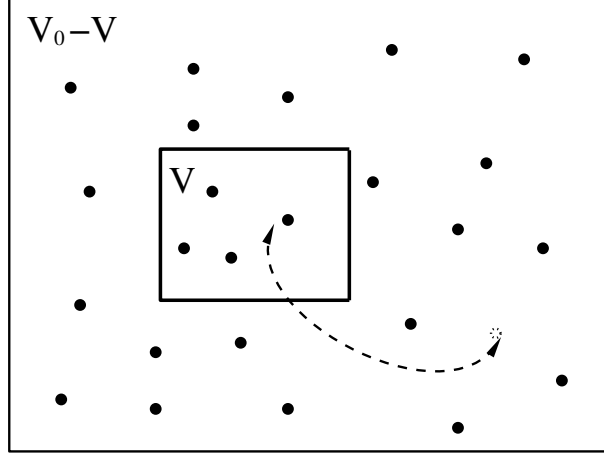


Figure 6. Ideal gas ($M - N$ particles, volume $V_0 - V$) can exchange particles with a N -particle system (volume V).

i from a reduced coordinate $\mathbf{s}_i$ in the volume $V_0 - V$ to the same reduced coordinate in volume V , then the potential energy function $\mathcal{U}$ changes from $\mathcal{U}(\mathbf{s}^N)$ to $\mathcal{U}(\mathbf{s}^{N+1})$. The expression for the total partition function of the system, including all possible distributions of the M particles over the two subvolumes is

$$Q(M, V, V_0, T) = \sum_{N=0}^M \frac{V^N (V_0 - V)^{M-N}}{\Lambda^{3M} N! (M - N)!} \int d\mathbf{s}^{M-N} \int d\mathbf{s}^N \exp[-\beta \mathcal{U}(\mathbf{s}^N)]. \quad (19)$$

We now write the probability density to find a system with $M - N$ particles at reduced coordinates $\mathbf{s}^{M-N}$ in volume $V' \equiv V_0 - V$ and N particles at reduced coordinates $\mathbf{s}^N$ in volume V :

$$\mathcal{N}(\mathbf{s}^M; N) = \frac{V^N V'^{M-N}}{Q(M, V, V', T) \Lambda^{3M} N! (M - N)!} \exp[-\beta \mathcal{U}(\mathbf{s}^N)]. \quad (20)$$

Let us now consider a trial move in which a particle is transferred from V' to the same scaled coordinate in V . First we should make sure that we construct an underlying Markov chain that is symmetric. Symmetry, in this case, implies that the a priori probability to move a particle from V' to V should be equal to the a priori probability of the reverse move. The probability of acceptance of a trial move in which we move a particle to or from volume V is determined by the ratio of the corresponding probability densities (20):

$$\alpha(N \rightarrow N + 1) = \frac{V(M - N)}{V'(N + 1)} \exp(-\beta[\mathcal{U}(\mathbf{s}^{N+1}) - \mathcal{U}(\mathbf{s}^N)]) \quad (21)$$

$$\alpha(N + 1 \rightarrow N) = \frac{V'(N + 1)}{V(M - N)} \exp(-\beta[\mathcal{U}(\mathbf{s}^N) - \mathcal{U}(\mathbf{s}^{N+1})]). \quad (22)$$

Now let us consider the limit that the ideal gas system is very much larger than the interacting system: $M \rightarrow \infty$, $V' \rightarrow \infty$, $(M/V') \rightarrow \rho$. Note that for an ideal gas the chemical

potential μ is related to the particle density ρ by

$$\mu = k_B T \ln \Lambda^3 \rho.$$

Therefore, in the limit $(M/N) \rightarrow \infty$, the partition function (19) becomes

$$Q(\mu, V, T) \equiv \sum_{N=0}^{\infty} \frac{\exp(\beta\mu N) V^N}{\Lambda^{3N} N!} \int d\mathbf{s}^N \exp[-\beta\mathcal{U}(\mathbf{s}^N)], \quad (23)$$

and the corresponding probability density

$$\mathcal{N}_{\mu VT}(\mathbf{s}^N; N) \propto \frac{\exp(\beta\mu N) V^N}{\Lambda^{3N} N!} \exp[-\beta\mathcal{U}(\mathbf{s}^N)]. \quad (24)$$

Equations 23 and 24 are the basic equations for Monte Carlo simulations in the grand-canonical ensemble. Note that, in these equations, all explicit reference to the ideal gas system has disappeared.

3.2 Monte Carlo Simulations

In a grand-canonical simulation, we have to sample the distribution (24). Acceptable trial moves are

1. *Displacement of particles.* A particle is selected at random and given a new conformation (for example in the case of atoms a random displacement). This move is accepted with a probability

$$\text{acc}(s \rightarrow s') = \min \left(1, \exp\{-\beta[\mathcal{U}(\mathbf{s}'^N) - \mathcal{U}(\mathbf{s}^N)]\} \right). \quad (25)$$

2. *Insertion and removal of particles.* A particle is inserted at a random position or a randomly selected particle is removed. The creation of a particle is accepted with a probability

$$\text{acc}(N \rightarrow N+1) = \min \left[1, \frac{V}{\Lambda^3(N+1)} \exp\{\beta[\mu - \mathcal{U}(N+1) + \mathcal{U}(N)]\} \right] \quad (26)$$

and the removal of a particle is accepted with a probability

$$\text{acc}(N \rightarrow N-1) = \min \left[1, \frac{\Lambda^3 N}{V} \exp\{-\beta[\mu + \mathcal{U}(N-1) - \mathcal{U}(N)]\} \right]. \quad (27)$$

The chemical potential of the reservoir can be related to the pressure of the reservoir. Algorithm 3 shows the basic structure of a simulation in the grand-canonical ensemble.

3.3 Justification of the Algorithm

It is instructive to demonstrate that the acceptance rules (25)–(27) indeed lead to a sampling of distribution (24). Consider a move in which we start with a configuration with N particles and move to a configuration with $N+1$ particles by inserting a particle in the system. Recall that we have to demonstrate that detailed balance is obeyed:

$$K(N \rightarrow N+1) = K(N+1 \rightarrow N),$$

Algorithm 3
[Basic Grand-Canonical Ensemble Simulation]

PROGRAM mc_gc	basic μ VT-ensemble simulation
do icycl=1,ncycl	perform ncycl MC cycles
ran=int(ranf()*(npart+nexc))+1	
if (ran.le.npart) then	
call mcmove	displace a particle
else	
call mcexc	exchange a particle with the reservoir
endif	
if (mod(icycl,nsamp).eq.0)	
+ call sample	sample averages
enddo	
end	

Comments:

1. This algorithm ensures that, after each MC step, detailed balance is obeyed. Per cycle we perform on average `npart` attempts to displace particles and `nexc` attempts to exchange particles with the reservoir.
2. Subroutine `mcmove` attempts to displace a particle (Algorithm 2 at page 36), subroutine `mcexc` attempts to exchange a particle with a reservoir (Algorithm 4), and subroutine `sample` samples quantities every `nsamp` cycle.

with

$$K(N \rightarrow N+1) = \mathcal{N}(N) \times \alpha(N \rightarrow N+1) \times \text{acc}(N \rightarrow N+1).$$

In Algorithm 3 at each Monte Carlo step the probability that an attempt is made to remove a particle is equal to the probability of attempting to add one:

$$\alpha_{\text{gen}}(N \rightarrow N+1) = \alpha_{\text{gen}}(N+1 \rightarrow N),$$

where the subscript “gen” refers to the fact that α measures the probability to generate this trial move. Substitution of this equation together with equation 24 into the condition of detailed balance gives:

$$\begin{aligned} \frac{\text{acc}(N \rightarrow N+1)}{\text{acc}(N+1 \rightarrow N)} &= \frac{\exp[\beta\mu(N+1)]V^{N+1} \exp[-\beta\mathcal{U}(\mathbf{s}^{N+1})]}{\Lambda^{3(N+1)}(N+1)!} \\ &\quad \times \frac{\Lambda^{3N}N! \exp[\beta\mathcal{U}(\mathbf{s}^N)]}{\exp(\beta\mu N)V^N} \\ &= \frac{\exp(\beta\mu)V}{\Lambda^3(N+1)} \exp\{-\beta[\mathcal{U}(\mathbf{s}^{N+1}) - \mathcal{U}(\mathbf{s}^N)]\}. \end{aligned}$$

It is straightforward to show that acceptance rules (26) and (27) obey this condition.

Algorithm 4

[Attempt to Exchange a Particle with a Reservoir]

SUBROUTINE mcexc	attempt to exchange a particle with a reservoir
if (ranf().lt.0.5) then	decide to remove or add a particle
if (npart.eq.0) return	test whether there is a particle
o=int(npart*ranf()+1)	select a particle to be removed
call ener(x(o),eno)	energy particle o
arg=npart*exp(beta*eno)	acceptance rule (27)
+ / (zz*vol)	
if (ranf().lt.arg) then	accepted: remove particle o
x(o)=x(npart)	
npart=npart-1	
endif	
else	
xn=ranf()*box	new particle at a random position
call ener(xn,enn)	energy new particle
arg=zz*vol*exp(-beta*enn)	acceptance rule (26)
+ / (npart+1)	
if (ranf().lt.arg) then	accepted: add new particle
x(npart+1)=xn	
npart=npart+1	
endif	
endif	
return	
end	

Comment:

1. We have defined: $zz = \exp(\beta\mu)/\Lambda^3$. The subroutine `ener` calculates the energy of a particle at a given position.

The most salient feature of the grand-canonical Monte Carlo technique is that in such simulations the chemical potential μ is imposed, while the number of particles N is a fluctuating quantity. During the simulation we may measure other thermodynamic quantities, such as the pressure P , the average density $\langle\rho\rangle$, or the internal energy $\langle\mathcal{U}\rangle$. As we know μ , we can derive all other thermal properties, such as the Helmholtz free energy or the entropy. This may seem surprising. After all, the Metropolis sampling cannot be used to sample absolute free energies and related quantities. Yet, with grand-canonical Monte Carlo we seem to be doing precisely that. The answer is that, in fact, we do not. What we measure is not an absolute but a relative free energy. In grand-canonical Monte Carlo, we are equating the chemical potential of a molecule in an ideal gas at density ρ (for the ideal gas case we know how to compute μ) and the chemical potential of the same species in an interacting system at density ρ' .

Grand-canonical Monte Carlo works best if the acceptance of trial moves by which

particles are added or removed is not too low. For atomic fluids, this condition effectively limits the maximum density at which the method can be used to about twice the critical density. Special tricks are needed to extend the grand-canonical Monte Carlo method to somewhat higher densities³². Grand-canonical Monte Carlo is easily implemented for mixtures and inhomogeneous systems, such as fluids near interfaces. In fact, some of the most useful applications of the grand-canonical Monte Carlo method are precisely in these areas of research. Although the grand canonical Monte Carlo technique can be applied to simple models of nonspherical molecules, special techniques are required since the method converges very poorly for all but the smallest polyatomic molecules.

4 Beyond Metropolis

4.1 Introduction

The original Metropolis Monte Carlo scheme⁶ was designed to perform single-particle trial moves. For most simulations, such moves are perfectly adequate. However, in some cases it is more efficient to perform moves in which the coordinates of many particles are changed. A case in point is the sampling of polymer conformations. The conventional Metropolis algorithm is ill-suited for polymer simulations because the natural dynamics of polymers is dominated by topological constraints (chains cannot cross). Hence, any algorithm that mimics the real motion of macromolecules will suffer from the same problem. For this reason, many algorithms have been proposed to speed up the Monte Carlo sampling of polymer conformations (see *e.g.* ref.³⁵). The Configurational-Bias Monte Carlo (CBMC) method is a dynamic MC scheme that makes it possible to achieve large conformational changes in a single trial move that affects a large number of monomeric units^{36–39}.

4.2 CBMC

The CBMC method is based on the Rosenbluth sampling scheme^{41,36,37} for lattice systems. In this scheme, the molecular conformation is built up step-by-step, in such a way that, at every stage, the next monomeric unit is preferentially added in a direction that has a large Boltzmann weight. This increases the probability of generating a trial conformation that has no hard-core overlaps. As explained below, the probability of acceptance of the trial conformation is given by the ratio of the ‘Rosenbluth weights’ of the new and the old conformations. Whereas the original Rosenbluth scheme was devised for polymers on a lattice, the CBMC scheme will also work for chain molecules in continuous space. The advantage of the CBMC algorithm over many of the other, popular algorithms is that it can be used in cases where particle-insertion and particle-removal trial moves are essential. This is the case in Grand-Canonical and Gibbs-Ensemble simulations. In addition, the CBMC can be used in the simulation of grafted chains and ring polymers.

4.2.1 Detailed balance

Before explaining the CBMC scheme, it is useful to recall the general recipe to construct a Monte Carlo algorithm. It is advisable (although not strictly obligatory⁴²) to start from the

condition of detailed balance:

$$P_o \times P_{gen}(o \rightarrow n) \times P_{acc}(o \rightarrow n) = P_n \times P_{gen}(n \rightarrow o) \times P_{acc}(n \rightarrow o) , \quad (28)$$

where P_o (P_n) is the Boltzmann weight of the old (new) conformation, P_{gen} denotes the *a priori* probability to generate the trial move from o to n , and P_{acc} is the probability that this trial move will be accepted. From Eqn. 28 it follows that

$$\frac{P_{acc}(o \rightarrow n)}{P_{acc}(n \rightarrow o)} = \exp(-\beta\Delta U) \frac{P_{gen}(n \rightarrow o)}{P_{gen}(o \rightarrow n)} , \quad (29)$$

where $\exp(-\beta\Delta U)$ is the ratio of the Boltzmann weights of the new and old conformations. If we use the Metropolis rule to decide on the acceptance of MC trial moves, then Eqn 29 implies

$$P_{acc}(o \rightarrow n) = \min \left(1, \exp(-\beta\Delta U) \frac{P_{gen}(n \rightarrow o)}{P_{gen}(o \rightarrow n)} \right) . \quad (30)$$

Ideally, by biasing the probability to generate a trial conformation in the right way, we could make the term on the right-hand side of Eqn. 30 always equal to unity. In that case, every trial move will be accepted. This ideal situation can be reached in rare cases⁴⁰ Configurational bias Monte Carlo does not achieve this ideal situation. However it does lead to enhanced acceptance probability of trial moves that involve large conformational changes.

In CBMC, chain configurations are generated by successive insertion of the bonded segments of the chain. When the positions of the segments are chosen at random, it is very likely, that one of the segments will overlap with another particle in the fluid, which results in rejection of the trial move. The Rosenbluth sampling scheme increases the insertion probability by looking one step ahead. On lattices, the availability (i.e. the Boltzmann factor) of all sites adjacent to the previous segment can be tested. In continuous space, there are in principle an infinite number of positions that should be tested (e.g. in the case of a chain molecule with rigid bonds, all points on the surface of a sphere with a radius equal to the bond length). Of course, it is not feasible to scan an infinite number of possibilities. Surprisingly, it turns out that it is possible to construct a correct Monte Carlo scheme for off-lattice models in which only a finite number of trial segments (k), is selected either at random or, more generally, drawn from the distribution of bond-lengths and bond-angles of the ‘ideal’ chain molecule.

During a CBMC trial move, a polymer conformation is generated segment-by-segment. At every step, k trial segments are generated (k is, in principle arbitrary and can be chosen to optimize computational efficiency). One of these segments, say i , is selected with a probability

$$P_i = \frac{\exp(-\beta u_i)}{\sum_{j=1}^k \exp(-\beta u_j)}$$

where u_j is the change in the potential energy of the system that would result of this particular trial segment was added to the polymer. The probability of generating a complete conformation Γ_{new} consisting of ℓ segments is then

$$P(\Gamma)_{new} = \prod_{n=1}^{\ell} \frac{\exp(-\beta u_i(n))}{\sum_{j=1}^k \exp(-\beta u_j(n))}$$

To keep the equations simple, we only consider the expression for one of the ℓ segments.

$$P_{gen}(\{j\}) = d\Gamma_j P_{id}(\Gamma_j) \left[\prod_{j' \neq j}^k d\Gamma_{j'} P_{id}(\Gamma_{j'}) \right] \frac{\exp(-\beta u_{ext}(j))}{\sum_{j'=1}^k \exp(-\beta u_{ext}(j'))} \quad (31)$$

In order to compute the acceptance probability of this move, we have to consider what happens in the reverse move. Then we start from conformation j and generate a set of k trial directions that includes i . When computing the acceptance probability of the forward move, we have to impose detailed balance. However, detailed balance in this case means not just that in equilibrium the number of moves from i to j is equal to the number of reverse move, but even that the rates are equal *for any given set of trial directions for the forward and reverse moves*. This condition we call ‘super-detailed balance’. Super-detailed balance implies that we can only decide on the acceptance of the forward move if we also generate a set of $k-1$ trial directions around the old conformation i . We denote the probability to generate this set of $k-1$ trial orientations by $P_{rest}(\{i\})$, where the sub-script ‘rest’ indicates that this is the set of orientations around, but excluding, i . This allows us to compute the ratio $w_j^{(t)}/w_i^{(o)}$ of the Rosenbluth weights for forward and reverse moves:

$$w_j^{(t)} = \frac{\exp(-\beta u_{ext}^{(t)}(j) + \sum_{j' \neq j}^{k-1} \exp(-\beta u_{ext}^{(t)}(j'))}{k}$$

and

$$w_i^{(o)} = \frac{\exp(-\beta u_{ext}^{(o)}(i) + \sum_{i' \neq i}^{k-1} \exp(-\beta u_{ext}^{(o)}(i'))}{k}$$

The superscript (t) and (o) distinguish the trial conformation from the old conformation. The acceptance probability is determined by the ratio $x \equiv w_j^{(t)}/w_i^{(o)}$ (actually, for a molecule of ℓ segments, we should compute a product of such factors). Let us assume that $w_j^{(t)} < w_i^{(o)}$. In that case, $P_{acc}(i \rightarrow j) = x$ while $P_{acc}(j \rightarrow i) = 1$. Next, let us check whether detailed balance is satisfied. To do so, we write down the explicit expressions for K_{ij} and K_{ji} .

$$K_{ij} = N_i P_{gen}(\{j\}) P_{rest}(\{i\}) w_j^{(t)} / w_i^{(o)}$$

and

$$K_{ji} = N_j P_{gen}(\{i\}) P_{rest}(\{j\}) 1$$

In addition, we use the fact that

$$P_{gen}(\{j\}) P_{rest}(\{i\}) w_j^{(t)} \sim N_j P_{rest}(\{j\}) P_{rest}(\{i\})$$

and

$$P_{gen}(\{i\}) P_{rest}(\{j\}) w_j^{(t)} \sim N_i P_{rest}(\{i\}) P_{rest}(\{j\})$$

It then follows immediately that

$$K_{ij}w_i^{(o)} = N_i P_{gen}(\{j\}) P_{rest}(\{i\}) w_j^{(t)} \quad (32)$$

$$\begin{aligned} &= \text{constant} \times N_i N_j P_{rest}(\{i\}) P_{rest}(\{j\}) \\ &= N_j P_{gen}(\{i\}) P_{rest}(\{j\}) w_i^{(o)} \\ &= K_{ji}w_i^{(o)} \end{aligned} \quad (33)$$

Hence, K_{ij} is indeed equal to K_{ji} . Note that, in this derivation, the number of trial directions, k , was arbitrary. The procedure sketched above is valid for a complete regrowth of the chain, but it is also possible to regrow only part of a chain, i.e. to cut a chain at a (randomly chosen) point and regrow the cut part of the chain either at the same site or at the other end of the molecule. Clearly, if only one segment is regrown and only one trial direction is used, CBMC reduces to the reptation algorithm (at least, for linear homopolymers).

We still have to consider the choice for the number of trial directions at the i -th regrowth step, k_i . Too many trial directions increase the cost of a simulation cycle, but too few trial directions lower the acceptance rate, and increase the simulation length. There exist simple guidelines that allow us to select k_i for every segment such that it optimizes the efficiency of the simulation⁴⁴.

4.3 Beyond CBMC

The CBMC method has several drawbacks. First of all, as the scheme looks ahead only one step at a time, it is likely to end up in “dead-alleys”. Secondly, for long chain molecules, the Rosenbluth weight of the trial conformation tends to become quite small. Hence, much of the computational effort may be wasted on “doomed” configurations.

4.4 DPERM

Grassberger and co-workers⁴⁵ have suggested adding two ingredients to the Rosenbluth scheme to improve its efficiency: “pruning” and “enrichment”. The basic rationale behind pruning is that it is not useful to spend much computer time on the generation of conformations that have a low Rosenbluth weight. Therefore, it is advantageous to discard (“prune”) such irrelevant conformations at an early stage. The idea behind enrichment is to make multiple copies of partially grown chains that have a large statistical weight^{45,46}, and to continue growing these potentially relevant, chains. The algorithm that combines these two features is called the Pruned-Enriched Rosenbluth Method (PERM). The examples presented by Grassberger and co-workers^{47,48} indicate that the PERM approach can be very useful to estimate the thermal equilibrium properties of long polymers. The main limitation of both the original Rosenbluth method and the PERM algorithm is that they are “static” Monte Carlo schemes. Such schemes can simulate single polymer chains very efficiently, but are less suited to simulate systems consisting of many polymer chains: at each step, one would have to simultaneously generate the conformations of all the chains in the system. On the other hand, in a dynamic scheme, one can conveniently choose a new point in phase space by only changing one chain at each step of the algorithm.

Just as CBMC is the “dynamic” version of Rosenbluth sampling. Similarly, one can construct a dynamic version of the PERM algorithm: DPERM⁴⁹.

The static PERM algorithm uses the Rosenbluth algorithm to generate the chains except that now pruning and enrichment are added. These ingredients are implemented as follows. At any step of the creation of a chain, if the partial Rosenbluth weight $W(j) = \prod_{i=1}^j w_i$ of a configuration is below a lower threshold $W^<(j)$, there is a probability of 50% to terminate the generation of this conformation. If the conformation survives this pruning step, its Rosenbluth weight is doubled $W^*(j) = 2 * W(j)$. Enrichment occurs when the partial Rosenbluth weight of a conformation $W(j) = \prod_{i=1}^j w_i$ exceeds an upper threshold $W^>(j)$. In that case, k copies of the partial chain are generated, each with a weight $W^*(j) = W(j)/k$. All these copies subsequently grow independently (subject to further pruning and enrichment).

The DPERM algorithm is the dynamic generalization of the PERM algorithm. As in the CBMC algorithm, we bias the acceptance of trials conformations to recover a correct Boltzmann sampling of chain conformations.

Thus, starting from an old configuration, we create a trial conformation and calculate the probability to generate it. Starting from the condition for detailed balance, we then derive the expression for the probability to accept or reject a new trial conformation. As we use the Rosenbluth method to generate chains, the probability to grow a particular conformation is :

$$P_{gen}(\text{chain}) = \prod_{i=1}^l \frac{e^{-\beta u^{(i)}(n)}}{w_i} \quad (34)$$

Whenever the re-weighted Rosenbluth partial weight $W^*(j)$ of the chain drops below the lower threshold $W^<(j)$, the chain has a 50% probability of being deleted. Let us assume that this happens m times. Then, the total probability to generate a *particular* conformation is :

$$P_{gen}(\text{chain}) = \frac{1}{2^m} \prod_{i=1}^l \frac{e^{-\beta u^{(i)}(n)}}{w_i} \quad (35)$$

and the re-weighted Rosenbluth weight of such a chain would be :

$$W^*(\text{chain\_new}) = 2^m * W(\text{chain\_new}) \quad (36)$$

$$\text{with } W(\text{chain\_new}) = \prod_{i=1}^l w_i \quad (37)$$

Whenever the Rosenbluth partial weight exceeds the upper threshold, k copies of the chain are created with the Rosenbluth weight $W^*(j) = W(j)/k$, which leads to the creation of a set of chains : this is a deterministic procedure. At every stage during the growth of the chain, others chains will branch off. The probability to grow the entire family of chains that is generated in one DPERM move can be written as :

$$P_{gen}(\text{chain\_new}) * P_{gen}(\text{rest\_new}) \quad (38)$$

Where $P_{gen}(\text{rest\_new})$ describes the product of the probabilities involved in generating all the other pieces of chains that branch off from the main chain. If we now call p the number of times the Rosenbluth weight exceeds the upper threshold during the generation of the *given* trial configuration, the probability to generate this *particular* chain is :

$$P_{gen}(\text{chain\_new}) = k^p \prod_{i=1}^l \frac{e^{-\beta u^{(i)}(n)}}{w_i} \quad (39)$$

and its re-weighted Rosenbluth weight is :

$$W^*(\text{chain\_new}) = \frac{1}{k^p} W(\text{chain\_new}) \quad (40)$$

Here k is the number of copies that are created each time the Rosenbluth weight exceeds the upper threshold. In Eqn. 39, the first term of the right hand side, describes the usual probability to generate a given chain following the Rosenbluth method. The factor k^p comes from the fact that the new chain could be any of the chains in the set so that the probability to generate a given chain is multiplied by this term. And we deduce Eqn. 40 from the fact that, each time we make some copies, the Rosenbluth weight is divided by k .

If we now also take into account the possibility that the chain can be pruned, then Eqn. 39 becomes :

$$\begin{aligned} P_{gen}(\text{chain\_new}) &= \frac{k^p}{2^m} \prod_{i=1}^l \frac{e^{-\beta u^{(i)}(n)}}{w_i} \\ &= \frac{k^p}{2^m} \frac{e^{-U(n)}}{W(\text{chain\_new})} \end{aligned} \quad (41)$$

And Eqn. 40 becomes :

$$W^*(\text{chain\_new}) = \frac{2^m}{k^p} W(\text{chain\_new}) \quad (42)$$

Note that Eqn. 41 and Eqn. 42 respectively reduce to Eqn. 35 and Eqn. 36 in the absence of enrichment ($p = 0$) and to Eqn. 39 and Eqn. 40 in the absence of pruning ($m = 0$).

We now choose to select the new trial chain from the set of chains created by the DPERM move with a probability given by :

$$P_{\text{choose\_new}} = W^*(\text{chain\_new}) / W_{total}(\text{new}) \quad (43)$$

where $W^*(\text{chain\_new})$ is the re-weighted Rosenbluth weight mentioned in Eqn. 42 and W_{total} is the sum of all such weights

$$W_{total}(\text{new}) = \sum_{\text{set}} W_{\text{chain}}^* \quad (44)$$

Eqn. 43 implies that we are most likely to choose the best chain (the one with the largest re-weighted Rosenbluth weight) of the set as the next Monte Carlo trial conformation.

Assuming that we start from an old configuration denoted by the subscript *old*, we generate a new configuration following the scheme described above and, we accept this move with the following acceptance rule :

$$acc(\text{old} \rightarrow \text{new}) = \min(1, \frac{W_{total}(\text{new})}{W_{total}(\text{old})}) \quad (45)$$

To calculate $W_{total}(\text{old})$, one has to “retrace” the old chain : the chain is first clear and reconstructed following the procedure described above to determine its weight. This is exactly analogous to what is done in the configurational-bias Monte Carlo scheme.

The demonstration that this scheme satisfies detailed balance is given in Ref.⁴⁹

This algorithm contains even more free parameters than CBMC. We choose them to optimize computational efficiency. In practice, this usually means that we wish to generate “enough” potentially successful trial chains, but not too many. Typically, the number of chains that survive at the end of a trial move should be $\mathcal{O}(1)$.

4.5 Recoil-growth

The recoil growth (RG) scheme is a dynamic Monte Carlo algorithm that was also developed with the dead-alley problem in mind^{50,51}. The algorithm is related to earlier static MC schemes due to Meirovitch⁵² and Alexandrowicz and Wilding⁵³. The basic strategy of the method is that it allows us to escape from a trap by “recoiling back” a few monomers and retrying the growth process using another trial orientation. In contrast, the CBMC scheme looks only one step ahead. Once a trial orientation has been selected, we cannot “deselect” it, even if it turns out to lead into a dead alley. The recoil growth scheme looks several monomers ahead to see whether traps are to be expected before a monomer is irrevocably added to the trial conformation. In this way we can alleviate (but not remove) the dead-alley problem. In principle, one could also do something similar with CBMC by adding a sequence of l monomers per step. However, as there are k possible directions for every monomer, this would involve computing k^l energies per group. Even though many of these trial monomers do not lead to acceptable conformations, we would still have to compute all interaction energies.

The RG algorithm is best explained by considering a totally impractical, but conceptually simple schemer that has the same effect. We place the first monomer at a random position. Next, we generate k trial positions for the second monomer. From each of these trial positions, we generate k trial positions for the third monomer. At this stage, we have generated k^2 “trimer” chains. We continue in the same manner until we have grown k^{l-1} chains of length l . Obviously, most of the conformations thus generated have a vanishing Boltzmann factor and are, therefore, irrelevant. However, some may have a reasonable Boltzmann weight and it is these conformations that we should like to find. To simplify this search, we introduce a concept that plays an important role in the RG algorithm: we shall distinguish between trial directions that are “open” and those that are “closed.” To decide whether a given trial direction, say b , for monomer j is open, we compute its energy $u_j(b)$. The probability that trial position b is open is given by

$$p_j^{\text{open}}(b) = \min(1, \exp[-\beta u_j(b)]), \quad (46)$$

For hard-core interactions, the decision whether a trial direction is open or closed is unambiguous, as $p_j^{\text{open}}(b)$ is either zero or one. For continuous interactions we compare $p_j^{\text{open}}(b)$ with a random number between 0 and 1. If the random number is less than $p_j^{\text{open}}(b)$, the direction is open; otherwise it is closed. We now have a tree with k^{l-1} branches but many of these branches are “dead,” in the sense that they emerge from a “closed” monomer. Clearly, there is little point in exploring the remainder of a branch if it does not correspond to an “open” direction. This is where the RG algorithm comes in. Rather than generating a host of useless conformations, it generates them “on the fly.” In addition, the algorithm uses a cheap test to check if a given branch will “die” within a specified number of steps (this number is denoted by l_{max}). The algorithm then randomly chooses among the available open branches. As we have only looked a distance l_{max} ahead, it may still happen that

we have picked a branch that is doomed. But the probability of ending up in such a dead alley is much lower than that in the CBMC scheme.

In practice, the recoil growth algorithm consists of two steps. The first step is to grow a new chain conformation using only “open” directions. The next step is to compute the weights of the new and the old conformations. The following steps are involved in the generation of a new conformation:

1. The first monomer of a chain is placed at a random position. The energy of this monomer is calculated (u_1). The probability that this position is “open” is given by Eqn. 46. If the position is closed we cannot continue growing the chain and we reject the trial conformation. If the first position is open, we continue with the next step.
2. A trial position b_{i+1} for monomer $i + 1$ is generated starting from monomer i . We compute the energy of this trial monomer $u_{i+1}(b)$ and, using Eqn. 46, we decide whether this position is open or closed. If this direction is closed, we try another trial position, up to a maximum of k trial orientations. As soon as we find an open position we continue with step 3.

If not a single open trial position is found, we make a recoil step. The chain retracts one step to monomer $i - 1$ (if this monomer exists), and the unused directions (if any) from step 2, for $i - 1$, are explored. If all directions at level $i - 1$ are exhausted, we attempt to recoil to $i - 2$. The chain is allowed to recoil a total of $l_{\max}$ steps, i.e., down to length $i - l_{\max} + 1$.

If, at the maximum recoil length, all trial directions are closed, the trial conformation is discarded.

3. We have now found an “open” trial position for monomer $i + 1$. At this point monomer $i - l_{\max}$ is permanently added in the new conformation; i.e., a recoil step will not reach this monomer anymore.
4. Steps 2 and 3 are repeated until the entire chain has been grown.

In the naive version of the algorithm sketched above, we can consider the above steps as a procedure for searching for an open branch on the existing tree. However, the RG procedure does this by generating the absolute minimum of trial directions compatible with the chosen recoil distance $l_{\max}$.

Once we have successfully generated a trial conformation, we have to decide on its acceptance. To this end, we have to compute the weights, $W(n)$ and $W(o)$, of the new and the old conformations, respectively. This part of the algorithm is more expensive. However, we only carry it out once we know for sure that we have successfully generated a trial conformation. In contrast, in CBMC it may happen that we spend much of our time computing the weight factor for a conformation that terminates in a dead alley.

In the RG scheme, the following algorithm is used to compute the weight of the new conformation:

1. Consider that we are at monomer position i (initially, of course, $i = 1$). In the previous stage of the algorithm, we have already found that at least one trial direction is available (namely, the one that is included in our new conformation). In addition, we may have found that a certain number of directions (say k_c) are closed—these are the

ones that we tried but that died within $l_{\max}$ steps. We still have to test the remaining $k_{\text{rest}} \equiv k - 1 - k_c$ directions. We randomly generate k_{rest} trial positions for monomer $i + 1$ and use the recoil growth algorithm to test whether at least one “feeler” of length $l_{\max}$ can be grown in this direction (unless $i + l_{\max} > l$; in that case we only continue until we have reached the end of the chain). Note that, again, we do *not* explore all possible branches. We only check if there is at least *one* open branch of length $l_{\max}$ in each of the k_{rest} directions. If this is the case, we call that direction “available”. We denote the total number of available directions (including the one that corresponds to the direction that we had found in the first stage of the algorithm) by m_i . In the next section we shall derive that monomer i contributes a factor $w_i(n)$ to the weight of the chain, where $w_i(n)$ is given by

$$w_i(n) = \frac{m_i(n)}{p_i^{\text{open}}(n)}$$

and $p_i^{\text{open}}(n)$ is given by Eqn. 46.

2. Repeat the previous step for all i from 1 to $l - 1$. The expression for the partial weight of the final monomer seems ambiguous, as $m_l(n)$ is not defined. An easy (and correct) solution is to choose $m_l(n) = 1$.
3. Next compute the weight for the entire chain:

$$W(n) = \prod_{i=1}^{\ell} w_i(n) = \prod_{i=1}^{\ell} \frac{m_i(n)}{p_i^{\text{open}}(n)}. \quad (47)$$

For the calculation of the weight of the old conformation, we use almost the same procedure. The difference is that, for the old conformation, we have to generate $k - 1$ additional directions for every monomer i . The weight is again related to the total number of directions that start from monomer i and that are “available”, i.e., that contain at least one open feeler of length $l_{\max}$:

$$W(o) = \prod_{i=1}^{\ell} w_i(o) = \prod_{i=1}^{\ell} \frac{m_i(o)}{p_i^{\text{open}}(o)}.$$

Finally, the new conformation is accepted with a probability:

$$\text{acc}(o \rightarrow n) = \min(1, \exp[-\beta U(n)]W(n) / \exp[-\beta U(o)]W(o)), \quad (48)$$

where $U(n)$ and $U(o)$ are the energies of the new and old conformations, respectively. It can easily be demonstrated (see ^{50,51}) that this scheme generates a Boltzmann distribution of conformations.

5 Relative Merits

A comparison between CBMC and the RG algorithm was made by Consta *et al.*⁵¹, who studied the behavior of Lennard-Jones chains in solution. The simulations showed that for relatively short chains ($\ell = 10$) at a density of $\rho = 0.2$, the recoil growth scheme was a factor of 1.5 faster than CBMC. For higher densities $\rho = 0.4$ and longer chains $N = 40$

the gain could be as large as a factor 25. This illustrates the fact that the recoil scheme is still efficient, under conditions where CBMC is likely to fail. For still higher densities or still longer chains, the relative advantage of RG would be even larger. However, the bad news is that, under those conditions, *both* schemes become very inefficient. A similar comparison has been made between CBMC and DPERM⁴⁹. For the cases studied, DPERM was no worse than CBMC, but also not much better. However, the fact that pruning can be performed on any chain property (i.e. not necessarily the Rosenbluth weight), may make the method attractive in special cases.

The basic idea behind both DPERM and RG is that these algorithms aim to avoid investing computational effort in the generation of trial moves that are, in the end, rejected. Houdayer⁵⁴ has proposed an algorithm that aims to achieve the same. In this algorithm, the trial move is a so-called “wormhole” move, where a polymer grows in one part of the simulation box while it shrinks at its original location. The growth-shrinkage process is carried out using a reptation-like algorithm. This algorithm has the advantage that, even if trial moves to the new state are rejected, the “old” state has also changed. This speeds up relaxation. In addition, the computing effort for the wormhole scheme appears to scale favorably with polymer size for long polymers (namely as N^n ⁵⁵, rather than as $\exp(cN)$). However, for short chains, the existing schemes are almost certainly more efficient (in ref.⁵⁴, the comparison with CBMC is made for a particularly inefficient CBMC-parameter choice). For longer chains, the wormhole scheme really should win. However, in that regime, all schemes are extremely costly. Nevertheless, as pointed out in ref.⁵⁴, a combination of the various algorithms would probably be more efficient than any one of them alone.

Acknowledgments

The work of the FOM Institute is part of the research program of the “Stichting voor Fundamenteel Onderzoek der Materie” (FOM), which is financially supported by the “Nederlandse Organisatie voor Wetenschappelijk Onderzoek” (NWO).

References

1. D. Frenkel and B. Smit *Understanding Molecular Simulations: from Algorithms to Applications*, Academic Press, San Diego, 2nd edition (2002).
2. G. Ciccotti, D. Frenkel, and I.R. McDonald. *Simulation of Liquids and Solids*. North-Holland, Amsterdam, 1987.
3. J.A. Prins, Onze voorstelling omtrent de bouw van de stof. Inaugural Address, 11-10-1928, Rijksuniversiteit Groningen *Physica*, 8:257, 1928.
4. O.K. Rice. *J. Chem. Phys.*, 12:1, 1944.
5. J.D. Bernal. Bakerian lecture. *Proc. Roy. Soc.*, 280:299, 1964.
6. N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.N. Teller, and E. Teller. Equation of state calculations by fast computing machines. *J. Chem. Phys.*, 21:1087–1092, 1953.
7. E. Fermi, J.G. Pasta, and S.M. Ulam. Studies of non-linear problems. *LASL Report*, LA-1940, 1955.

8. B.J. Alder and T.A. Wainwright. In I. Prigogine, editor, *Proceedings of the International Symposium on Statistical Mechanical Theory of Transport Processes (Brussels, 1956)*. Interscience, New York, 1958.
9. J.B. Gibson, A.N. Goland, M. Milgram, and G.H. Vineyard. Dynamics of radiation damage. *Phys. Rev.*, 120:1229–1253, 1960.
10. G.H. Vineyard. Autobiographical remarks of G.H. Vineyard. In P.C. Gehlen, J.R. Beeler, and R.I. Jaffe, editors, *Interatomic Potentials and Simulation of Lattice Defects*, pages xiii–xvi. Plenum, New York, 1972.
11. A. Rahman. Correlations in the motion of atoms in liquid argon. *Phys. Rev.*, 136:A405–A411, 1964.
12. L. Verlet. Computer ‘experiments’ on classical fluids. i. thermodynamical properties of Lennard-Jones molecules. *Phys. Rev.*, 159:98–103, 1967.
13. J.A. Barker and R.O. Watts. Structure of water: A Monte Carlo calculation. *Chem. Phys. Lett.*, 3:144–145, 1969.
14. I.R. McDonald and K. Singer. Calculation of the thermodynamic properties of liquid argon from Lennard-Jones parameters by a Monte Carlo method, *Discuss. Faraday Soc.*, 43:40, 1967.
15. W.W. Wood. Early history of computer simulation in statistical mechanics. In G. Ciccotti and W.G. Hoover, editors, *Molecular Dynamics Simulations of Statistical Mechanics Systems*, pages 2–14. Proceedings of the 97th Int. “Enrico Fermi” School of Physics, North Holland, Amsterdam, 1986.
16. N.G. van Kampen. *Stochastic Processes in Physics and Chemistry*. North-Holland, Amsterdam, 1981.
17. M.P. Allen and D.J. Tildesley. *Computer Simulation of Liquids*. Clarendon Press, Oxford, 1987.
18. J.J. Erpenbeck and W.W. Wood. Molecular dynamics techniques for hard-core systems. In B.J. Berne, editor, *Statistical Mechanics, Part B*, pages 1–40. Plenum, New York, 1976.
19. J.L. Lebowitz, J.K. Percus, and L. Verlet. Ensemble dependence of fluctuations with application to machine computations. *Phys. Rev.*, 153:250–254, 1967.
20. W.W. Wood. Monte Carlo calculations for hard disks in the isothermal-isobaric ensemble. *J. Chem. Phys.*, 48:415–434, 1968.
21. I.R. McDonald. NpT-ensemble Monte Carlo calculations for binary liquid mixtures. *Mol. Phys.*, 23:41–58, 1972.
22. R. Najafabadi and S. Yip. Observation of finite-temperature strain transformation (f.c.c. $\leftrightarrow$ b.c.c.) in Monte Carlo simulation of iron. *Scripta Metall.*, 17:1199–1204, 1983.
23. G.E. Norman and V.S. Filinov. Investigation of phase transitions by a Monte-Carlo method. *High Temp. (USSR)*, 7:216–222, 1969.
24. D.J. Adams. Chemical potential of hard-sphere fluids by Monte Carlo methods. *Mol. Phys.*, 28:1241–1252, 1974.
25. M. Creutz. Microcanonical Monte Carlo simulation. *Phys. Rev. Lett.*, 50:1411–1414, 1983.
26. A.Z. Panagiotopoulos. Direct determination of phase coexistence properties of fluids by Monte Carlo simulation in a new ensemble. *Mol. Phys.*, 61:813–826, 1987.
27. D.J. Adams. Grand canonical ensemble Monte Carlo for a Lennard-Jones fluid. *Mol.*

- Phys.*, 29:307–311, 1975.
28. D.J. Adams. Calculating the low temperature vapour line by Monte Carlo. *Mol. Phys.*, 32:647–657, 1976.
 29. D.J. Adams. Calculating the high-temperature vapour line by Monte Carlo. *Mol. Phys.*, 37:211–221, 1979.
 30. L.A. Rowley, D. Nicholson, and N.G. Parsonage. Monte Carlo grand canonical ensemble calculation in a gas-liquid transition region for 12-6 argon. *J. Comp. Phys.*, 17:401–414, 1975.
 31. J. Yao, R.A. Greenkorn, and K.C. Chao. Monte Carlo simulation of the grand canonical ensemble. *Mol. Phys.*, 46:587–594, 1982.
 32. M. Mezei. A cavity-biased (T,V, μ) Monte Carlo method for the computer simulation of fluids. *Mol. Phys.*, 40:901–906, 1980.
 33. J.P. Valleau and L.K. Cohen. Primitive model electrolytes. I. Grand canonical Monte Carlo computations. *J. Chem. Phys.*, 72:5935–5941, 1980.
 34. W. van Megen and I.K. Snook. The grand canonical ensemble Monte Carlo method applied to the electrical double layer. *J. Chem. Phys.*, 73:4656–4662, 1980.
 35. Monte Carlo and Molecular Dynamics Simulations in Polymer Science Edited by K. Binder, OUP, Oxford, 1995.
 36. J. Harris and S.A. Rice. A lattice model of a supported monolayer of amphiphilic molecules: Monte Carlo simulations *J. Chem. Phys.* **88**, 1298–1306 (1988).
 37. J.I. Siepmann and D. Frenkel, Configurational-bias Monte Carlo: a new sampling scheme for flexible chains *Mol.Phys.* **75**, 59–70 (1992).
 38. D. Frenkel, G.C.A.M. Mooij, and B. Smit, Novel scheme to study structural and thermal properties of continuously deformable molecules *J.Phys. Condensed Matter* **4**,3053–3076 (1992).
 39. J.J. de Pablo, M. Laso, and Suter U.W, Simulation of polyethylene above and below the melting point *J. Chem. Phys.* **96**, 2395–2403 (1992).
 40. R.H Swendsen and J.S. Wang, Nonuniversal critical dynamics in Monte Carlo simulations, *Phys. Rev. Lett.* **58**, 86–88 (1987).
 41. M.N. Rosenbluth and A.W. Rosenbluth, Monte Carlo simulations of the average extension of molecular chains *J.Chem. Phys.* **23**, 356–359 (1955).
 42. V.I. Manousiouthakis and M.W. Deem, Strict Detailed Balance is Unnecessary in Monte Carlo Simulation *J. Chem. Phys.*, **110**,2753–2758(1999).
 43. G.C.A.M. Mooij, D. Frenkel, and B. Smit, Direct simulation of phase equilibria of chain molecules *J. Phys. Condensed Matter* **4**, L255–L259 (1992).
 44. G.C.A.M. Mooij and D.Frenkel, A systematic optimization scheme for configurational bias Monte Carlo *Molecular Simulation* **17**,41–55(1996).
 45. P. Grassberger, Pruned-enriched Rosenbluth method: Simulations of theta polymers of chain length up to 1,000,000 *Phys. Rev. E* **56**, 3682–3693 (1997).
 46. U. Bastolla, H. Frauenkron, E. Gerstner, P. Grassberger, and W. Nadler, Testing a new Monte Carlo algorithm for protein folding *Proteins: Struc. Func. Gen.* **32**,52–66 (1998).
 47. H. Frauenkron, U. Bastolla, E. Gerstner, P. Grassberger, and W. Nadler, New Monte Carlo Algorithm for Protein Folding *Phys. Rev. Lett.* **80**,3149–3152 (1998).
 48. U. Bastolla and P. Grassberger, **xxx** *J. Stat. Phys.* **89**,1061 (1997).
 49. N. Combe, T.J.H.Vlugt. P.R. ten Wolde and D. Frenkel, Dynamic pruned-enriched

- Rosenbluth method *Mol. Phys.* **101**, 16751682 (2003).
50. S.Consta, T.J.H. Vlugt, J.Wichers Hoeth, B.Smit, and D.Frenkel, Recoil growth algorithm for chain molecules with continuous interactions *Mol. Phys.* **97**,1243–1254(1999).
 51. S. Consta, N.B. Wilding, D. Frenkel, and Z. Alexandrowicz, Recoil growth: An efficient simulation method for multi-polymer systems *J. Chem. Phys.* **110**,3220–3228(1999).
 52. H. Meirovitch, Statistical Properties Of The Scanning Simulation Method For Polymer-Chains *J. Chem. Phys.* **89**,2514–2522,(1988).
 53. Z. Alexandrowicz and N.B. Wilding, Simulation of polymers with rebound selection *J. Chem. Phys.* **109**,5622–5626(1998).
 54. J. Houdayer, The wormhole move: A new algorithm for polymer simulations *J. Chem. Phys.* **116**, 1783–1787(2002).
 55. Houdayer⁵⁴ suggests that the computing time per accepted move should scale as N^2 . However, as the problem is asymptotically that of a one-dimensional random walk with absorbing boundaries, it is more plausible that, for very long chains, the computing time per accepted move should scale as N^3 .

Advanced Simulations for Hydrodynamic Problems: Lattice Boltzmann and Dissipative Particle Dynamics

Burkhard Dünweg

Max-Planck-Institut für Polymerforschung,
55021 Mainz, Germany
E-mail: duenweg@mpip-mainz.mpg.de

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 lecture 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 lecture, we will deal with the dynamic behavior of polymers¹ and discuss the simplest systems only: We will focus on *linear*, *flexible*, and *uncharged* macromolecules, disregarding 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 sub-field 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 restrictive, 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. The generic phase diagram of such a solution has been presented in the contribution by J. Baschnagel; one sees that even the statics is quite non-trivial, giving rise to a host of scaling laws and various crossovers. Fig. 1 shows this diagram again, augmented by information on the dynamics. One sees that the understanding the dynamics is even more demanding than the statics, and that our present state of understanding is based on three fundamental models: The Rouse model², the Zimm model³, and the reptation model¹. In the present lecture we will discuss the Rouse and the Zimm model, while

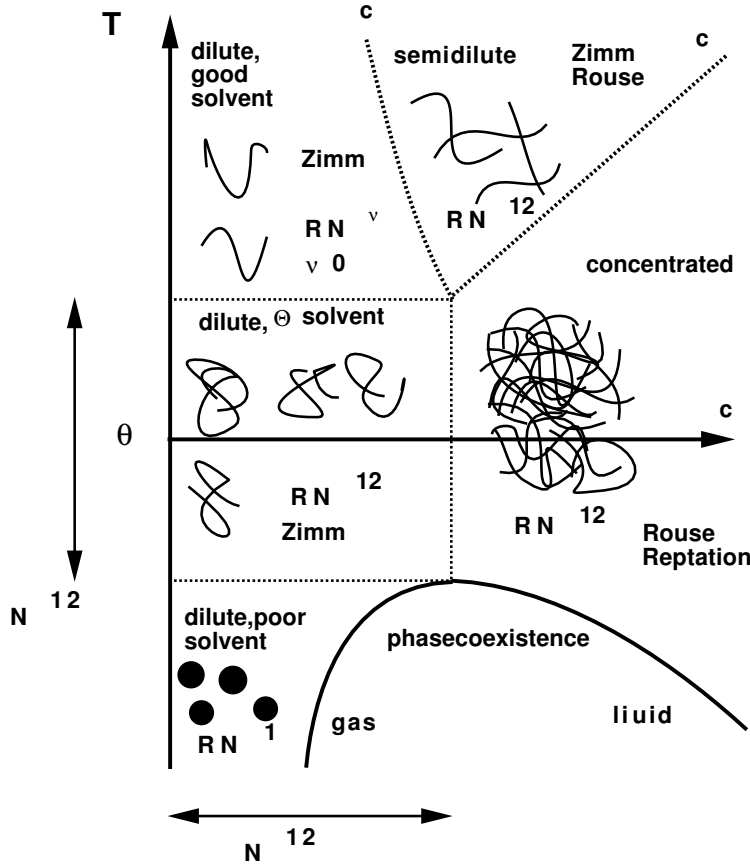


Figure 1. Phase diagram of a polymer solution, in the plane concentration vs. solvent quality. The various dynamic regimes and crossovers are indicated.

we refer the reader to the contributions by J. Baschnagel and K. Kremer for the reptation model.

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 increased¹⁵.

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 lecture will also disregard solvent quality effects and focus on good solvents only.

The methods to be discussed in this lecture 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 coil¹⁹. We model the chain as a sequence of N “monomers” with positions $\mathbf{r}_i$, $i = 1, \dots, N$. 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 $\mathbf{r}_i$ 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 ,

$$\langle b^2 \rangle = \langle (\mathbf{r}_{i+1} - \mathbf{r}_i)^2 \rangle. \quad (1)$$

It is however also possible to combine several repeat units into a “super-monomer”. 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

$$R \sim bN^\nu. \quad (2)$$

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,

$$\langle R_E^2 \rangle = \langle (\mathbf{r}_N - \mathbf{r}_1)^2 \rangle, \quad (3)$$

the gyration radius

$$\langle R_G^2 \rangle = \frac{1}{N} \sum_i \langle (\mathbf{r}_i - \mathbf{R}_{CM})^2 \rangle \quad (4)$$

($\mathbf{R}_{CM}$ denoting the chain's center of mass, $\mathbf{R}_{CM} = N^{-1} \sum_i \mathbf{r}_i$), and the hydrodynamic radius

$$\left\langle \frac{1}{R_H} \right\rangle = \frac{1}{N^2} \sum_{i \neq j} \left\langle \frac{1}{r_{ij}} \right\rangle, \quad (5)$$

where $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$.

Another important way to characterize the conformations is via the single-chain static structure factor $S(k)$, defined as

$$S(k) = \frac{1}{N} \left\langle \left| \sum_i \exp(i\mathbf{k} \cdot \mathbf{r}_i) \right|^2 \right\rangle = \frac{1}{N} \left\langle \sum_{ij} \exp(i\mathbf{k} \cdot (\mathbf{r}_i - \mathbf{r}_j)) \right\rangle. \quad (6)$$

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 $S(k) = N f(kR)$. 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

$$P(\{\mathbf{r}_i\}) \propto \exp\left(-\frac{V(\{\mathbf{r}_i\})}{k_B T}\right), \quad (7)$$

k_B and T denoting the Boltzmann constant and the absolute temperature, respectively.

For a Gaussian chain, V is just a harmonic potential:

$$V = \frac{3k_B T}{2b^2} \sum_{i=1}^{N-1} (\mathbf{r}_{i+1} - \mathbf{r}_i)^2. \quad (8)$$

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^* where overlap starts to happen¹⁹ is estimated via the requirement that an arrangement of unperturbed SAWs is just space-filling:

$$c^* \sim \frac{N}{R^3} \sim \frac{N}{b^3 N^{3\nu}} \sim b^{-3} N^{-(3\nu-1)}. \quad (9)$$

Solutions with concentration $c \gg c^*$ (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

$$\xi \sim b (cb^3)^{-\frac{\nu}{3\nu-1}} \propto c^{-0.77}. \quad (10)$$

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 essentially modifications. The Rouse model is based on quite a number of assumptions, some of which are easy to justify, while some are far from obvious.

Firstly, there is a colossal reduction in the number of degrees of freedom. While a true Newtonian description requires the positions and momenta of all particles, the Rouse model only takes into account the positions of the monomers of a single chain, $\mathbf{r}_i$, $i = 1, \dots, N$, and assumes the following overdamped Langevin equation of motion:

$$\frac{d}{dt}\mathbf{r}_i = \frac{1}{\zeta}\mathbf{F}_i + \boldsymbol{\rho}_i, \quad (11)$$

where ζ is the monomer friction coefficient, $\boldsymbol{\rho}_i$ the random displacement (per unit time) acting on monomer i , and $\mathbf{F}_i$ the force acting on monomer i , derived from the effective potential V (see previous subsection):

$$\mathbf{F}_i = -\frac{\partial V}{\partial \mathbf{r}_i}. \quad (12)$$

The random displacements are Gaussian white noise satisfying the standard fluctuation-dissipation theorem,

$$\langle \rho_i^\alpha \rangle = 0 \quad \langle \rho_i^\alpha(t) \rho_j^\beta(t') \rangle = 2 \frac{k_B T}{\zeta} \delta_{ij} \delta_{\alpha\beta} \delta(t - t'), \quad (13)$$

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 is safe only at first, but not at second glance. The motivation 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 tails²¹ we know that this is not quite true: Since momentum is a conserved quantity, it can only be transported away, but not simply destroyed. On long time and length scales, this transport of momentum can be described by the Stokes equation (hydrodynamic equation of motion for an incompressible fluid, where the nonlinear term is neglected):

$$\rho \frac{\partial}{\partial t} \mathbf{u} = \eta \nabla^2 \mathbf{u}, \quad \nabla \cdot \mathbf{u} = 0, \quad (14)$$

where ρ is the fluid density, η its viscosity, and $\mathbf{u}$ 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.

Looking again at the Rouse equation of motion, Eq. 11, this would mean that a description in terms of a simple friction coefficient is not strictly correct. Instead, one would have to use an appropriate memory function. However, the Rouse model is known to work (approximately) for dense polymer systems (short-chain melts). In this case, 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, 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. 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.

Conversely, 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, which is actually much more important than the self-correlation, is called “hydrodynamic interaction”. Taking these correlations into account, one obtains the Zimm model, which is discussed below.

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 to this fact. The reptation model, where the entanglements with the other chains are replaced by a tube of a certain diameter¹, 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^{22,23}, 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^{24,25}, 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

$$\mathbf{X}_p = \sqrt{2} N^{-1/2} \sum_{i=1}^N \mathbf{r}_i \cos \left[\frac{p\pi}{N} (i - 1/2) \right], \quad p = 1, \dots, N - 1, \quad (15)$$

whose equations of motion decouple. Each mode is characterized by a mean square am-

plitude

$$\langle \mathbf{X}_p^2 \rangle = \frac{b^2}{4 \sin^2 \left(\frac{p\pi}{2N} \right)}, \quad (16)$$

and a relaxation function

$$\langle \mathbf{X}_p(t) \cdot \mathbf{X}_p(0) \rangle = \langle \mathbf{X}_p^2 \rangle \exp \left(-\frac{t}{\tau_p} \right), \quad (17)$$

where the mode relaxation time τ_p is given by

$$\tau_p^{-1} = \frac{12k_B T}{\zeta b^2} \sin^2 \left(\frac{p\pi}{2N} \right). \quad (18)$$

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

$$\tau_R \propto R^z. \quad (19)$$

We therefore see $z = 4$ for the RW Rouse model.

The mean square displacement of a single monomer, $\langle \Delta \mathbf{r}^2 \rangle$, can be written exactly as a complicated expression, which however behaves asymptotically as $\langle \Delta \mathbf{r}^2 \rangle \propto t^{1/2}$ 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

$$D = \frac{k_B T}{N\zeta}. \quad (20)$$

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 Eq. 20 is derived immediately. The scaling law is

$$D \propto N^{-1} \propto R^{-1/\nu}. \quad (21)$$

Furthermore, one estimates the longest relaxation time τ_R via the consideration that the object will just move its own size within τ_R :

$$D\tau_R \propto R^2 \quad \tau_R \propto R^{2+1/\nu}, \quad (22)$$

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 $\langle \Delta \mathbf{r}^2 \rangle \sim R^2$ for $t \sim \tau_R$ fixes the exponent as

$$\langle \Delta \mathbf{r}^2 \rangle \propto t^{2/z}. \quad (23)$$

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

$$S(k, t) = \frac{1}{N} \left\langle \sum_{ij} \exp(i\mathbf{k} \cdot (\mathbf{r}_i(t) - \mathbf{r}_j(0))) \right\rangle, \quad (24)$$

satisfies the scaling relation

$$S(k, t) = k^{-1/\nu} f(k^2 t^{2/z}) \quad (25)$$

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 $\mathbf{r}_i$ is *highly correlated* with that at position $\mathbf{r}_j$, 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 $\mathbf{r}_i$ to $\mathbf{r}_j$), 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 $\mathbf{F}$ acting on a monomer at $\mathbf{r}_i$ 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 is¹ the so-called Oseen tensor

$$\mathbf{u}(\mathbf{r}) = \overleftrightarrow{T}(\mathbf{r} - \mathbf{r}_i) \cdot \mathbf{F} \quad (26)$$

$$\overleftrightarrow{T}(\mathbf{r}) = \frac{1}{8\pi\eta r} (\overleftrightarrow{1} + \hat{r} \otimes \hat{r}), \quad (27)$$

where $\hat{r}$ is the unit vector in the direction of $\mathbf{r}$.

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):

$$\frac{d}{dt} \mathbf{r}_i = \sum_j \overleftrightarrow{\mu}_{ij} \cdot \mathbf{F}_j + \boldsymbol{\rho}_i \quad (28)$$

where $\overleftrightarrow{\mu}_{ij}$ is the mobility tensor, and the stochastic displacements satisfy the relation

$$\langle \boldsymbol{\rho}_i(t) \otimes \boldsymbol{\rho}_j(t') \rangle = 2k_B T \overleftrightarrow{\mu}_{ij} \delta(t - t'). \quad (29)$$

We only consider mobility tensors which are divergence-free, and hence we need not worry about “spurious drift” terms¹. Now, since the monomers are essentially just “embedded” in the surrounding flow, we can identify (at least approximately) $\overleftrightarrow{\mu}_{ij} = \overleftrightarrow{\mu}(\mathbf{r}_i - \mathbf{r}_j)$ (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

$$D^{(K)} = \frac{D_0}{N} + \frac{k_B T}{6\pi\eta} \left\langle \frac{1}{R_H} \right\rangle, \quad (30)$$

and this differs only marginally from the long-time value²⁶. Since the second term strongly dominates in the long-chain limit, we find the scaling law

$$D \propto \frac{1}{R}, \quad (31)$$

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 whose physics had not fully been understood until very recently when a computer simulation¹⁵ clarified the last remaining puzzles.

The underlying question is: How does the system get rid of its hydrodynamic correlations? Early attempts^{27,28} 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 $\mathbf{u}$ experiences a friction force per unit volume of $-\zeta c \mathbf{u}$, such that the Stokes equation is modified to

$$\rho \frac{\partial}{\partial t} \mathbf{u} = \eta \nabla^2 \mathbf{u} - \zeta c \mathbf{u}. \quad (32)$$

Its Green’s function now exhibits a Debye-Hückel-like decay $\propto (1/r) \exp(-r/\xi_H)$, 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 Gennes²⁹ has criticized 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 puzzle²⁹ 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^{30,31}; 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 chains³², where a clear Zimm-like contribution was found. The results of our recent simulation¹⁵ 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, $\tau_\xi \sim \eta\xi^3/(k_B T)$. 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 Ref.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 $k^2 t^{2/z}$. Indeed we found Zimm behavior ($z = 3$) at short times and small length scales, while the data show RW

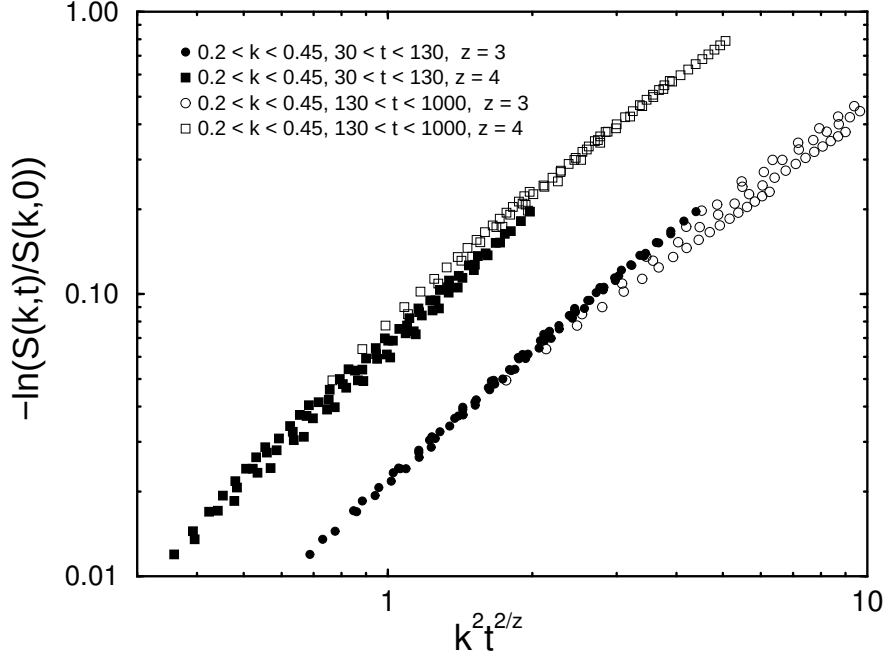


Figure 2. Scaling plot of single-chain dynamic structure factor data, for both Rouse and Zimm scaling, taken from Ref.¹⁵. For more details, see the text and Ref.¹⁵.

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, but for a challenging application like the system of Ref.¹⁵ it is necessary to choose and design the method carefully.

The most straightforward approach would be Brownian dynamics, where just Eqn. 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 Ref.²⁶ 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 interactions³³. 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)^{34–45}, 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 Ref.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 (see contribution by J. Baschnagel), 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)⁴⁶, where collisions are implemented via local stochastic updating rules which conserve energy and momentum. In practice, this is done by decomposing the system into boxes, and calculating the center-of-mass velocity of each box. Then every particle velocity belonging to the box (in the center-of-mass reference frame) is subjected to the same stochastic rotation. For further details on MPCD, see the original literature⁴⁶.

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)⁴⁷. The latter has become also quite popular for soft matter systems, in particular colloidal suspensions^{48–56}, 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)⁵⁷, 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:

$$\frac{d}{dt}q_i = \frac{\partial \mathcal{H}}{\partial p_i} \quad (33)$$

$$\frac{d}{dt}p_i = -\frac{\partial \mathcal{H}}{\partial q_i}, \quad (34)$$

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:

$$\frac{d}{dt}q_i = \frac{\partial \mathcal{H}}{\partial p_i} \quad (35)$$

$$\frac{d}{dt}p_i = -\frac{\partial \mathcal{H}}{\partial q_i} - \zeta_i \frac{\partial \mathcal{H}}{\partial p_i} + \sigma_i f_i; \quad (36)$$

here ζ_i is the friction coefficient for the i th 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. Refs.^{58,59} or⁶⁰). For the present case, one obtains

$$\frac{\partial}{\partial t}P = \mathcal{L}P, \quad (37)$$

where $\mathcal{L}$ is the Fokker–Planck operator, which is naturally decomposed into two parts,

$$\mathcal{L} = \mathcal{L}_H + \mathcal{L}_{SD}, \quad (38)$$

where the first part refers to the Hamiltonian part of the dynamics (it is nothing but the Liouville operator),

$$\begin{aligned} \mathcal{L}_H &= -\sum_i \frac{\partial}{\partial q_i} \frac{\partial \mathcal{H}}{\partial p_i} + \sum_i \frac{\partial}{\partial p_i} \frac{\partial \mathcal{H}}{\partial q_i} \\ &= -\sum_i \frac{\partial \mathcal{H}}{\partial p_i} \frac{\partial}{\partial q_i} + \sum_i \frac{\partial \mathcal{H}}{\partial q_i} \frac{\partial}{\partial p_i}, \end{aligned} \quad (39)$$

while the second part is due to friction and noise,

$$\mathcal{L}_{SD} = \sum_i \frac{\partial}{\partial p_i} \left[\zeta_i \frac{\partial \mathcal{H}}{\partial p_i} + \sigma_i^2 \frac{\partial}{\partial p_i} \right]. \quad (40)$$

In order to describe a system in thermal equilibrium, the Boltzmann distribution must be the stationary solution of the Fokker–Planck equation:

$$\mathcal{L} \exp(-\beta \mathcal{H}) = 0, \quad (41)$$

where $\beta = 1/(k_B T)$. For the Hamiltonian part, this relation is identically fulfilled. Therefore, the condition results in

$$\sum_i \frac{\partial}{\partial p_i} \left[\zeta_i \frac{\partial \mathcal{H}}{\partial p_i} - \beta \sigma_i^2 \frac{\partial \mathcal{H}}{\partial p_i} \right] \exp(-\beta \mathcal{H}) = 0. \quad (42)$$

Hence the relation

$$\sigma_i^2 = k_B T \zeta_i \quad (43)$$

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^{60,61} 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 Sec. 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^{40–45}. We will from now on exclusively focus on the thermostat aspect of DPD. As Espanol and Warren³⁶ 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^{62,63}). 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 Ref.⁶⁴). The DPD thermostat solves this problem in a very natural and straightforward way⁶⁵.

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

$$\sigma^2(r) = k_B T \zeta(r), \quad (44)$$

in close analogy to SD. The function has a finite range, such that only near neighbors are taken into account.

Defining $\mathbf{r}_{ij} = \mathbf{r}_i - \mathbf{r}_j = r_{ij} \hat{\mathbf{r}}_{ij}$, we then obtain the friction force on particle i by

projecting the relative velocities on the interparticle axes:

$$\mathbf{F}_i^{(fr)} = - \sum_j \zeta(r_{ij}) [(\mathbf{v}_i - \mathbf{v}_j) \cdot \hat{\mathbf{r}}_{ij}] \hat{\mathbf{r}}_{ij}; \quad (45)$$

it is easy to see that the relation $\sum_i \mathbf{F}_i^{(fr)} = 0$ holds. Similarly, we get the stochastic forces along the interparticle axes:

$$\mathbf{F}_i^{(st)} = \sum_j \sigma(r_{ij}) \eta_{ij}(t) \hat{\mathbf{r}}_{ij}, \quad (46)$$

where the noise η_{ij} satisfies the relations $\eta_{ij} = \eta_{ji}$, $\langle \eta_{ij} \rangle = 0$, and $\langle \eta_{ij}(t) \eta_{kl}(t') \rangle = 2(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk})\delta(t-t')$, such that different pairs are statistically independent. As before, one easily shows $\sum_i \mathbf{F}_i^{(st)} = 0$. The equations of motion,

$$\frac{d}{dt} \mathbf{r}_i = \frac{1}{m_i} \mathbf{p}_i, \quad (47)$$

$$\frac{d}{dt} \mathbf{p}_i = \mathbf{F}_i + \mathbf{F}_i^{(fr)} + \mathbf{F}_i^{(st)}, \quad (48)$$

where m_i is the mass of the i th particle, and vecp_i its momentum, therefore indeed conserve the total momentum, as the conservative forces $\mathbf{F}_i$ satisfy Newton's third law. The Fokker-Planck operator can then be written as

$$\mathcal{L} = \mathcal{L}_H + \mathcal{L}_{DPD}, \quad (49)$$

where $\mathcal{L}_H$ again describes the Hamiltonian part with $\mathcal{L}_H \exp(-\beta\mathcal{H}) = 0$ (cf. Eq. 39), and $\mathcal{L}_{DPD}$ is given by

$$\begin{aligned} \mathcal{L}_{DPD} &= \sum_{ij} \zeta(r_{ij}) \hat{\mathbf{r}}_{ij} \cdot \frac{\partial}{\partial \mathbf{p}_i} \left[\hat{\mathbf{r}}_{ij} \cdot \left(\frac{\partial \mathcal{H}}{\partial \mathbf{p}_i} - \frac{\partial \mathcal{H}}{\partial \mathbf{p}_j} \right) \right] \\ &\quad - \sum_{i \neq j} \sigma^2(r_{ij}) \left(\hat{\mathbf{r}}_{ij} \cdot \frac{\partial}{\partial \mathbf{p}_i} \right) \left(\hat{\mathbf{r}}_{ij} \cdot \frac{\partial}{\partial \mathbf{p}_j} \right) \\ &\quad + \sum_i \sum_{j(\neq i)} \sigma^2(r_{ij}) \left(\hat{\mathbf{r}}_{ij} \cdot \frac{\partial}{\partial \mathbf{p}_i} \right)^2 \\ &= \sum_i \sum_{j(\neq i)} \hat{\mathbf{r}}_{ij} \cdot \frac{\partial}{\partial \mathbf{p}_i} \left[\zeta(r_{ij}) \hat{\mathbf{r}}_{ij} \cdot \left(\frac{\partial \mathcal{H}}{\partial \mathbf{p}_i} - \frac{\partial \mathcal{H}}{\partial \mathbf{p}_j} \right) \right. \\ &\quad \left. + \sigma^2(r_{ij}) \hat{\mathbf{r}}_{ij} \cdot \left(\frac{\partial}{\partial \mathbf{p}_i} - \frac{\partial}{\partial \mathbf{p}_j} \right) \right]. \end{aligned} \quad (50)$$

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 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 ; $\mathbf{r}$ denotes its sites. Furthermore, we introduce a finite (small) set of (dimensionless) vectors $\mathbf{c}_i$, such that $a\mathbf{c}_i$ 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 $\mathbf{c}_i$ 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 $(a/h)\mathbf{c}_i$. An object residing on a certain lattice site $\mathbf{r}$, and having the velocity $(a/h)\mathbf{c}_i$, would thus be moved to site $\mathbf{r} + a\mathbf{c}_i$ within one time step. A commonly used model is the 18-velocity model, where the vectors $\mathbf{c}_i$ correspond to the nearest and next-nearest neighbors. Sometimes an additional velocity $\mathbf{c}_i = 0$ is included (19-velocity model); this is however not necessary for simulating incompressible flow. The algorithm now works with real-valued variables $n_i(\mathbf{r}, t)$, denoting the “number of particles” which reside on site $\mathbf{r}$ at time t and have the velocity $(a/h)\mathbf{c}_i$. Denoting the particle mass with m , we find for the mass density at site $\mathbf{r}$ at time t

$$\rho(\mathbf{r}, t) = \frac{m}{a^3} \sum_i n_i(\mathbf{r}, t) \quad (51)$$

and for the momentum density

$$\mathbf{j}(\mathbf{r}, t) = \frac{m}{a^2 h} \sum_i n_i(\mathbf{r}, t) \mathbf{c}_i. \quad (52)$$

We can also introduce the streaming velocity $\mathbf{u}$ at site $\mathbf{r}$ via $\mathbf{u} = \mathbf{j}/\rho$. 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 lecture, 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 developed⁶⁶, 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 $\mathbf{j}$ on each lattice site.
2. From ρ and $\mathbf{j}$, 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 $\mathbf{j}$ 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 $\mathbf{r}$. 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 $\sum_i n_i \mathbf{c}_i = \sum_i n_i^{eq} \mathbf{c}_i$ 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:

$$n_i \rightarrow n_i + \sum_j L_{ij}(n_j - n_j^{eq}). \quad (53)$$

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 quickly⁴⁹. In order to ensure mass and momentum conservation, the matrix should satisfy the conditions $\sum_i L_{ij} = 0$ and $\sum_i \mathbf{c}_i L_{ij} = 0$.

4. Streaming: The populations are displaced to new sites according to their velocities:

$$n_i(\mathbf{r}, t) \rightarrow n_i(\mathbf{r} + a\mathbf{c}_i, t + h). \quad (54)$$

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 ansatz⁴⁹

$$n_i^{eq} = \frac{a^3}{m} \rho \left(A_i + B_i \mathbf{c}_i \cdot \mathbf{u} \frac{h}{a} + C_i u^2 \frac{h^2}{a^2} + D_i (\mathbf{c}_i \cdot \mathbf{u})^2 \frac{h^2}{a^2} \right), \quad (55)$$

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 $\mathbf{j}$, as mentioned above.
- The stress tensor constructed from n_i^{eq} ,

$$\overset{\leftrightarrow}{\Pi} = \frac{m}{a^3} \frac{a^2}{h^2} \sum_i n_i^{eq} \mathbf{c}_i \otimes \mathbf{c}_i, \quad (56)$$

should have the hydrodynamic form

$$\overset{\leftrightarrow}{\Pi} = \rho c_s^2 \overset{\leftrightarrow}{1} + \rho \mathbf{u} \otimes \mathbf{u}; \quad (57)$$

here we have assumed the equation of state of an ideal gas with sound velocity c_s (other equations of state can be implemented⁶⁷).

- 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 $\mathbf{u} = 0$ both shells should contain the same number of particles. This is useful for numerical stability⁴⁹.

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

$$\overset{\leftrightarrow}{\Pi} = \frac{m}{a^3} \frac{a^2}{h^2} \sum_i n_i \mathbf{c}_i \otimes \mathbf{c}_i \quad (58)$$

towards $\overset{\leftrightarrow}{\Pi}^{eq}$. We now require that this process exhibits two relaxation rates, one for the trace of $\overset{\leftrightarrow}{\Pi}$, 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 require⁴⁹ 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 $A_i, \dots, D_i$ are used)⁴⁹. Via a Chapman–Enskog expansion one can show that this procedure yields hydrodynamic behavior in the macroscopic limit⁴⁹ *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 hydrodynamics⁶⁸, 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 literature⁴⁹.

When coupling this to a system of Brownian particles, one can use two methods: The original approach by Ladd^{49,50} 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 interactions⁵².

For polymer solutions, we found a point-particle approach^{69,8} 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 $-\zeta \mathbf{v}$ ($\mathbf{v}$ particle velocity), but rather $-\zeta(\mathbf{v} - \mathbf{u})$, where $\mathbf{u}$ 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 shown⁷⁰ 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 $\mathbf{v}$ and constant average force $\mathbf{F}$ through a fluid globally at rest. Our simulation procedure tells us that the force should be $\mathbf{F} = \zeta(\mathbf{v} - \mathbf{u})$, $\mathbf{u}$ 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

$$\mu = \frac{1}{\zeta} + \frac{1}{g\eta a}; \quad (59)$$

this relation has been checked numerically⁸. 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 $a \rightarrow 0$ 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 A Final Remark

Although the presented material is highly selective, strongly reflecting my own research, I hope the present lecture 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 for soft matter systems, with emphasis on hydrodynamics, is a quite active field of current research, and far from being closed.

Acknowledgments

It is a pleasure to 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.

References

1. M. Doi and S. F. Edwards. *The Theory of Polymer Dynamics*. Oxford University Press, Oxford, 1986.
2. P. E. Rouse. *J. Chem. Phys.*, 21:1272, 1953.
3. B. H. Zimm. *J. Chem. Phys.*, 24:269, 1956.
4. C. Pierleoni and J.-P. Ryckaert. *J. Chem. Phys.*, 96:8539, 1992.
5. W. Smith and D. C. Rapaport. *Mol. Sim.*, 9:25, 1992.
6. B. Dünweg and K. Kremer. *J. Chem. Phys.*, 99:6983, 1993.

7. A. G. Schlijper, P. J. Hoogerbrugge, and C. W. Manke. *J. Rheol.*, 39:567, 1995.
8. P. Ahlrichs and B. Dünweg. *J. Chem. Phys.*, 111:8225, 1999.
9. N. A. Spenley. *Europhys. Lett.*, 49:534, 2000.
10. A. Malevanets and J. M. Yeomans. *Europhys. Lett.*, 52:231, 2000.
11. K. Kremer and G. S. Grest. *J. Chem. Phys.*, 92:5057, 1990.
12. W. Paul, K. Binder, D. W. Heermann, and K. Kremer. *J. Phys. II (Paris)*, 1:37, 1991.
13. W. Paul, K. Binder, D. W. Heermann, and K. Kremer. *J. Chem. Phys.*, 95:7726, 1991.
14. M. Pütz, K. Kremer, and G. S. Grest. *Europhys. Lett.*, 49:735, 2000.
15. P. Ahlrichs, R. Everaers, and B. Dünweg. *Phys. Rev. E*, 64:040501 (R), 2001.
16. R. Chang and A. Yethiraj. *J. Chem. Phys.*, 114:7688, 2001.
17. C. F. Abrams, N.-K. Lee, and S. P. Obukhov. *Europhys. Lett.*, 59:391, 2002.
18. N. Kikuchi, A. Gent, and J. M. Yeomans. *Europ. Phys. J. A*, 9:63, 2002.
19. P. G. de Gennes. *Scaling Concepts in Polymer Physics*. Cornell University Press, Ithaca, 1979.
20. A. Yu. Grosberg and A. R. Khokhlov. *Statistical Physics of Macromolecules*. AIP Press, New York, 1994.
21. B. Alder and T. Wainwright. *Phys. Rev. A*, 1:18, 1970.
22. W. Paul, G. D. Smith, D. Y. Yoon, B. Farago, S. Rathgeber, A. Zirkel, L. Willner, and D. Richter. *Phys. Rev. Lett.*, 80:2346, 1998.
23. W. Paul. *Chemical Physics*, 284:59, 2002.
24. V. G. Rostashvili, M. Rehkopf, and T. A. Vilgis. *Europ. Phys. J. B*, 6:497, 1998.
25. M. Guenza. *Phys. Rev. Lett.*, 88:025901, 2002.
26. B. Liu and B. Dünweg. *J. Chem. Phys.*, 118:8061, 2003.
27. S. F. Edwards and K. F. Freed. *J. Chem. Phys.*, 61:1189, 1974.
28. K. F. Freed and S. F. Edwards. *J. Chem. Phys.*, 61:3626, 1974.
29. P.-G. de Gennes. *Macromolecules*, 9:594, 1976.
30. P. Wiltzius, H. R. Haller, and D. S. Carnell. *Phys. Rev. Lett.*, 53:834, 1984.
31. P. Wiltzius and D. S. Carnell. *Phys. Rev. Lett.*, 56:61, 1986.
32. D. Richter, K. Binder, B. Ewen, and B. Stühn. *J. Phys. Chem.*, 88:6618, 1984.
33. A. S. Sangani and G. Mo. *Phys. Fluids*, 8:1990, 1996.
34. P. J. Hoogerbrugge and J. M. V. A. Koelman. *Europhys. Lett.*, 19:155, 1992.
35. J. M. V. A. Koelman and P. J. Hoogerbrugge. *Europhys. Lett.*, 21:369, 1993.
36. P. Español and P. Warren. *Europhys. Lett.*, 30:191, 1995.
37. P. Español. *Phys. Rev. E*, 52:1734, 1995.
38. R. D. Groot and P. B. Warren. *J. Chem. Phys.*, 107(11):4423, 1997.
39. P. Español. *Phys. Rev. E*, 57:2930, 1998.
40. I. Pagonabarraga, M. J. H. Hagen, and D. Frenkel. *Europhys. Lett.*, 42:377, 1998.
41. J. B. Gibson, K. Chen, and S. Chynoweth. *Int. J. Mod. Phys. C*, 10:241, 1999.
42. G. Besold, I. Vattulainen, M. Karttunen, and J. M. Polson. *Phys. Rev. E*, 62:R7611, 2000.
43. I. Vattulainen, M. Karttunen, G. Besold, and J. M. Polson. *J. Chem. Phys.*, 116:3967, 2002.
44. P. Nikunen, M. Karttunen, and I. Vattulainen. *Comp. Phys. Comm.*, 153:407, 2003.
45. T. Shardlow. *SIAM J. Sci. Comp.*, 24:1267, 2003.
46. A. Malevanets and R. Kapral. *J. Chem. Phys.*, 110:8605, 1999.
47. S. Succi. *The Lattice Boltzmann Equation for Fluid Dynamics and Beyond*. Oxford

University Press, Oxford, 2001.

48. A. J. C. Ladd. *Phys. Rev. Lett.*, 70:1339, 1993.
49. A. J. C. Ladd. *J. Fluid Mech.*, 271:285, 1994.
50. A. J. C. Ladd. *J. Fluid Mech.*, 271:311, 1994.
51. A. J. C. Ladd and R. Verberg. *J. Stat. Phys.*, 104:1191, 2001.
52. A. J. C. Ladd, Hu Gang, J. X. Zhu, and D. A. Weitz. *Phys. Rev. E*, 52:6550, 1995.
53. C. P. Lowe, D. Frenkel, and A. J. Masters. *J. Chem. Phys.*, 103:1582, 1995.
54. M. H. J. Hagen, I. Pagonabarraga, C. P. Lowe, and D. Frenkel. *Phys. Rev. Lett.*, 78:3785, 1997.
55. M. H. J. Hagen, D. Frenkel, and C. P. Lowe. *Physica A*, 272:376, 1999.
56. M. W. Heemels, M. H. J. Hagen, and C. P. Lowe. *J. Comput. Phys.*, 164:48, 2000.
57. T. Schneider and E. Stoll. *Phys. Rev. B*, 17:1302, 1978.
58. S. Chandrasekhar. *Rev. Mod. Phys.*, 15:1, 1943.
59. H. Risken. *The Fokker–Planck Equation*. Springer–Verlag, Berlin, 1984.
60. B. Dünweg. Langevin methods. In B. Dünweg, D. P. Landau, and A. I. Milchev, editors, *Computer Simulations of Surfaces and Interfaces*. Kluwer, Dordrecht, 2003.
61. B. Dünweg. *J. Chem. Phys.*, 99:6977, 1993.
62. F. Müller-Plathe. *J. Chem. Phys.*, 106:6082, 1997.
63. F. Müller-Plathe. *Phys. Rev. E*, 59(5):4894, 1999.
64. D. J. Evans and G. P. Morriss. *Statistical Mechanics of Nonequilibrium Liquids*. Academic Press, London, 1990.
65. T. Soddemann, B. Dünweg, and K. Kremer. *Phys. Rev. E*, 68:046702, 2003.
66. T. Ihle and D. M. Kroll. *Comp. Phys. Comm.*, 129:1, 2000.
67. M. R. Swift, S. E. Orlandini, W. R. Osborn, and J. M. Yeomans. *Phys. Rev. E*, 54:5041, 1996.
68. L. D. Landau and E. M. Lifshitz. *Fluid mechanics*. Addison-Wesley, Reading, 1959.
69. P. Ahlrichs and B. Dünweg. *Int. J. Mod. Phys. C*, 9:1429, 1998.
70. V. Lobaskin and B. Dünweg, submitted to New J. Phys.

Monte Carlo Simulation of Polymers: Coarse-Grained Models

J. Baschnagel, J. P. Wittmer, and H. Meyer

Institut Charles Sadron,
6, rue Boussingault, 67083 Strasbourg Cedex, France
E-mail: {baschnag, jwittmer, hmeyer}@ics.u-strasbg.fr

A coarse-grained simulation model eliminates microscopic degrees of freedom and represents a polymer by a simplified structure. A priori, two classes of coarse-grained models may be distinguished: those which are designed for a specific polymer and reflect the underlying atomistic details to some extent, and those which retain only the most basic features of a polymer chain (chain connectivity, short-range excluded-volume interactions, etc.). In this article we mainly focus on the second class of generic polymer models, while the first class of specific coarse-grained models is only touched upon briefly. Generic models are suited to explore general and universal properties of polymer systems, which occur particularly in the limit of long chains. The simulation of long chains represents a challenging problem due to the large relaxation times involved. We present some of the Monte Carlo approaches contrived to cope with this problem. More specifically, our review contains two main parts. One part (Sec. 5) deals with local and non-local updates of a polymer. While local moves allow to extract information on the physical polymer dynamics from Monte Carlo simulations, the chief aim of non-local moves is to accelerate the relaxation of the polymers. We discuss some examples for such non-local moves: the slithering-snake algorithm, the pivot algorithm, and its recently suggested variant, the double-pivot algorithm, which is particularly suited for the simulations of concentrated polymer solutions or melts. The second part (Sec. 6) focuses on modern Monte Carlo methods that were inspired by the Rosenbluth-Rosenbluth algorithm proposed in the 1950s to simulate self-avoiding walks. The modern variants discussed comprise the configuration-bias Monte Carlo method, its recent extension, the recoil-growth algorithm, and the pruned-enriched Rosenbluth method, an algorithm particularly adapted to the simulation of attractively interacting polymers.

1 Introduction

Polymers are macromolecules in which N monomeric repeat units are connected to form long chains.^a Experimentally the chain length N is large, typically $10^3 \lesssim N \lesssim 10^5$. The size of a chain ($\sim 10^3 \text{ \AA}$) thus exceeds that of a monomer ($\sim 1 \text{ \AA}$) by several orders of magnitude. However, contrary to granular materials,² the chain is not so large that thermal energy^b would be unimportant. Not at all! Thermal energy is the important energy scale for polymers. It provokes conformational transitions so that the polymer can assume a multitude of different configurations at ambient conditions.^c

^aMore precisely, this definition refers to “linear homopolymers”, i.e., linear chain molecules consisting of one monomer species only. By contrast, polymer chemistry can nowadays synthesize various other topologies, such as cyclic, star- or H-polymers. For a very commendable review on the physical chemistry of polymers see Ref. 1.

^bHere, we mean the thermal energy supplied at ambient temperature, i.e., $k_B T = 4.1 \cdot 10^{-21} \text{ J}$ for $T = 300 \text{ K}$.

^cPolymers are a paradigm for “soft matter” materials or “complex fluids”. Roughly speaking, “soft matter” consists of materials whose constituents have a mesoscopic size (microscopic scale $\sim 1 \text{ \AA} \ll$ mesoscopic object $\sim 10^2 - 10^4 \text{ \AA} \ll$ macroscopic scale $\sim 1 \text{ mm}$) and for which $k_B T$ is the important energy scale (whence the softness at ambient conditions). Examples other than polymers are colloidal suspensions, liquid crystals, or fluid membranes.³

Changes of the configurations occur on very different scales, ranging from the local scale of a bond to the global scale of the chain.^{4,5} This separation of length scales entails simplifications and difficulties. Simplifications arise on large scales where the chain exhibits universal behavior. That is, properties which are independent of chemical details.^{6,7} These properties may be studied by simplified, “coarse-grained” models, e.g. via computer simulations. For simulations the large-scale properties, however, also give rise to a principal difficulty. Long relaxation times are associated with large chain lengths.⁷⁻⁹

The present chapter focuses on some of the Monte Carlo approaches to cope with this difficulty. Why Monte Carlo? Within a computational framework it appears natural to address dynamical problems via the techniques of Molecular Dynamics (see Ref. 10). A Molecular Dynamics (MD) simulation numerically integrates the equations of motion of the (polymer) system, and thereby replicates, authentically, its (classical) dynamics. As the polymer dynamics ranges from the (fast) local motion of the monomers to (slow) large-scale rearrangements of a chain, there is a large spread in time scales. The authenticity of MD thus carries a price: Efficient equilibration and sampling of equilibrium properties becomes very tedious –sometimes even impossible– for long chains. At that point, Monte Carlo simulations may provide an alternative. Monte Carlo moves are not bound to be local. They can be tailored to alter large portions of a chain, thereby promising efficient equilibration. The discussion of such moves is one of the gists of this review.

Outline and Prerequisites. The plan of the chapter is as follows: We begin by gathering necessary background information, both as to polymer physics (Sec. 2) and as to the Monte Carlo method (Sec. 3). Then, we present the simulation models (Sec. 4), which have been used to develop and to study various Monte Carlo algorithms. The discussion of these algorithms (Secs. 5 and 6) represents the core of the chapter. Section 5 deals with local moves, allowing to study the physical polymer dynamics via Monte Carlo, and non-local moves (slithering-snake algorithm, pivot algorithm, double-pivot algorithm), aiming at speeding up the relaxation of the chains. Section 6 discusses the Rosenbluth-Rosenbluth method for simulating self-avoiding walks and some of its modern variants (pruned-enriched Rosenbluth method, configuration-bias Monte Carlo, recoil-growth algorithm). The last section (Sec. 7) briefly recapitulates the different methods and gives some advice when to employ which algorithm. Finally, the appendix 7 reviews a recently proposed approach to systematically derive coarse-grained models for specific polymers.

Our presentation is based on the following prerequisites:

- We will restrict our attention to homopolymers, i.e., to polymers consisting of one monomer species only. However, (some of) the algorithms discussed may also be applied e.g. to polymer blends or block-copolymers (see Ref. 11).
- The chains are monodisperse, i.e, N is constant.
- We do not consider long-range (e.g., electrostatic) or specific (e.g., H-bonds) interactions between the monomers. These interactions are treated in other chapters (e.g., see Refs. 12, 13).
- We do not treat the solvent molecules explicitly. They are indirectly accounted for by the interactions between the monomers. The neglect of the solvent does not affect

the static properties of chains in dilute solution. However, it does affect their physical dynamics (see Ref. 14).

2 A Primer to Polymer Physics

2.1 A Polymer in Good Solvent

To substantiate the remarks of the introduction about the large-scale properties of polymers let us consider a specific example, a dilute solution of polyethylene. Polyethylene consists of CH_2 -monomers which are joined to form a linear polymer (Fig. 1). A configuration of the chain may be specified by the positions of the monomers^d $\mathbf{x} = (\mathbf{r}_1, \dots, \mathbf{r}_N)$. Thermodynamic properties are calculated by averaging an observable A over all configurations

$$\langle A \rangle = \frac{1}{\mathcal{Z}} \int d\mathbf{x} A(\mathbf{x}) \exp [-\beta U(\mathbf{x})] . \quad (1)$$

Here $\beta = k_B T$, $\mathcal{Z}$ is the partition function and $U(\mathbf{x})$ the interaction potential. We assume that $U(\mathbf{x})$ can be split into two parts:^e

$$U(\mathbf{x}) = \sum_{i=1}^{N-1} \underbrace{U_0(\mathbf{b}_i, \dots, \mathbf{b}_j, \dots, \mathbf{b}_{i+i_{\max}})}_{\text{"short-range": } \ell, \theta, \phi, \dots} + \underbrace{U_1(\mathbf{x}, \text{solvent})}_{\text{"long-range"}}, \quad (2)$$

where $\mathbf{b}_i = \mathbf{r}_{i+1} - \mathbf{r}_i$ denotes the bond vector from the i th to the $(i+1)$ th monomer.

The first term of Eq. (2), U_0 , depends on the chemical nature of the polymer. It comprises the potentials of the bond length ℓ , the bond angle θ , the torsional angle ϕ , etc. (Fig. 1).¹⁷ These potentials lead to correlations between the bond vectors $\mathbf{b}_i$ and $\mathbf{b}_j$. Typically, the correlations are of short range: they only extend up to some bond vector $\mathbf{b}_{j=i+i_{\max}}$ with $i_{\max} \ll N$.

Although distant monomers along the backbone of the chain are thus orientationally decorrelated, they can still come close in space. The resulting interaction is long-range along the chain backbone (Fig. 1). In Eq. (2), it is accounted for by the second term U_1 .^{6,7} U_1 depends strongly on the quality of the solvent.^f In good solvents the monomers effectively repel one another (they want to be surrounded by solvent molecules), whereas they attract each other if the solvent cannot dissolve the polymer (bad solvent).

Due to its long-range character, one intuitively expects U_1 to influence the large-scale behavior of the chain more strongly than U_0 . A possible test of this idea is to estimate how the size of a chain scales with N . Common measures of the chain size are the mean-square end-to-end distance R_e^2 or the radius of gyration R_g^2 (Fig. 1)

$$R_e^2 = \langle (\mathbf{r}_N - \mathbf{r}_1)^2 \rangle, \quad R_g^2 = \frac{1}{N} \sum_{i=1}^N \langle (\mathbf{r}_i - \mathbf{R}_{\text{cm}})^2 \rangle, \quad (3)$$

^dHere, we adopt a description in terms of a so-called “united atom model”. The united atom model represents a CH_2 -group by a single, spherical interaction site and does not distinguish between inner (CH_2) and end monomers (CH_3).¹⁵ Furthermore, we neglect the momenta of the monomers to specify the configuration, as we assume the observables and interaction potentials to depend on positions only.

^eEquation (1) does not contain the degrees of freedom of the solvent. They are assumed to be integrated out. Thus, $U(\mathbf{x})$ is an effective potential—in fact, a free energy—depending on the properties of the solvent.

^fIn Eq. (2) we assume that U_0 is independent of the solvent quality.

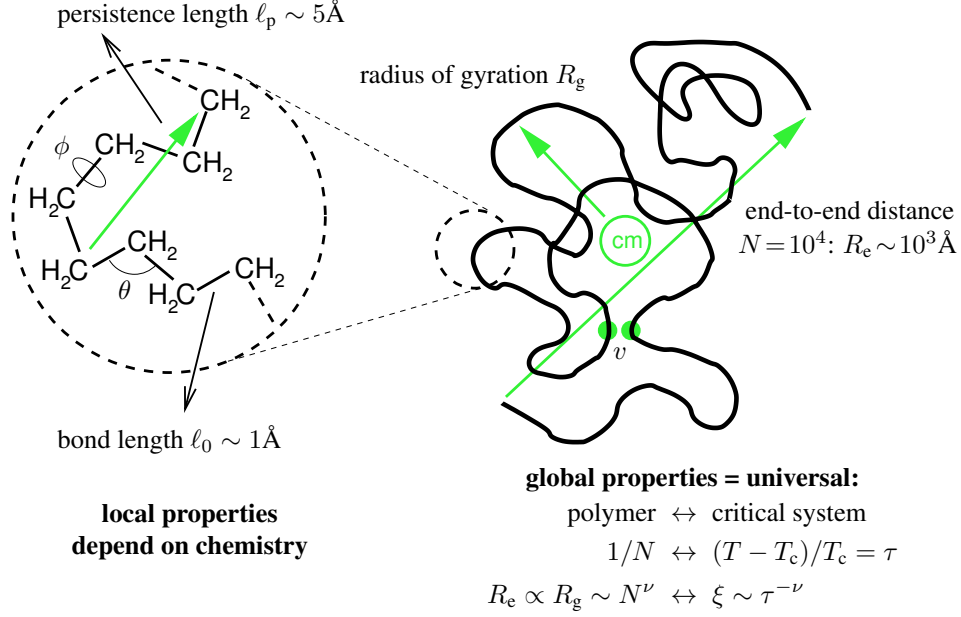


Figure 1. Schematic illustration of polyethylene. The local properties of the polymer depend on its microscopic degrees of freedom: the bond length ℓ , the bond angle θ , and the torsional angle ϕ . Because the potential of the bond length is fairly “stiff”, ℓ may be kept fixed at its equilibrium value ℓ_0 in a modeling approach. By contrast, the potential of the torsional angle is much “softer”. Thus, ϕ , which characterizes rotations about a middle C-C bond, mainly determines the local conformation of the chain. All degrees of freedom (ℓ, θ, ϕ) determine the intrinsic stiffness of the chain. The stiffness reflects the persistence of orientational correlations along the backbone of the chain. Orientational correlations decouple on the length scale of the “persistence length” ℓ_p . For typical chain lengths, $N \sim 10^4$, ℓ_p is much smaller than the end-to-end distance R_e or the radius of gyration R_g . (R_g measures the average distance of a monomer from the center of mass (cm) of the chain.) Thus, the chain appears flexible on length scales larger than ℓ_p . If the polymer is dissolved in a good solvent, distant monomers (filled grey circles) repel each other when they come in contact. That is, the excluded-volume parameter v , measuring the effective interaction between distant monomers along the chain, is positive. Under these conditions (i.e., linear polymer with some flexibility and repulsive monomer-monomer interactions) a correspondence between the large-scale properties of the polymer and a critical system close to its phase transition can be established.^{6,16} $1/N$ may be identified with the reduced distance, τ , to the critical temperature T_c of the phase transition, and R_e or R_g scale with N as the correlation length ξ of the order parameter does with τ . ν is a universal critical exponent, often called “Flory exponent” in polymer science.

where $\mathbf{R}_{\text{cm}}$ is the position of the chain’s center of mass. Because $R_e \propto R_g$ we focus on R_e in the sequel to illustrate the role played by U_0 and U_1 .

Let $\hat{b}_i$ denote the unit vector associated with the bond $\mathbf{b}_i$ of fixed length ℓ_0 (Fig. 1). Then, quite generally, we may write R_e^2 as

$$R_e^2 = \ell_0^2 \sum_{i=1}^{N-1} \sum_{j=1}^{N-1} \langle \hat{b}_i \cdot \hat{b}_j \rangle = 2\ell_0^2 \sum_{i=1}^{N-1} \sum_{k=0}^{N-1-i} \langle \hat{b}_i \cdot \hat{b}_{i+k} \rangle - (N-1)\ell_0^2. \quad (4)$$

Apparently, the large-scale behavior of R_e depends on the range of orientational correlations between bond vectors. Two cases may be distinguished:⁹

⁹In part, the subsequent discussion closely follows that on p. 148 of Ref. 18.

1. If $\langle \hat{b}_i \cdot \hat{b}_{i+k} \rangle$ is “short-range”, i.e., if it decays more rapidly than $1/k$ for large k , the second term converges in the large- N limit. Then,

$$R_e^2 = N\ell_0^2 \left[2 \sum_{k=0}^{\infty} \langle \hat{b}_1 \cdot \hat{b}_{1+k} \rangle - 1 \right] =: N\ell_0^2 \left[2 \frac{\ell_p}{\ell_0} - 1 \right] \quad (N \rightarrow \infty), \quad (5)$$

where we introduce the persistence length ℓ_p in the last term. (ℓ_p measures the “persistence” of orientational correlations along the backbone and thus the intrinsic stiffness of the chain; see Fig. 1). Equation (5) shows that short-range orientational correlations only affect the prefactor—they renormalize the bond length to $b = \ell_0[2(\ell_p/\ell_0) - 1]^{1/2}$ (b is called “effective bond length”⁷)—but they do not change the scaling of R_e with N . The scaling is always “random-walk-like”:^h $R_e \sim N^{1/2}$.⁷ In polymer science, a chain exhibiting this random-walk-like behavior is commonly referred to as an “ideal chain”.

Of course, the finite-range correlations, assumed for U_0 in Eq. (2), are also of short range. Thus, provided $U_1 = 0$, the end-to-end distance of a (long) chain is given by $R_e = bN^{1/2}$, irrespective of the precise form of U_0 . The microscopic degrees of freedom, ℓ, θ, ϕ , determine the prefactor, the effective bond length b , but not the scaling with N . Therefore, if we are interested in studying large-scale properties, we can replace a chemically realistic model for polyethylene by a much simpler “coarse-grained model”, which is microscopically unrealistic, but correctly reproduces the large- N behavior. An example for such a coarse-grained model is a “bead-spring model”, where N effective monomers (“beads”) are connected by harmonic springs of average length b (Fig. 2).

2. However, if $\langle \hat{b}_i \cdot \hat{b}_{i+k} \rangle$ decays as $1/k$ or more slowly (as $1/k^y$ with $y < 1$) due to long-range correlations, the scaling behavior of R_e^2 is changed. Instead of $R_e^2 \sim N$ we find

$$R_e^2 \sim N\ell_0^2 \int_0^N dk \langle \hat{b}(k) \cdot \hat{b}(0) \rangle \sim \begin{cases} N^{2-y} & (y < 1), \\ N \ln N & (y = 1). \end{cases} \quad (6)$$

Thus, long-range correlations lead to a “swelling” of the chain size with respect to a pure random walk.

Such long-range correlations are embodied in the potential U_1 in Eq. (2). For a polymer in a good solvent a swelling of the chain dimension is physically reasonable. As soon as two (distant) monomers come close in space, they repel each other. On the level of the coarse-grained bead-spring model we can incorporate this repulsive interaction by writing U_1 as (see e.g. Ref. 7 or the lucid discussion on pp. 16–20 of Ref. 16)

$$U_1(\mathbf{r}^N) = \int d^3\mathbf{r} \frac{1}{2} k_B T v \rho(\mathbf{r})^2 \quad \text{with} \quad \rho(\mathbf{r}) = \sum_{i=1}^N \delta(\mathbf{r} - \mathbf{r}_i). \quad (7)$$

^hBy the term “random-walk-like” we mean the diffusional motion of a Brownian particle. This motion can be thought of as resulting from the addition of many small displacements in random directions so that the overall mean-square displacement of the particle in time t , $R^2(t) = \langle [\mathbf{r}(t) - \mathbf{r}(0)]^2 \rangle$, scales as $R^2 \sim t$. This allows for the following identifications in regard to polymer physics: $R \leftrightarrow R_e$ and $t \leftrightarrow N$.

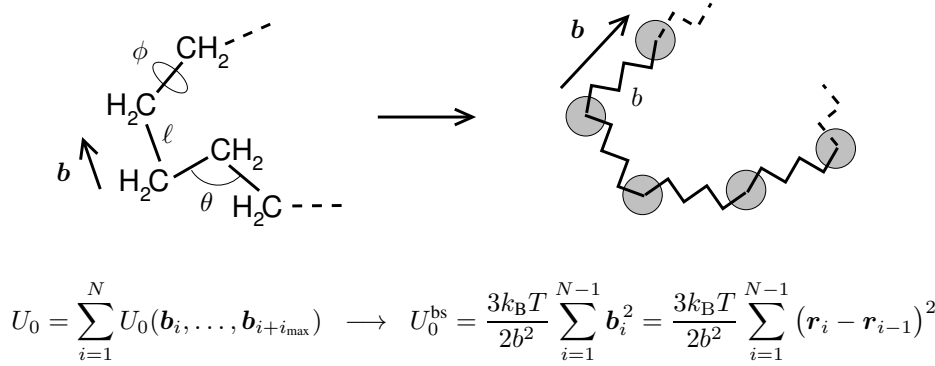


Figure 2. From a chemically realistic model to a coarse-grained bead-spring model. Local properties of the realistic model are determined by its microscopic degrees of freedom: ℓ , θ , and ϕ . On the global level of the chain, however, the influence of the microscopic degrees of freedom can be lumped into one parameter, the effective bond length b . The microscopic degrees of freedom do not determine the scaling of the end-to-end distance, which, under the sole effect of U_0 , is given by $R_e = bN^{1/2}$ (“ideal chain”). This behavior may be recovered from Eq. (1) when calculating R_e with the potential U_0^{bs} of a coarse-grained bead-spring model. This model identifies the monomers with spherical “beads” which are bound to one another by harmonic springs with force constant $3k_{\text{B}}T/b^2$. (This bead-spring model is often called “Gaussian chain” model in the polymer literature.⁷⁾

Here, $\rho(\mathbf{r})$ is the monomer density at point $\mathbf{r}$ and v (> 0) is the excluded-volume parameter. v measures the strength of the repulsion of a binary contact between two beads. Because a binary contact occurs with probability $\rho(\mathbf{r})^2$, Eq. (7) expresses the total energy penalty resulting from the repulsive contacts of all beads in the chain.

From the previous discussion of U_0 and U_1 the following conclusion may be drawn: When focusing on the large-scale properties of linear polymers with some flexibility and predominantly repulsive interactions we may forego a microscopic description in favor of a coarse-grained model. An example is the bead-spring model introduced above (Fig. 2), which is characterized by two parameters, b and v . Another possibility is a self-avoiding walk (SAW) on a (hyper-cubic) lattice. That is, a random walk which is not allowed to visit an already occupied lattice site again (see Sec. 4.1). The replacement “microscopic model $\rightarrow$ SAW” is permissible because a linear polymer in good solvent can be shown to correspond to a critical system which undergoes a phase transition for $N \rightarrow \infty$ (Fig. 1). It belongs to the universality class of the n -vector model in the limit $n \rightarrow 0$.^{6,16} This implies that the large- N behavior is determined by critical exponents. For instance,

$$R_g \propto R_e = bN^\nu \quad \text{or} \quad \mathcal{Z} \sim \mu^N N^{\gamma-1} \quad (N \rightarrow \infty), \quad (8)$$

where the partition function $\mathcal{Z}$ counts the number of N -step SAW’s starting at the origin and ending anywhere. The connectivity constant μ and the bond length b are non-universal. They depend on the polymer and the external conditions (temperature, solvent, etc.). By contrast, the critical exponents ν and γ are universal. They only depend on the dimension of space.ⁱ Thus, they can be determined for all polymers by studying this (simple) model.

ⁱIn the course of the research on critical phenomena it has become clear that all systems with short-range, isotropic

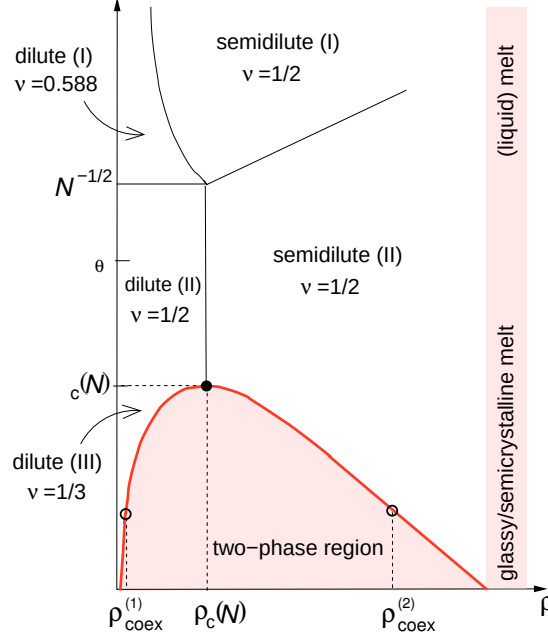


Figure 3. Schematic phase diagram of flexible polymers (see Chap. 9 of Ref. 16 or Chap. 4 of Ref. 23). For small monomer density ρ the solution is dilute. Three different regimes may be distinguished according to the temperature T : swollen chains [Eq. (8), $T > T_\Theta$: dilute (I)], nearly ideal chains [Eq. (10), $T \approx T_\Theta$: dilute (II)], and collapsed chains [Eq. (11), $T < T_\Theta$: dilute (III)]. There is an interval ΔT around the Θ -point of order $\Delta T/T_\Theta \sim 1/\sqrt{N}$, where the chains are nearly ideal. Whereas the chains may be considered as being isolated in dilute solution, they strongly overlap in the semidilute regimes. For $T \leq T_c(N)$ phase separation in a dilute phase of collapsed chains and a semidilute solution of nearly ideal chains occurs. If the monomer density approaches 1, we obtain a polymer melt. At high T the melt is a (viscous) liquid, whereas at low T it may become a glassy²⁴ or a semicrystalline²⁵ solid, depending on the ability of the polymer to form ordered structures or not.

In fact, the currently most precise values of ν and γ (see footnote on page 112) have been obtained from high-precision Monte Carlo simulations of SAW's.^{21,22}

2.2 Phase Diagram of a Polymer Solution

The utility of coarse-grained models to investigate the statistical physics of polymer systems is not limited to the previous example. A dilute solution in a good solvent is just one region in the phase diagram. The phase diagram of flexible polymers is schematically shown in Fig. 3. Out of the various regimes we choose to discuss two cases in more detail, a chain in another than good solvent and (high-temperature) polymer melts. In the following sections we concentrate on those cases because novel Monte Carlo approaches have been applied to them.

interactions, the same dimension of space d , and the same dimensionality n of the order parameter ($n = 1$: scalar, $n \geq 2$: n -dimensional vector) have critical exponents which depend only on (d, n) and take the same values as those of the n -vector model.^{19,20}

A Chain in a Θ -Solvent or a Bad Solvent. To extend the discussion of the good solvent to other solvents let us reconsider Eq. (7). This equation corresponds to the first term of a virial expansion in the monomer density $\rho(\mathbf{r})$. That is,^{7,16}

$$U_1 = \int d^3\mathbf{r} \left[\frac{1}{2}k_B T v \rho(\mathbf{r})^2 + \frac{1}{6}k_B T w \rho(\mathbf{r})^3 + \dots \right]. \quad (9)$$

This identifies the excluded-volume parameter v with the second virial coefficient. In general, the virial coefficients depend on temperature T . The second virial coefficient vanishes at some temperature, called “ Θ -temperature T_Θ ” in polymer science, and behaves as $v = v_0(1 - T_\Theta/T)$ close to the Θ -point ($v_0 = \text{const.} > 0$). This implies that we can tune the solvent quality by temperature. In addition to the case of a good solvent ($T > T_\Theta$) two further cases may be distinguished:

1. *Θ -solvent* ($T = T_\Theta$): Since binary interactions are absent [but ternary interactions are present: $w > 0$ in Eq. (9)], the polymer behaves nearly as an ideal chain:¹⁶

$$R_e \propto R_g \sim \sqrt{N} \quad (+ \ln N \text{ corrections}). \quad (10)$$

2. *Bad solvent* ($T < T_\Theta$): Since the binary interactions are attractive, the polymer is collapsed to a dense sphere of monomers, implying that the average monomer density ρ inside the sphere is of order 1. Thus,

$$\rho \approx \frac{N}{R_g^3} \sim 1 \quad \Rightarrow \quad R_e \propto R_g \sim N^\nu \quad \text{with} \quad \nu = \frac{1}{3}. \quad (11)$$

The simulation of this situation is complicated because the time to equilibrate the chain and to sample equilibrium properties from many independent configurations becomes exceedingly long. Two factors are responsible for that. On the one hand, the local dynamics is sluggish (maybe even glass-like) due to the dense packing of monomers that strongly attract each other. On the other hand, the polymer is entangled with itself. Bonds cannot pass through each other. These topological constraints may also lead to slow dynamics for long chains.

The Size of a Chain in a Polymer Melt. In a good solvent a chain expands with respect to the ideal state, owing to long-range monomer-monomer repulsions. This is peculiar to dilute solutions. In a dense liquid of chains, a “polymer melt”, the situation is quite different. One can show^{6,7,26} that the intra-chain excluded-volume interactions are screened by the presence of the surrounding polymers. Thus, a chain in a melt behaves on large scales as an ideal chain, implying $R_e \propto R_g \sim N^{1/2}$ (see Fig. 4).

This ideality, first proposed by Flory,¹⁷ appears fairly unexpected. Some feeling why this should be so may be obtained from the following argument: Inside the volume of a chain ($\sim R_g^3$) the monomer density resulting from the N monomers of the chain is very small. For ideal chains it is of order $N/R_g^3 \sim N^{-1/2}$, whereas it scales as $\sim N^{-0.764}$ under good solvent conditions (using Eq. (8) and $\nu = 0.588$). We see that in dilute solution, swelling reduces the monomer density inside the chain and thus the total interaction energy [see Eq. (7)]. However, no energetic advantage may be gained in a melt because the overall monomer density is $\rho \sim 1$. Swelling would reduce the number of intra-chain contacts, but this reduction must be compensated by inter-chain contacts to keep ρ constant. Thus, a chain has to have $N^{1/2}$ contacts with other chains, which is huge in the large- N limit. This strong interpenetration of the chains suppresses the expansion of an individual chain.

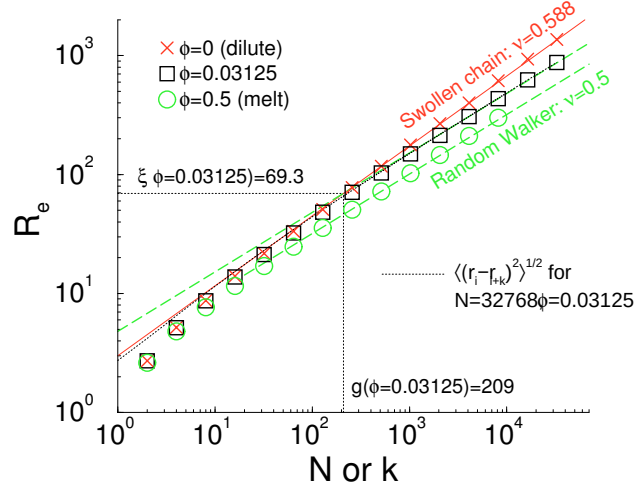


Figure 4. End-to-end distance R_e versus chain length N for the (athermal) bond-fluctuation model which will be discussed in more detail in Secs. 4.1 and Sec. 5. Results for three volume fractions (of occupied lattice sites) are given, illustrating the dilute ($\phi = 0$), the semidilute ($\phi = 0.03125$) and the melt ($\phi = 0.5$) limits of the schematic phase diagram (Fig. 3). Using the slithering-snake algorithm (Sec. 5.2) it is possible to simulate chains containing up to $N = 32768$ monomers for $\phi < 0.5$. Since the slithering-snake algorithm becomes less efficient at high densities (Sec. 5.2), the recently proposed double-pivot algorithm, described in Sec. 5.3, was harnessed to probe systems of higher densities ($\phi = 0.5$). Periodic boxes of linear size $L = 512$ and containing up to 2^{22} monomers are required to eliminate finite-size effects. Such periodic boundary conditions are not needed for single chains. Here, an infinite box was used ($L(\phi = 0) = \infty$). As only excluded-volume interactions are taken into account, good solvent statistics applies in dilute solution. The chains are swollen, as indicated by the exponent $\nu = 0.588$ (solid line), which fits the data over three orders of magnitude. In the opposite (so-called) melt limit long-range correlations appear to be screened down to small chain lengths of about $N \approx 10$ (grey dashed line).²⁷ Both chain statistics are visible for the intermediate density ($\phi = 0.03125$): Small chains ($N \ll g$, $R_e \ll \xi$) are swollen (solid line) and long chains are Gaussian (dashed line). The intercept of both lines defines the size ξ of the “excluded volume blob”^{6,7} and the number of monomers g that the blob contains. The indicated numbers are specific to the volume fraction (and persistence length) given, but are independent of chain length. For a given density ξ corresponds to the chain size where the coils start to overlap. Also presented in the figure is the spatial distance $\langle (r_i - r_{i+k})^2 \rangle^{1/2}$ along the longest chain for $\phi = 0.03125$ (dotted line). With the exception of small N or k (i.e., $N, k \ll 10$) this distance is, within the numerical accuracy of the data, identical to $R_e(N)$ with $N = k$. This agreement also demonstrates that the difference between a segment of a long chain and a chain having the same length as the segment becomes irrelevant for distances larger than ξ . In precisely this sense the (long-range) excluded volume interactions are screened in semidilute solutions and in melts. Mean-field descriptions become appropriate on the level of coarse-grained (Gaussian) chains of blobs.^{6,7}

2.3 Dynamics of Polymer Melts: Rouse and Reptation Models

The Rouse Model. As a monomer in a dilute solution moves, it creates a vortex (“wave”) in the solvent. The solvent transports the “wave” which is transported to other monomers of the chains so that a coupling between the motion of (distant) monomers arises (see Ref. 14). This long-range hydrodynamic interaction becomes screened by other chains when the concentration of the solution increases.⁷ In a dense melt, hydrodynamic interactions are completely suppressed. Thus, it is generally believed that the Rouse theory^{6,7} provides a viable description for the long-time behavior of polymer dynamics in a melt, provided entanglements with other chains, giving rise to reptation dynamics,^{7,9} are not important

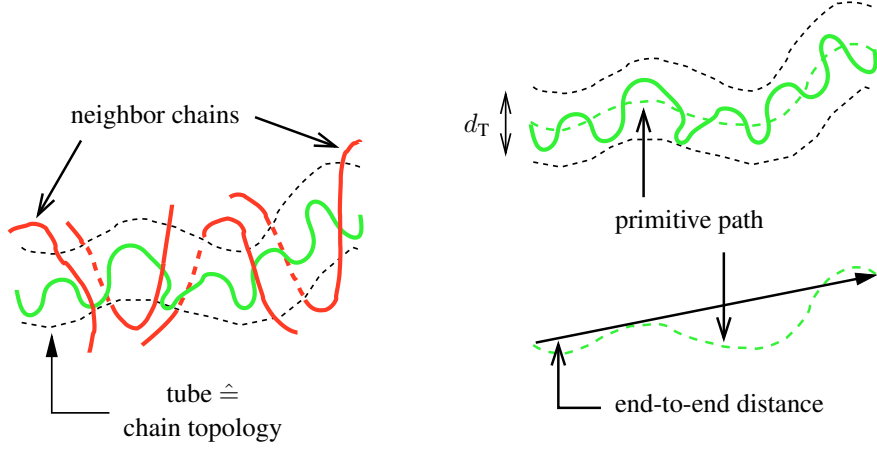


Figure 5. Sketch of the reptation concept for the dynamics of long-chain polymer melts.^{7,9} The chain is supposed to be enclosed in a “tube” formed by its neighbors. The tube may be characterized by an axis, the primitive path. The tube confines the motion of the enclosed chain: It predominantly moves along the primitive path. Perpendicular excursions are suppressed beyond the tube diameter d_T . The tube diameter is larger than the effective bond length b : $d_T = b\sqrt{N_e}$, where the “entanglement length” $N_e \gg 1$. The primitive path represents the shortest connection between the chain ends, which respects the topology imposed on the enclosed chain by the entanglements with its neighbors. The length L of the primitive path is thus larger than R_e , which is the shortest connection between the chain ends in space. L varies linearly with N : $Ld_T = R_e^2$ so that $L = d_T(N/N_e)$.

(see Fig. 5 and also below).^j

The Rouse theory assumes the chains to be ideal and models them as a sequence of Brownian beads, connected by harmonic springs and subjected to a local random force and a local friction.^{6,7} This bead-spring model is characterized by two parameters: the effective bond length b and the monomer mobility m . The mobility, or more precisely $1/m$, measures the time it takes a bead to diffuse over the distance b . Thus, the diffusion coefficient of a monomer is proportional to mb^2 . As the center of mass (CM) of a chain does not experience any external force other than the antagonistic friction and random forces, the theory predicts that the CM diffuses freely at all times

$$g_3(t) = \left\langle [\mathbf{R}_{\text{cm}}(t) - \mathbf{R}_{\text{cm}}(0)]^2 \right\rangle = 6D_N t, \quad (12)$$

where $\mathbf{R}_{\text{cm}}(t)$ denotes the position of the CM at time t . The diffusion coefficient of a chain is by a factor of N slower than that of a monomer, i.e.,

$$D_N \sim \frac{mb^2}{N}. \quad (13)$$

From Eqs. (12,13) the longest relaxation time τ_N can be obtained. Arguing that a chain is relaxed when its CM has diffused over a distance of the order of its own size, we find

$$g_3(\tau_N) \sim D_N \tau_N \stackrel{!}{\sim} R_g^2 \sim b^2 N \quad \Rightarrow \quad \tau_N \sim \frac{N^2}{m}, \quad (14)$$

^jTo our knowledge, there is no established derivation of the Rouse model from a microscopic theory. For a recent attempt see Ref. 28.

where the ideality of the chain was exploited.

Strongly Entangled Polymers and Reptation Model. The single-chain picture proposed by the Rouse theory is supposed to be valid as long as entanglements with other chains do not dominate the polymer dynamics. This is believed to be the case for short chains, for which N is smaller than the entanglement length N_e (Fig. 5). For $N \gg N_e$ the prevailing picture is that a chain is enclosed in a temporary “tube” formed by its neighbors. Entanglements force the enclosed chain to diffuse along the contour of the tube having a length of $L \sim N$ (“reptation”; see Fig. 5).^{6,7} Because the curvilinear diffusion through the tube is presumed to be Rouse-like, reptation theory predicts the relaxation time of the chain to scale with N as

$$\tau_N \sim \frac{L^2}{(mb^2/N)} \sim N^3 \quad (15)$$

so that the diffusion coefficient of the CM in space is given by [Eq. (14)]

$$g_3(\tau_N) \sim D_N \tau_N \stackrel{!}{\sim} R_g^2 \sim N \quad \Rightarrow \quad D_N \sim \frac{1}{N^2}. \quad (16)$$

Experimentally, one finds a still stronger dependence: $\tau_N \sim N^{\approx 3.4}$ and $D_N \sim N^{\approx -2.4}$.^k

Clearly, simulation methods which attempt to model the true physical dynamics, such as Molecular Dynamics or Monte Carlo algorithms employing local random moves, must suffer from these long relaxation times. Various alternative Monte Carlo methods have been proposed to efficiently equilibrate dense polymer melts. We will present some of these approaches (Sec. 5 and Sec. 6).

3 Monte Carlo Methods: A Brief Overview

In equilibrium statistical mechanics thermodynamic properties are calculated as ensemble averages over all points $\mathbf{x}$ in a high-dimensional configuration space Γ .^l In the canonical ensemble the average of an observable $A(\mathbf{x})$ is given by

$$\langle A \rangle = \int d\mathbf{x} A(\mathbf{x}) P_{\text{eq}}(\mathbf{x}) = \frac{1}{\mathcal{Z}} \int d\mathbf{x} A(\mathbf{x}) \exp[-\beta U(\mathbf{x})]. \quad (17)$$

In general, the integral cannot be solved analytically. Monte Carlo (MC) simulations provide a numerical approach to this problem by generating a random sample of configuration-space points $\mathbf{x}_1, \dots, \mathbf{x}_m, \dots, \mathbf{x}_M$ according to some distribution $P_s(\mathbf{x})$. $\langle A \rangle$ is then estimated by^{31–33}

$$\bar{A} = \frac{\sum_{m=1}^M A(\mathbf{x}_m) e^{-\beta U(\mathbf{x}_m)} / P_s(\mathbf{x}_m)}{\sum_{m=1}^M e^{-\beta U(\mathbf{x}_m)} / P_s(\mathbf{x}_m)} = \frac{\sum_{m=1}^M A(\mathbf{x}_m) W(\mathbf{x}_m)}{\sum_{m=1}^M W(\mathbf{x}_m)}, \quad (18)$$

^kThis exponent varies very little –if at all– with the chemical properties of the (linear) polymer.^{29,30}

^lWe assume that the momenta can be integrated out, since the observables only depend on the positions of the particles.

where we introduced the “weight” $W(\mathbf{x}) = P_{\text{eq}}(\mathbf{x})/P_s(\mathbf{x})$. Note that, while $\langle A \rangle$ is a number, $\bar{A}$ is still a random variable. Whether $\bar{A}$ represents a good estimate for $\langle A \rangle$ depends on the total number M of configurations used and, for a given M , on the choice of $P_s(\mathbf{x})$.

To see this in more detail^m let us define the mean value with respect to P_s by

$$\langle (\cdot) \rangle_s = \int d\mathbf{x} (\cdot) P_s(\mathbf{x}) . \quad (19)$$

For large M , the average of $\bar{A}$ and its variance $\text{var}_s(\bar{A})$ may be estimated from the small-fluctuation approximations³⁴

$$\left\langle \frac{Y}{Z} \right\rangle_s \approx \frac{\langle Y \rangle_s}{\langle Z \rangle_s} \left[1 - \frac{\langle YZ \rangle_s}{\langle Y \rangle_s \langle Z \rangle_s} + \frac{\langle Z^2 \rangle_s}{\langle Z \rangle_s^2} \right] , \quad (20)$$

$$\text{var}_s \left(\frac{Y}{Z} \right) \approx \frac{1}{\langle Z \rangle_s^2} \left[\langle Y^2 \rangle_s - 2 \frac{\langle Y \rangle_s \langle YZ \rangle_s}{\langle Z \rangle_s} + \frac{\langle Y \rangle_s^2 \langle Z^2 \rangle_s}{\langle Z \rangle_s^2} \right] . \quad (21)$$

This givesⁿ

$$\langle \bar{A} \rangle_s \approx \langle A \rangle - \frac{1}{M} \left[\langle WA \rangle - \langle W \rangle \langle A \rangle \right] , \quad \text{var}_s(\bar{A}) \approx \frac{1}{M} \left\langle W(A - \langle A \rangle)^2 \right\rangle . \quad (22)$$

Equation (22) shows that $\bar{A}$ provides an unbiased estimate of $\langle A \rangle$ in the limit $M \gg 1$ unless $\langle W \rangle \gg 1$, i.e., unless $P_s(\mathbf{x})$ is very different from $P_{\text{eq}}(\mathbf{x})$. When $P_s(\mathbf{x})$ deviates considerably from $P_{\text{eq}}(\mathbf{x})$, it predominantly samples configuration-space points, which are not representative of the thermal equilibrium. One could try to compensate this inefficient sampling by making M larger and larger. However, on the one hand this renders the simulation very time-consuming. On the other hand, there is no guarantee that the maximum M one is willing (or able) to simulate suffices to outweigh the error incurred by the inadequate choice of $P_s(\mathbf{x})$.

Thus, $P_s(\mathbf{x})$ should approximate $P_{\text{eq}}(\mathbf{x})$ as closely as possible to obtain meaningful results from MC simulations. To this end, two approaches may be distinguished:^{34,35}

1. *Static MC methods*: Static methods generate a sequence of *statistically independent* configuration-space points from the distribution $P_s(\mathbf{x})$. In this case one has to tune the algorithm cleverly so that the weights $W(\mathbf{x})$ do not get out of hand. Examples how to achieve this will be discussed in Sec. 6.
2. *Dynamic MC methods*: Dynamic methods generate a sequence of *correlated* configuration-space points via some stochastic process which has $P_{\text{eq}}(\mathbf{x})$ as its unique equilibrium distribution. In practice, this process is always taken to be a Markov process.^{32,33} The defining property of a Markov process is that it has no “memory”. That is, the probability for the occurrence of the future configuration $\mathbf{x}$ depends only on the present configuration $\mathbf{x}'$ and not on the other configurations that the process visited in the past.

Dynamic MC methods have become a widely used simulation technique, to which we will also heavily refer in the following sections. So, we provide a brief introduction here (many more details may be found in Ref. 36).

^mIn part, our discussion closely follows Sec. 2.3 of Ref. 34.

ⁿNote that $\langle W \rangle_s = 1$, $\langle WA \rangle_s = \langle A \rangle$, etc. In Eq. (22) the $\approx$ -sign means that there are corrections of $O(1/M^2)$ which we have neglected.

Let us assume that the configuration space is discrete and that the Markov process evolves in this space in discrete time steps $\Delta t (= 1)$. The time evolution of this Markov chain may be characterized by the “master equation” for the probability $P(\mathbf{x}, t)$ to find the system in the state $\mathbf{x}$ at time t

$$P(\mathbf{x}, t+1) - P(\mathbf{x}, t) = \sum_{\mathbf{x}' \neq \mathbf{x}} \left[w(\mathbf{x}|\mathbf{x}')P(\mathbf{x}', t) - w(\mathbf{x}'|\mathbf{x})P(\mathbf{x}, t) \right]. \quad (23)$$

Here, $w(\mathbf{x}|\mathbf{x}')$ denotes the transition probability from $\mathbf{x}'$ to $\mathbf{x}$ which is independent of time. (In the continuous time limit ($\Delta t \rightarrow 0$) it becomes a “transition rate”, i.e., a transition probability per unit time.) Equation (23) expresses the balance between the flux of all other states $\mathbf{x}'$ towards $\mathbf{x}$ (first term of the rhs), leading to an increase of $P(\mathbf{x})$, and the flux away from $\mathbf{x}$ (second term of the rhs) which diminishes $P(\mathbf{x})$. Note that only terms with $\mathbf{x} \neq \mathbf{x}'$ contribute. We can rewrite Eq. (23) by including the missing term for $\mathbf{x} = \mathbf{x}'$ if we take into account that $w(\mathbf{x}|\mathbf{x}')$ is normalized. Since a transition from $\mathbf{x}'$ to some state $\mathbf{x}$, including $\mathbf{x}'$, will occur with certainty, $w(\mathbf{x}|\mathbf{x}')$ satisfies

$$\sum_{\mathbf{x}'} w(\mathbf{x}|\mathbf{x}') = 1. \quad (24)$$

Inserting Eq. (24) into Eq. (23) the master equation takes the following form

$$P(\mathbf{x}, t+1) = \sum_{\mathbf{x}'} w(\mathbf{x}|\mathbf{x}')P(\mathbf{x}', t). \quad (25)$$

For the application of these results to statistical physics it is necessary that $P(\mathbf{x}, t)$ converges to a unique stationary distribution, irrespective of the initial configuration of the system, in the long-time limit and that this distribution is the (canonical) equilibrium distribution $P_{\text{eq}}(\mathbf{x})$. Thus, the right-hand side of Eq. (23) must vanish for $P(\mathbf{x}', t) = P_{\text{eq}}(\mathbf{x}')$. Certainly, this is the case if each term of the sum vanishes separately. This leads to the condition of “detailed balance” (see Refs. 31–33, 36)

$$w(\mathbf{x}|\mathbf{x}')P_{\text{eq}}(\mathbf{x}') = w(\mathbf{x}'|\mathbf{x})P_{\text{eq}}(\mathbf{x}). \quad (26)$$

To exploit this condition in MC algorithms the transition probability may be split into two independent parts: First, we propose a transition from $\mathbf{x}'$ to $\mathbf{x}$ according to some probability $P_{\text{pro}}(\mathbf{x}' \rightarrow \mathbf{x})$. Then, this move will be accepted or rejected with probabilities $\text{acc}(\mathbf{x}' \rightarrow \mathbf{x})$ and $1 - \text{acc}(\mathbf{x}' \rightarrow \mathbf{x})$, respectively. So, we have

$$\frac{w(\mathbf{x}|\mathbf{x}')}{w(\mathbf{x}'|\mathbf{x})} = \frac{P_{\text{pro}}(\mathbf{x}' \rightarrow \mathbf{x}) \text{acc}(\mathbf{x}' \rightarrow \mathbf{x})}{P_{\text{pro}}(\mathbf{x} \rightarrow \mathbf{x}') \text{acc}(\mathbf{x} \rightarrow \mathbf{x}')} = e^{-\beta[U(\mathbf{x}) - U(\mathbf{x}')]}. \quad (27)$$

To solve this equation for $\text{acc}(\mathbf{x}' \rightarrow \mathbf{x})$ we set

$$\text{acc}(\mathbf{x}' \rightarrow \mathbf{x}) = F \left(\frac{P_{\text{pro}}(\mathbf{x} \rightarrow \mathbf{x}') e^{-\beta U(\mathbf{x})}}{P_{\text{pro}}(\mathbf{x}' \rightarrow \mathbf{x}) e^{-\beta U(\mathbf{x}')}} \right). \quad (28)$$

From Eq. (27) we see that the function $F(x)$ satisfies $F(x)/F(1/x) = x$. One solution to this equation was proposed by Metropolis *et al.*:³⁷ $F(x) = \min(1, x)$. This leads to the “Metropolis criterion” for the acceptance probability

$$\text{acc}(\mathbf{x}' \rightarrow \mathbf{x}) = \min \left(1, \frac{P_{\text{pro}}(\mathbf{x} \rightarrow \mathbf{x}')}{P_{\text{pro}}(\mathbf{x}' \rightarrow \mathbf{x})} e^{-\beta[U(\mathbf{x}) - U(\mathbf{x}')] } \right). \quad (29)$$

The Metropolis criterion is the core of essentially all dynamic MC algorithms. It embodies detailed balance which guarantees that the simulation, irrespective of the initial configuration, converges to the canonical equilibrium distribution, provided the set of chosen Monte Carlo moves leads to ergodic sampling.^o

Detailed Balance versus Stationarity. Detailed balance is an important, but very strict criterion. Less stringent is the condition of stationarity [Eq. (25)]

$$P_{\text{eq}}(\mathbf{x}) = \sum_{\mathbf{x}'} w(\mathbf{x}|\mathbf{x}') P_{\text{eq}}(\mathbf{x}') , \quad (30)$$

implying that $P_{\text{eq}}(\mathbf{x})$ remains invariant under the Markov dynamics. Stationarity in conjunction with the ergodicity of chosen set of MC moves ensures a valid simulation.^{38,40}

In practice, this milder condition may be important. Imagine that we want to update a polymer chain consisting of N monomers and that each monomer can be displaced in N_{dis} directions. One possibility is to select a monomer and a direction randomly. Thus, $P_{\text{pro}}(\mathbf{x}' \rightarrow \mathbf{x}) = 1/(NN_{\text{dis}}) = P_{\text{pro}}(\mathbf{x} \rightarrow \mathbf{x}')$. This procedure obeys detailed balance: In the next move the same monomer and the reverse displacement may be chosen with the same a priori probability. On the other hand, one could also attempt to move one monomer after the other, proceeding regularly from monomer 1 to monomer N . This sequential updating scheme violates detailed balance: The next step never selects again the monomer whose displacement has just been attempted. So, the probability for the reverse move is zero.

However, sequential updating is a valid scheme if the individual steps obey detailed balance⁴⁰ or at least stationarity. To see that we can write the transition probability from $\mathbf{x}'$ to $\mathbf{x}$ for sequential updating as

$$w(\mathbf{x}|\mathbf{x}') = \sum_{\mathbf{z}_N} \cdots \sum_{\mathbf{z}_2} \sum_{\mathbf{z}_1} w^{(N)}(\mathbf{x}|\mathbf{z}_N) \cdots w^{(2)}(\mathbf{z}_2|\mathbf{z}_1) w^{(1)}(\mathbf{z}_1|\mathbf{x}') . \quad (31)$$

This means that the process passes sequentially first with probability $w^{(1)}(\mathbf{z}_1|\mathbf{x}')$ from $\mathbf{x}'$ to $\mathbf{z}_1$ by attempting to move the first monomer, then from $\mathbf{z}_1$ to $\mathbf{z}_2$ by attempting to move the second monomer, and so on until configuration $\mathbf{x}$ is reached. Multiplying Eq. (31) by $P_{\text{eq}}(\mathbf{x}')$ and summing over all $\mathbf{x}'$ we find

$$\begin{aligned} & \sum_{\mathbf{x}'} w(\mathbf{x}|\mathbf{x}') P_{\text{eq}}(\mathbf{x}') \\ &= \sum_{\mathbf{z}_N} \cdots \sum_{\mathbf{z}_2} \sum_{\mathbf{z}_1} w^{(N)}(\mathbf{x}|\mathbf{z}_N) \cdots w^{(2)}(\mathbf{z}_2|\mathbf{z}_1) \underbrace{\sum_{\mathbf{x}'} w^{(1)}(\mathbf{z}_1|\mathbf{x}') P_{\text{eq}}(\mathbf{x}')}_{= P_{\text{eq}}(\mathbf{z}_1)} \\ &= \dots = P_{\text{eq}}(\mathbf{x}) , \end{aligned} \quad (32)$$

^oBy “ergodic” sampling we mean that the probability of finding the system in configuration $\mathbf{x}$, starting from some state $\mathbf{x}'$ (including $\mathbf{x}$), is non-zero for all $\mathbf{x}$ after a sufficiently long time.³⁸ This definition is a bit dangerous because it conflicts with others in the literature. For instance, in mathematical texts on Markov chains (= discrete-time Markov processes with a discrete configuration space) our definition rather corresponds to an “irreducible and aperiodic” chain (there, “ergodic” is a synonym for “irreducible”).^{34,35,39} In Ref. 40 our definition would be termed “regular sampling”.

i.e., sequential sampling preserves the stationarity of the equilibrium distribution. Thus, it is a correct simulation procedure. This conclusion is important for a variety of MC methods which perform different trial moves in a fixed order.

4 Some Coarse-Grained Simulation Models

In Sec. 1 we introduced the term “coarse-grained model”. This was defined as a model which associates a group of chemical monomers with a “bead” (effective monomer) in order to eliminate microscopic degrees of freedom (bond length vibrations, etc.). Here, we refine our definition and distinguish between two types of coarse-grained models:

1. *The coarse-grained model is derived from a specific polymer.* In practice, this usually implies that the properties of the model (potential parameters, density, etc.) have to be adjusted to results from atomistic simulations of the polymer under consideration (see Appendix 7 for an example). The incentive to devise such models rests upon the fact that they may be simulated much more efficiently than their atomistic counterpart. Thus, it is tempting to split the simulation into two levels: First, one uses the coarse-grained model for equilibration and for the determination of large-scale properties. Then, atomistic details may be reinserted to allow for a thorough comparison with experiments. Recent attempts to perform such multi-scale approaches are described in Refs. 41, 42 (see also Appendix 7).
2. *The coarse-grained model has no direct connection to any specific polymer.* It is a generic model retaining only features common to all polymers of the same chain topology. For (uncharged) linear polymers these features are chain connectivity, excluded-volume interactions, and, additionally, monomer-monomer attractions if one wants to simulate Θ - or bad-solvent conditions (see Fig. 3). Many of these generic models, be it lattice or continuum models, have been introduced in the literature (see Refs. 43, 44 for comprehensive overviews). In the following we present those models in more detail, which will be discussed in Secs. 5, 6.

4.1 Lattice Models

The Self-Avoiding Walk. About 50 years ago Orr and Montroll⁴⁶ proposed the self-avoiding walk (SAW) as a model for a linear polymer in a good solvent. The SAW is defined on a discrete lattice, often on a square or simple cubic lattice (Fig. 6). Each monomer occupies one lattice site, the bond length equals the lattice constant, and the bond angles are restricted by the lattice geometry and by the repulsive hard-core monomer-monomer interaction (e.g. 90° and 180° for the cubic lattice, as immediate backfolding is forbidden). This model can be complemented by attractive interactions if, for instance, an energy gain $-\epsilon$ is associated with every occupied nearest neighbor pair.⁴⁷ In addition to excluded volume interactions the simulation then also has to take account of the Boltzmann factor $\exp(n_{\text{nn}}\epsilon/k_{\text{B}}T)$, where n_{nn} is the number of nearest neighbors.

To simulate the SAW by dynamic Monte Carlo one must first decide about the elementary moves that propose a new SAW configuration $\mathbf{x}$ from an old one $\mathbf{x}'$. The earliest suggestion⁴⁸ comprised one-bead excitations^{34, 43, 44} (Fig. 6). In these algorithms, one

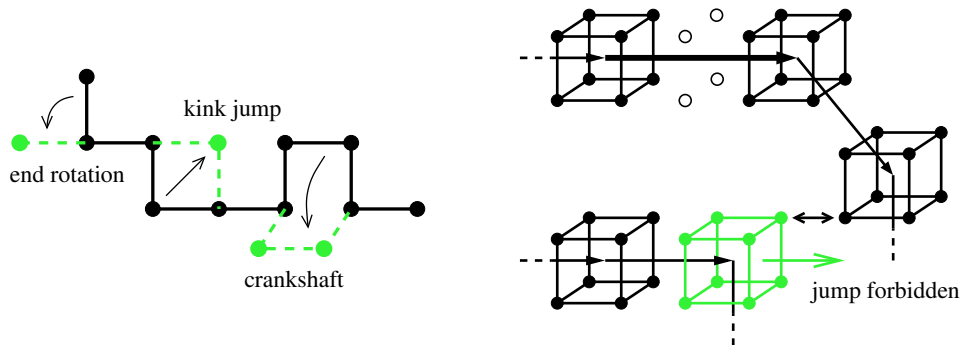


Figure 6. *Left figure*: Single-site self-avoiding walk (SAW) of chain length $N = 10$ on a simple cubic lattice (solid lines and black dots). The grey dots and the grey dashed lines indicate the moves discussed in the text: end-bond rotation, kink jumps and 90° crankshaft motion. *Right figure*: Sketch of a possible configuration of monomers in the 3D bond-fluctuation model (BFM). (A vectorized version of the BFM algorithm can be found in Ref. 45.) The bond vector $(3, 0, 0)$ (thick black arrow) blocks four lattice sites (marked by $\circ$) that are no longer available to other monomers due to the excluded volume interaction. This interaction also prevents the jump of the grey monomer in the direction of the large arrow ($\longrightarrow$), since the corners of the monomers, indicated by $\leftrightarrow$, would then occupy the same lattice site.

chooses a monomer at random. If the monomer is at the chain end, the bond to its neighbor is turned to a randomly selected lattice direction. Due to the fixed bond length an inner monomer is only mobile if its bond angle is 90° on the square or simple cubic lattice. In this case, one attempts a “kink-jump” motion, i.e., a one-bead flip to the opposite lattice site. End-bond rotation and kink jumps are accepted according to the Metropolis criterion if the target sites are empty.

These moves are special examples of the class of “local N -conserving moves”.³⁴ Quite generally, a “local move” alters the configuration of a small piece of the original SAW while leaving the remaining monomers unchanged. This definition opens the possibility to invent moves comprising more than one bead, such as two-bead or three-bead excitations. Figure 6 shows a common example, the 90° crankshaft motion (only possible in 3D). The crankshaft motion removes an important drawback of kink jumps. It introduces new bond vectors, whereas a kink jump does not. Therefore, if only end-bond rotations and kink jumps are allowed, new bond orientations have to diffuse from the ends toward the interior part of the chains. This algorithm is not very efficient in reshuffling the bond vectors and so in preparing independent configurations. The inclusion of crankshaft motions remedies this problem.

However, even then a disturbing feature remains. It has been proved that all local N -conserving algorithms for two- and three-dimensional SAW’s are not ergodic for large N .⁴⁹ There are dense configurations (“double cul-de-sac” in 2D, “knots” in 3D; see Ref. 34) which are completely frozen: They can neither be transformed into nor reached from other configurations. Whether this problem is serious in practice is a question that, to our knowledge, is not fully settled (see e.g. Ref. 43 or footnote 9 of Ref. 50). One can argue that, if one starts from an extended configuration –for instance, from a straight rod– and if one is interested in high- T properties only, non-ergodicity effects due to compact structures should be small. This argument may be true for short chains,^p but should fail for long

^pHere, it is not clear what “short” really means. For $N \lesssim 10^2$ the error incurred by using local N -conserving

ones, since it has been proved that the fraction of SAW's belonging to the ergodicity class^q of the straight rod is exponentially small in the large- N limit. Of course, if one is interested in low- T properties, problems with non-ergodicity might be sizable, even for small chain length.⁵¹

The Bond-Fluctuation Model. The bond-fluctuation model (BFM) was proposed^{52,53} as an alternative to a (single-site) SAW model, which retains the computational efficiency of the lattice without being plagued by severe ergodicity problems. The key idea is to increase the size of a monomer which now occupies, instead of a single site, a whole unit cell of the lattice (e.g. a square for the 2D- or a cube for the 3D hyper-cubic lattice; see Fig. 6). This enlarged monomer size has two important consequences:

1. A priori, many different bond vectors can occur. This multitude is restricted by two conditions. First, adjacent monomers may not overlap. This limits the bond length to $\ell \geq \ell_{\min} = 2$ (in units of the lattice constant). Second, the hard-core monomer-monomer interaction should suffice to prevent two bonds from intersecting each other in the course of the simulation. In 2D this only imposes an upper bound on the bond length, $\ell \leq \ell_{\max} = \sqrt{13}$,^{52,53} whereas in 3D, in addition to $\ell \leq \ell_{\max} = \sqrt{10}$, some smaller bond vectors also have to be excluded.⁵⁴ The resulting sets of allowed bond vectors are:

$$\begin{aligned} \{\mathbf{b}\} &= [2, 0], [2, 1], [2, 2], [3, 0], [3, 1], [3, 2] \quad (2\text{D}) , \\ \{\mathbf{b}\} &= [2, 0, 0], [2, 1, 0], [2, 1, 1], [2, 2, 1], [3, 0, 0], [3, 1, 0] \quad (3\text{D}) , \end{aligned} \quad (33)$$

where $[\cdot]$ denotes a class of bond vectors sharing the same length, but differing in direction. For instance, the class $[2, 0]$ ($[2, 0, 0]$) comprises all vectors with a length of 2 and direction along the lattice axis (4 directions in 2D, 6 in 3D). Equation (33) gives rise to 41 bond angles in 2D⁵⁵ and to 87 bond angles in 3D.⁵⁴ This has to be compared to 3 (2D) or 5 (3D) bond angles for the SAW model on the hypercubic lattice where a monomer is associated with a lattice site. Due to the multitude of different bond lengths and bond angles the BFM is much closer to continuous-space behavior than the single-site lattice model^r.⁵⁷

2. Ergodicity problems are much less severe than for the single-site SAW. For the BFM a local N -conserving move consists of selecting a monomer at random and of attempting a displacement by one lattice constant in a randomly chosen lattice direction. As these local jumps^s permit transitions between different vectors, the algorithm can escape from configurations where a single-site model would be frozen in.⁵² If the attempted displacement satisfies both the bond vectors constraints of Eq. (33) and the excluded volume interaction, the move is accepted. Of course, it is also possible

algorithms seems to be small (see Ref. 43 and the footnote 9 of Ref. 50).

^qBy "ergodicity class of a straight rod" we mean all mutually accessible configurations, one of which is the rod.
^rThe main advantage of lattice models is their computational efficiency. Longer length and times scales may be probed. However later on, the results of the simulation shall be compared to theories or experiences, which "live" in continuous space. So, the important question arises of how well the lattice algorithm approximates continuum properties. A general, intuitive answer is: The finer the lattice, i.e., the more sites are occupied by one particle, the closer the continuum limit should be realized. Recently, this statement was made more precise by the example of monatomic fluids interacting via a Lennard-Jones or a Buckingham potential.⁵⁶

^sLarger jumps distances were also tested (in 3D), but found less efficient in concentrated solutions.⁵⁴

to include a finite interaction energy. Then, the move is accepted according to the Metropolis criterion. A possible choice^t is to work with an energy $-\epsilon$ between pairs of monomers with distance $2 \leq r \leq \sqrt{6}$. This interval comprises all neighbors which contribute to the first peak of the pair-distribution function⁵⁹ in a dense polymer system.⁵⁴ This choice was made in studies of the Θ -point⁶⁰ and of the phase transition in binary polymer blends (see Ref. 11).

4.2 Continuum Models

Two Bead-Spring Models. A widely used continuum model is the bead-spring model introduced by Grest and Kremer.⁶¹ In this model nearest-neighbor monomers along the backbone of the chain are bonded to each other by a FENE (finitely extendible non-linear elastic) potential

$$U_F(\ell) = \begin{cases} -\frac{1}{2}k\ell_{\max}^2 \ln [1 - (\ell/\ell_{\max})^2] & \ell \leq \ell_{\max} , \\ \infty & \text{else} , \end{cases} \quad (34)$$

whereas all monomers, bonded and non-bonded ones, interact via a truncated and shifted Lennard-Jones (LJ) potential

$$U_{\text{LJ}}^{\text{ts}}(r) = \begin{cases} 4\epsilon [(\sigma/r)^{12} - (\sigma/r)^6] + C(r_{\text{cut}}) & \text{for } r \leq r_{\text{cut}} , \\ 0 & \text{else} , \end{cases} \quad (35)$$

where $C(r_{\text{cut}})$ ensures that the potential vanishes at the cut-off parameter r_{cut} . Such a cut-off is commonly employed to render the interaction short-ranged (Fig. 7).^u The parameter ϵ defines the energy scale and σ the length scale of the system. That is, we set $\epsilon = \sigma = 1$ (LJ units) in the following.

For small values of the bond length the FENE potential is harmonic (“elastic behavior”), i.e., $U_F(\ell) = k\ell^2/2$ for $0 \leq \ell \ll \ell_{\max}$, whereas the logarithmic divergence imposes $\ell < \ell_{\max}$ (“finite extensibility”). The parameters $\ell_{\max}$ and k have to be chosen such that the possibility of bond crossing becomes so unlikely that it never occurs. Reference 61 suggests $k = 30$ and $\ell_{\max} = 1.5$ (in LJ units). This has become a standard choice.

The FENE potential alone does not prevent monomers from overlapping. To realize excluded volume the LJ-interaction has to be taken into account also between bonded monomers. The superposition of the FENE- and the LJ-potentials yields a steep effective bond potential with a minimum at $\ell_0 \simeq 0.96$ (Fig. 7). The shape of the bond potential depends on the cut-off parameter of the LJ-interaction:

- If one takes $r_{\text{cut}} = r_{\text{min}} = 2^{1/6}$ ($C(r_{\text{cut}}) = 1$), i.e., as the minimum of the LJ-potential, the monomer-monomer interaction becomes purely repulsive. This model is commonly called “Kremer-Grest model”.⁶¹ For isolated chains it realizes good solvent conditions.

^tAnother choice uses a discretization of the Lennard-Jones potential.⁵⁸

^uFrom a computational point of view short-range interactions are convenient because the simulation can be speeded up by neighbor lists.^{31,62} However, as the truncation ignores the contribution of the tail of the potential, the error incurred must be corrected before comparing with results for the full potential. For instance, the truncation shifts the location of the critical point of the liquid-gas transition in a LJ-liquid (see Ref. 63 or Sec. 3.2.2 of Ref. 31 for details). To avoid these truncation effects some authors prefer to work with the full LJ-potential.⁶⁴

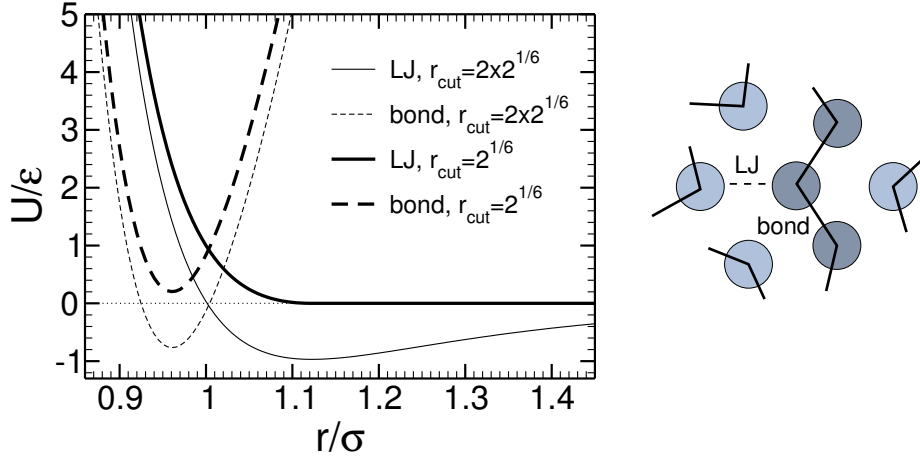


Figure 7. Bond and Lennard-Jones potentials versus the distance r between two monomers of the bead-spring model (for the bond potential $\mathbf{r} = \mathbf{b}$). The bond potential results from the superposition of Eqs. (34,35). For both cut-off parameters r_{cut} the bond potential was shifted by -20 to show it on the same scale as the LJ-potentials. The LJ-potential with $r_{\text{cut}} = 2^{1/6}$ is purely repulsive, whereas the potential with $r_{\text{cut}} = 2 \times 2^{1/6}$ has an attractive minimum at $r_{\text{min}} = 2^{1/6}$.

- The simulation of Θ - or bad solvents requires to incorporate part of the LJ-attraction by increasing r_{cut} . Obviously, there is freedom where to cut off the attractive part. One possibility is $r_{\text{cut}} = 2 \times 2^{1/6}$ ($C(r_{\text{cut}}) = 127/4096$).⁶⁵ This choice is a compromise between the wish to include the major part of the attractive interaction and the need to keep the potential short-ranged. The resulting phase diagram was studied in Ref. 66.

(Yet) Another Bead-Spring Model. If we recall the idea of the coarse-graining –a coarse-grained monomer stands for a group of chemical monomers– it appears plausible that coarse-grained monomers are softer than their chemical counterparts. Thus, an exponent smaller than 12 in Eq. (35) may be better suited to represent their repulsion. In fact, such an observation was made in a recent effort to develop a coarse-grained model for poly(vinyl alcohol) (see Ref. 67 and Appendix 7). This study also suggests the following generic model which may be considered as a variant of the Kremer-Grest model.

In this (“Kremer-Grest-like”) model non-bonded monomers interact via a purely repulsive 9-6 LJ-potential,

$$U_{9-6}^{\text{rep}}(r) = \begin{cases} \epsilon_0 [(\sigma_0/r)^9 - (\sigma_0/r)^6] + C(r_{\text{min}}) & \text{for } r \leq r_{\text{min}} = (3/2)^{1/3} \sigma_0, \\ 0 & \text{else,} \end{cases} \quad (36)$$

where $\epsilon_0 = 1.511$ and $C(r_{\text{min}}) = 4\epsilon_0/27$. These non-bonded interactions are excluded between nearest neighbors in the chain, which are connected to each other by a harmonic potential

$$U_{\text{bond}}(\ell) = \frac{1}{2}k(\ell - \ell_0)^2 \quad (k = 2141.84 \sigma_0^{-2}, \quad \ell_0 = 0.97 \sigma_0). \quad (37)$$

The equilibrium bond length ℓ_0 agrees with that of the Kremer-Grest model. The spring constant k has to be chosen so large to inhibit bond crossings (see Ref. 68 for further

discussion). A similar bond potential, in conjunction with Eq. (35) and $r_{\text{cut}} = 2^{1/6}$, has recently been used to study the effect of the bond length on the structure and dynamics of polymer melts.⁶⁹

Local Moves for Continuum Models. The continuum models are constructed for use in Molecular Dynamics simulations. However, simulation within Monte Carlo schemes is also possible. Similarly to the lattice models a local updating scheme can be realized by selecting a monomer and a direction at random and by attempting a displacement in the chosen direction. This proposition is again accepted according to the Metropolis criterion.

The size Δ of the displacement is a tunable parameter. It should neither be too small nor too large. If Δ is too small, many moves may be accepted, but the system advances only slowly in configuration space. Many displacements are thus needed to obtain well decorrelated configurations. On the other hand, if Δ is too large, many moves will be rejected and the decorrelation is also slow. A scheme how to optimize the choice for Δ is explained in Sec. 3.3 of Ref. 31.

5 Monte Carlo Methods for Polymers: From Local to Non-Local Moves

The method of importance sampling is based on a Markov process in configuration space. A priori, this stochastic dynamics is merely a numerical algorithm, aiming at an efficient sampling according to $P_{\text{eq}}(\mathbf{x})$. It need not correspond to the physical dynamics of the (polymer) system under consideration. An appealing consequence of this feature is the freedom to invent clever MC moves which decorrelate the configurations in the smallest (CPU) time possible. These non-physical moves serve to rapidly equilibrate the system and to produce statistically independent equilibrium configurations for the study of structural and thermodynamic properties. We will pursue this idea in Secs. 5.2, 5.3. In the following section we rather want to concentrate on local moves and the ensuing dynamic interpretation of the MC method.

5.1 Local Moves: Studying Dynamic Properties with Monte Carlo

By employing non-local moves we can explore the statics of the system, but information about its dynamic properties is lost. Of course, the equilibrated configurations could be used in a Molecular Dynamics (MD) simulation to analyze the dynamic properties. However, if one is not willing to do that, the question arises of under which conditions the MC dynamics can be realistic. The answer to this comprises two parts:

1. Certainly, one can only expect the MC dynamics to become reliable on length and time scales where the deterministic motion of the monomers has been damped by the interaction with the surrounding (other monomers and/or solvent). For instance, in a (classical) MD simulation the monomers move ballistically at early times, i.e., their displacement is proportional to t . This is a consequence of the underlying Newtonian dynamics in the limit of vanishing force. At short times the monomers behave as if they did not “feel” the bonding to their neighbors and the presence of other particles, that is, as if they were free particles. As time increases, the interaction with

the surrounding becomes important. The motion of the monomers is then a result of a multitude of individual collisions. This “averaging” over fast degrees of freedom gradually lends a stochastic character to the dynamics which ultimately becomes diffusive in the long-time limit.

2. The moves should be “physical”. Usually, this implies that they are local.^v Furthermore, the dynamics should not be dominated by the momenta which are absent in Monte Carlo. The latter condition is satisfied in dense melts, but not in dilute solution. In dilute solution the motion of distant monomers along the chain backbone are coupled via hydrodynamic interactions (see Ref. 14). and Sec. 2.3). Thus, we might expect that a local Monte Carlo algorithm reproduces Rouse dynamics where these long-range interactions are neglected.

This expectation can be verified by estimating the scaling of local MC algorithms with N . To this end, let us assume that the center of mass (CM) of an isolated chain, be it on a lattice or in the continuum, may be considered as a free Brownian particle. This is reasonable, since the CM does not experience any external force other than the random force of the heat bath (resulting from the compound effect of the random monomer hops and the acceptance criterion). So, it should diffuse freely [Eq. (12)]. The corresponding diffusion constant D_N depends on chain length. To estimate this dependence we can argue that the center of mass is displaced by $\sim b/N$, if one monomer moves over a distance of order b while the other monomers remain fixed. This elementary motion takes on average the time $1/m$ with m denoting again the (temperature, density, etc. dependent) mobility of the monomer. For the CM to diffuse over the distance b , N such random motions are needed. This takes the time $m \times N$, which we use as our time unit here.^w Utilizing Eq. (12) we then find $g_3(t=1) \sim (mN) \times (b/N)^2 \sim D_N$. So,

$$D_N \sim \frac{mb^2}{N^{\approx 1}}. \quad (38)$$

Inserting this result in Eq. (14) we obtain the relaxation time of a chain

$$\tau_N \sim \frac{N^{\approx 1+2\nu}}{m} \sim \begin{cases} N^{\approx 2} & (\text{ideal chain: } \nu = 0.5) , \\ N^{\approx 2.176} & (3\text{D excluded-volume chain: } \nu = 0.588) . \end{cases} \quad (39)$$

Equations (38,39) agree with the predictions of the Rouse theory [Eqs. (13,14)].

Monte Carlo Dynamics versus Molecular Dynamics: An Example. The previous arguments suggest that the MC dynamics, based on local moves, becomes realistic for time and length scales outside the microscopic regime (of a bond). We want to support this assertion by a comparison between MC and MD simulations.

Figure 8 shows the diffusion coefficient D_N of a chain versus chain length. D_N was derived from the long-time limit of Eq. (12) for both the BFM, simulated via MC, and the Kremer-Grest-like continuum model of Eqs. (36,37), simulated via MD. The figure

^vExamples for local moves of lattice models are given in Fig. 6. See the very end of Sec. 4.2 for a brief discussion of local moves in continuum models.

^wThis statement introduces the time unit τ_{mcs} of a “Monte Carlo step (MCS)”. A MCS is defined as the time it takes to give each of the N monomers the possibility to move once.^{32,33} Thus, we measure time in units of attempted elementary moves per monomer.

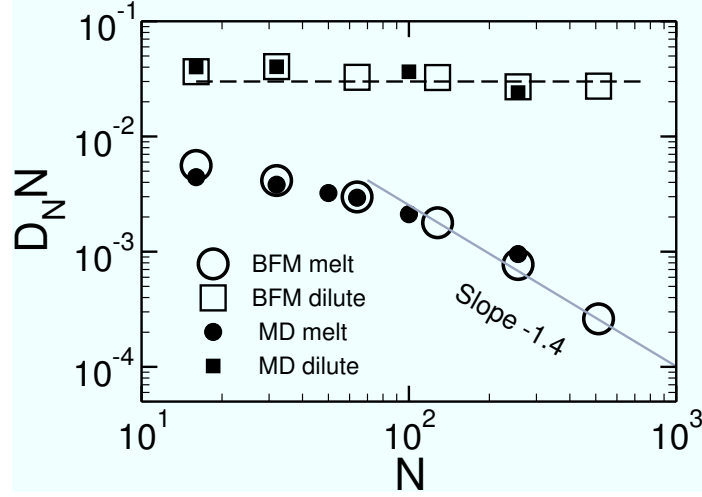


Figure 8. Diffusion coefficient D_N versus N . Two simulation methods are compared: The open symbols represent MC data of the (lattice) BFM, the filled symbols were obtained from MD simulations of a (continuum) Kremer-Grest-like model [Eqs. (36,37)]. For both models results obtained in 3D for a dilute solution and for a melt are shown. For the BFM this corresponds to the following volume fraction ϕ : $\phi = 0.0078$ (dilute), $\phi = 0.5$ (melt). For the Kremer-Grest-like model this corresponds to the following monomer densities ρ : $\rho = 0.0835$ (dilute), $\rho = 0.835$ (melt). Qualitatively, the MD simulations yield the same dependence of D_N on N . To illustrate this agreement the MD data were vertically shifted by an amount that optimizes the agreement with the MC results. (The shift factors are different for the dilute solution and the melt.) In dilute solution, we find Rouse behavior [Eq. (13)] for both methods. In the melt, the chains diffuse more slowly. The dependence of D_N on N is qualitatively compatible with the Rouse-to-reptation crossover when N passes the threshold N_e (Sec. 2.3). Quantitatively however, there are deviations. Particularly for large N , the decrease of D_N appears to be stronger than predicted by reptation theory [Eq. (16)]. Roughly, we find $D_N \sim N^{-2.4}$. Adapted from Ref. 70.

displays the results of the simulations for a dilute solution and a dense melt.^x Clearly, there is a high degree of accord between the results, illustrating that the BFM with local moves reproduces the realistic dynamics of the MD simulations. Thus, MC simulations can be more than just a versatile tool to calculate high-dimensional integrals. They may provide information on the dynamics of a system.^y

^xIn the BFM, density is commonly specified in terms of the volume fraction ϕ of lattice sites occupied by monomers. As a monomer comprises all sites of a unit cube, the monomer density ρ is smaller than ϕ by a factor 8, $\rho = \phi/8$. Although the value $\phi = 0.5$ appears small, the work by Paul *et al.*²⁷ established that the chains have melt-like properties at this density (see also Ref. 4). Since then, $\phi = 0.5$ has become a standard choice (in 3D). For the Kremer-Grest model, the work of Ref. 61 showed that a monomer density of $\rho = 0.85$, or a value close to this, is a good choice to realize melt-like behavior. We adopted this choice in our MD study. The MD simulations were done at constant volume and constant temperature (Langevin thermostat⁶²).

^yThere is further ample evidence for the correctness of this statement from other studies. For simple liquids of LJ-particles see e.g. Ref. 71. For polymers see the review in Ref. 8 or the comparison of MD simulations for polybutadiene and polyethylene with MC simulations of the BFM.⁷² Furthermore, Monte Carlo methods have been applied to simulate dynamic processes in such diverse fields as relaxation phenomena in spin and structural glasses, spinodal decomposition of mixtures, nucleation processes, diffusion-limited aggregation, etc. (see e.g. the textbooks of Binder and Heermann³² or Landau and Binder³³).

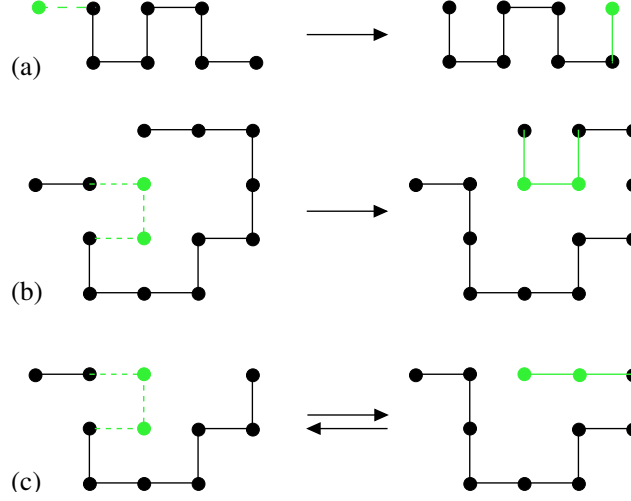


Figure 9. Slithering-snake move (a) and general reptation moves (b,c). Both moves are illustrated by the shrinkage-growth implementation. For the slithering-snake algorithm, a randomly chosen end bond (dashed line) is removed and then a new bond vector (also randomly chosen) is attached to the other chain end. For the general reptation algorithm, three moves are shown: Kink-kink transport (b) and kind-end/end-kink reptation (c). Kink-kink transport implies that a randomly chosen kink is shrunk to a bond and a new kink is inserted somewhere else along the chain. Kink-end reptation ($\rightarrow$) amounts to replacing a randomly chosen kink by a bond and to appending two new bond vectors (also randomly chosen) to the other chain end. End-kink reptation ($\leftarrow$) corresponds to the reverse “reaction”.

Relaxation Time and Computational Complexity. An important issue in any algorithm is its “computational complexity”. Quite generally, the computational complexity may be defined as the time required to solve a computational problem.⁷³ Here, the computational problem is to decorrelate chain configurations. According to Eq. (39) this takes a relaxation time $\tau_N \sim N^{1+2\nu}$ in units of the Monte Carlo step (MCS; see footnote on page 103). As a MCS comprises N attempted moves of a monomer, the computational complexity τ_{cc} scales with N as $\tau_{cc} = N\tau_N \sim N^{2+2\nu}$.

This rapid increase of τ_{cc} with chain length –called “critical slowing-down”³³– makes it difficult in practice to efficiently decorrelate configurations of long chains by local moves. In order to be able to simulate large chains with sufficient statistics, moves have to be implemented, which reduces ($\tau_{cc} \sim N^\alpha$ with $\alpha < 2 + 2\nu$) or even eliminate ($\tau_{cc} \sim N^0$) the critical slowing-down. These moves cannot be local, they have to act, in some way, on all monomers of the chain. In the following we want to discuss two examples of such global updates: bilocal moves and the pivot algorithm.

5.2 Bilocal Moves: The Slithering-Snake and the Extended Reptation Algorithms

A bilocal N -conserving move consists in altering the configuration of two small groups of consecutive monomers. The groups are usually far from one another along the backbone of the chain. Typical examples are the slithering-snake and the extended reptation algorithm:

- The *slithering-snake* (or *reptation*) algorithm removes a bond from one chain end,

adds a new one to the other end and shifts the inner monomers one bond up along the chain in direction of the new bond (Fig. 9a). As the positions of the inner monomers remain unchanged, the chains “slithers” along its contour during the MC move (whence the name of the algorithm).^z

- The *extended reptation* algorithm transports a kink or an end group via a slithering motion along the chain.^a Commonly utilized moves are: (1) “Kink-kink reptation”, which deletes a kink at some position along the chain and inserts a new one at another position (Fig. 9b). (2) “Kink-end reptation”, which removes a kink somewhere along the chain and adds two new bonds at one of the chain ends (Fig. 9c→). (3) “End-kink reptation”, the inverse of “kink-end reptation” (Fig. 9c←).

In the remainder of this section we will concentrate on the slithering-snake algorithm. Extended reptation is only discussed in comparison to the slithering-snake algorithm.

Implementation and Ergodicity. The slithering-snake and the extended reptation algorithms can be implemented in two ways: in a shrinkage-growth or a growth-shrinkage fashion. As growth-shrinkage is just the inverse of shrinkage-growth, we illustrate the procedure for the latter via the example of an isolated chain.^b For the slithering-snake algorithm one chain end is selected at random, the bond to its neighbor is cleaved, and a randomly chosen new bond vector is attached at the other end. If this move respects the excluded-volume condition in the athermal case and additionally passes the Metropolis test in the thermal case, it is accepted. Otherwise the old configuration is recounted. For the extended reptation algorithm the procedure is more complicated. Details may be found for the SAW on a hypercubic lattice in Refs. 51, 76 and for a continuum bond-fluctuation model in Ref. 77.

Usually, shrinkage-growth is preferred to the growth-shrinkage procedure because it is computationally more efficient. The reason for this is illustrated in Fig. 10. The nested configuration of Fig. 10a would be frozen, if a new bond had to be appended before an end bond may be removed. However, it can be unraveled when shrinkage is attempted first. Thus, the shrinkage-growth algorithm is less plagued by –though not exempt of– non-ergodicity effects. An example is provided by the double cul-de-sac configuration of Fig. 10b.^c It is frozen in the shrinkage-growth procedure, but not for the kink-end reptation move shown in Fig. 10c. In fact, kink-end/end-kink moves are known to be ergodic⁷⁶ (as well as other bilocal algorithms; see Ref. 50 for a thorough discussion).

Should one thus abandon the slithering-snake algorithm in favor of extended reptation? Usually, the answer is “No”. For the SAW on the hypercubic lattice problems with ergodicity arise to the constraints imposed by the small coordination number of the lattice. If

^zThe slithering-snake algorithm was invented by Kron in the 1960’s and later independently by Wall and Mandel.⁷⁴ For an overview of applications to SAW’s see e.g. Ref. 43 and to off-lattice models see e.g. Ref. 75.

^aThis generalization of the slithering-snake algorithm was first discussed in detail by Reiter.⁷⁶ More recently, algorithmic and statistical properties of extended reptation moves were analyzed and their implementation was discussed in Refs. 50, 51.

^bFor the multi-chain system the only difference to the isolated chain is that additionally one chain out of the n chains in the systems has to be chosen at random.

^cNon-ergodicity effects are less severe for the slithering-snake algorithm than for the N -conserving local moves discussed before. For the slithering-snake algorithm the ergodicity class of a straight rod contains at least a fraction of $N^{-(\gamma-1)/2}$ of all SAW configurations, whereas this fraction is exponentially small for the local algorithms.^{34,51}

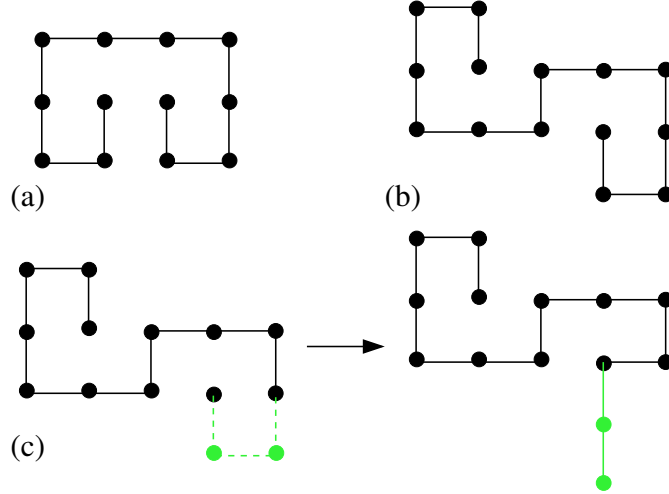


Figure 10. Configurations of 2D SAW's to illustrate the ergodicity problem of the slithering-snake algorithm (a,b) and its solution via extended reptation moves (c). Panel (a) shows a configuration that cannot be moved by slithering-snake moves, if chain growth is attempted first. However, it is not blocked in the shrinkage-growth scheme. By contrast, the configuration of panel (b) is frozen for both growth-shrinkage and shrinkage-growth moves. Panel (c) shows that this configuration may be dissolved by extended reptation moves, e.g. by kink-end reptation if the chain end, where the kink is, happens to be selected for the attachment of the two bonds.

many more bond vectors are a priori possible, as for the bond-fluctuation model or for (typical) continuum models, non-ergodicity should not represent a problem.^d

Relaxation Time and Computational Complexity: Isolated Chains. One expects that the slithering-snake algorithm is able to decorrelate configurations more efficiently than a local updating scheme, the speed-up factor being roughly of order N . This expectation results from the following heuristic argument: The elementary move of the algorithm may be interpreted as a shift of all monomers along the contour of the chain. For the CM this curvilinear motion has two consequences: (1) The curvilinear diffusion coefficient D_c should not depend on N , since all monomers are always shifted at once, irrespective of chain length. Thereby, the slithering-snake algorithm gains a factor of N in regard to the physical reptation dynamics, in which the curvilinear displacement is Rouse-like (Sec. 2.3). (2) An elementary move displaces the CM by $\sim b$ along the chain backbone. After N such moves, the CM has diffused curvilinearly a distance of the order of the contour length $L \propto Nb$. Thus, the relaxation time τ_N should be given by

$$\tau_N = \frac{L^2}{D_c} \Rightarrow \tau_N \sim N^{\approx 2}. \quad (40)$$

^dSee pp. 283/284 of Ref. 43 for further discussion of that point. Contrary to SAW's, the equilibrium configurations of collapsed chains are typically (very) dense. Quite generally with increasing density, the slithering-snake or the extended reptation algorithm become less efficient, as the "free volume" to add new bonds decreases (see e.g. Ref. 78 for a comparison of various algorithms to simulate high-density polymer systems and the subsequent discussion). However, the recent study of Ref. 51 for 2D SAW's with $N \leq 3200$ at the Θ -point suggests that extended reptation is almost as efficient as for pure SAW's with no attractive interactions.

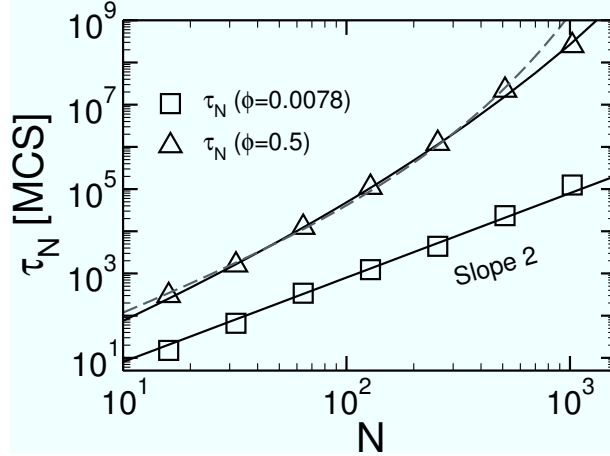


Figure 11. Relaxation time τ_N versus N . τ_N is defined by $g_3(\tau_N) = R_g^2$. In the dilute limit ($\phi = 0.0078$) $\tau_N \sim N^{\approx 2}$, as expected from Eq. (40). In the melt ($\phi = 0.5$), the increase of τ_N with N is stronger. The stretched exponentials are motivated by the activated reptation hypothesis:^{79,82,83} $\tau_N \approx N^2 \exp(0.8N^{1/3})$ (bold line) provides a better description than $\tau_N \approx N^2 \exp(0.074N^{2/3})$ (dashed grey line). Adapted from Ref. 80.

With respect to the computational complexity (page 104) one expects $\tau_{cc} \propto \tau_N$. There is no extra factor of N , as in the case of local moves, for the slithering-snake algorithm. The algorithm is bilocal. It takes a time of order 1 to check and update the chain ends.^e

Note that Eq. (40) is independent of the conformational properties of the chain, contrary to Eq. (39) (which depends on ν). Thus, it should be valid for both 2D and 3D dilute polymer solutions as well as for dense melts. While for the slithering-snake algorithm⁴³ and for some extended reptation algorithms^{51,76} the scaling found for τ_N is very close to Eq. (40), the behavior in dense systems is quite different.^{79,80} The influence of density on the slithering-snake dynamics^f has recently been studied by the bond-fluctuation model.⁸⁰ The following paragraphs briefly summarize some results of this work.

From Dilute Solutions to Dense Melts: A Case Study by the BFM. Reference 80 describes simulations for athermal systems containing chains of length $16 \leq N \leq 1024$ at different volume fractions ϕ . ϕ ranges from dilute solutions to dense melts ($\phi \approx 0.5$; see footnote on page 104). Figure 11 compares the relaxation time τ_N in dilute solution with that in the melt. In dilute solution, the simulation results agree with the prediction of Eq. (40), $\tau_N \sim N^{\approx 2}$. This implies that the assumption of independent, free diffusive motion, which underlies Eq. (40), is well borne out. If this assumption was also true in the melt, the sole effect of density would be to slow down the monomer mobility m . However, the

^eHere, we assume that the time to shift the monomer index along the chain is implemented in a way that it also takes a time of order 1 only.

^fFor local moves the set of allowed bond vectors automatically prevents bonds from crossing each other in the course of the simulation (see Sec. 4.1). If slithering-snake moves are considered, the uncrossability of the bonds has to be checked explicitly to avoid configurations which cannot be attained or unraveled by local updates. Bond crossing can occur if vectors from the classes $[2, 2, 1]$ or $[3, 1, 0]$ are selected.⁸¹

dependence of τ_N on N should not change. Figure 11 shows that this is not true. At $\phi = 0.5$, τ_N increases exponentially with N . This strong slowing-down of the dynamics reflects correlations between the motion of the chains.

The importance of such intermolecular interactions for the polymer dynamics was first discussed by Deutsch.⁷⁹ However, Deutsch goes beyond a mere interpretation of the dynamic properties of the slithering-snake algorithm. He identifies the slithering dynamics with the physical dynamics along the primitive path in the reptation model (see Fig. 5). This suggests an attractive application: The slithering-snake algorithm mimics the back and forth reptation motion of real chains without modeling the (time consuming) local monomer fluctuations around the primitive path. It focuses on the long time behavior of very large chains, where all of these local motions have already relaxed. This suggests that the slithering-snake dynamics may be interpreted in terms of theories proposed for the dynamics of strongly entangled polymer melts, such as the one of Deutsch.⁷⁹

The main results of this theory may be summarized as follows: A chain can reptate through the network of its neighbors only as long as the end monomer does not enter a dense region which prohibits any further forward move. The only way out of the trap is to partially retract and to explore the environment for new pathways. These intermolecular interactions create a free energy barrier which temporarily localizes the chain in the region it initially occupied, and protracts the relaxation. Further relaxation in a dense region could only occur if the chain end encounters another end which moves out of its way. This implies that the portion of the chain, which altered its initial configuration while exploring the environment, should span the typical distance between chain ends d_{end} . Let there be g monomers in this portion. Then, by exploiting the ideality of the chains in the melt, we have $g = (d_{\text{end}}/b)^2 \sim N^{2/3}$ because the density of chain ends scales as $\rho/N \propto d_{\text{end}}^{-3}$. Thus, g is large for long chains. If we now assume that the monomers have to overcome the free energy barrier $g\Delta\mu$, where $\Delta\mu$ is the difference in the monomer chemical potential between the newly explored environment and the region of the initial chain configuration, the barrier is large and the relaxation dynamics should be activated. Thus, $\tau_N \propto N^2 \exp[\text{const } N^{2/3}]$. This is the main prediction by Deutsch. The assumption of a finite $\Delta\mu$ was challenged by Semenov⁸² who suggested that the barrier is due to fluctuations of the molecular field rather than to a permanent chemical potential difference (see also Ref. 83). This picture implies that the barrier should be proportional to $\sqrt{g}$ so that $\tau_N \propto N^2 \exp[\text{const } N^{1/3}]$.

The simulation data of Fig. 11 appear to agree with the latter prediction better than with the original one of Ref. 79 (at least for the chain lengths simulated up to now). Certainly, more work is needed to test these predictions.

Slithering-Snake versus Local Moves. From a merely computational point of view Fig. 11 appears to indicate that the slithering-snake dynamics is not very efficient in equilibrating dense melts. Its relaxation time increases with N more strongly than a power law which is typically found for local updating schemes.⁸ Nevertheless, simulations of the BFM for short chains ($N = 10, 20$) suggest that the slithering-snake algorithm decorrelates melt configurations ($\phi \approx 0.5$) very efficiently.^{84,85}

This point certainly needs more studies. Work in this direction was done in Ref. 80. Figure 12 shows a preliminary result for the diffusion coefficient D_N as a function of chain length. D_N was obtained from simulations employing a mixture of local and slithering-

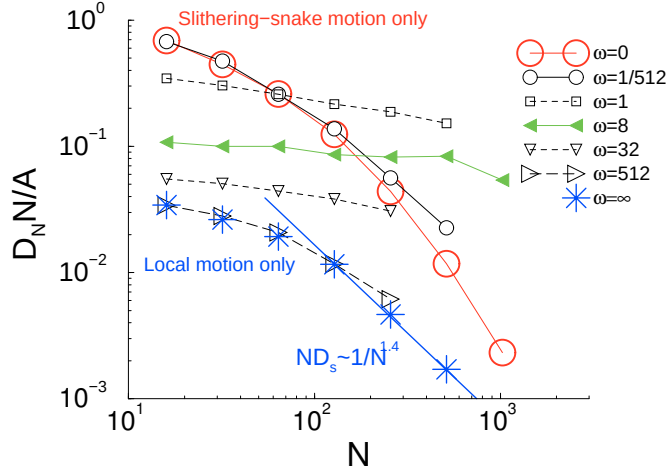


Figure 12. Spatial diffusion coefficient D_N versus N for the BFM at $\phi = 0.5$. Different ratios ω of local to slithering-snake moves are compared: $\omega = 0$ corresponds to pure slithering-snake dynamics, $\omega = \infty$ to the pure local dynamics. The diffusion coefficient is scaled by N/A , where A is the acceptance rate. For both local and slithering-snake dynamics the acceptance rate is roughly $A \approx 0.1$ for all N at $\phi = 0.5$. The data for $\omega \ll 1$ and for $\omega \gg 1$ are very similar to the pure slithering-snake ($\omega = 0$) and the pure local limit ($\omega = \infty$), respectively. For $\omega \approx 8$, ND_N/A is approximately independent of N . This may define a reasonable choice of ω for efficient equilibration of longer chains by local and slithering-snake moves. Adapted from Ref. 80.

snake moves. This introduces, as an additional parameter, the ratio ω of local to slithering-snake moves. The figure indicates that pure slithering-snake dynamics ($\omega = 0$) equilibrates short chains more efficiently than pure local dynamics ($\omega = \infty$), in accord with the observations made in Refs. 84,85. By contrast, with increasing chain length D_N slows down exponentially for the slithering-snake algorithm, as expected due to $D_N \sim R_g^2/\tau_N$, whereas the local dynamics exhibits a crossover from Rouse-like, $D_N \sim 1/N$, to reptation-like behavior, $D_N \sim 1/N^{\approx 2.4}$ (see Sec. 2.3). If this trend persists, the pure slithering-snake algorithm will become inefficient to equilibrate long chains. However, one can speculate that the addition of local moves weakens the confinement imposed by neighboring chains on the slithering-snake dynamics in the large- N limit. Indeed, this seems to be borne out by the data. For short chains ($N < 64$) D_N decreases monotonously with increasing ω , since local moves are less efficient in exploring the configuration space and the confinement is negligible. As N increases, one finds a non-monotonous behavior. The dynamics first becomes more rapid, as local moves are added. This effect appears to saturate at $\omega \approx 10$. Larger values of ω causes the diffusion coefficient to decrease again strongly (at fixed N). This implies that a judicious (model-dependent) choice of ω is crucial if one wants to equilibrate a melt of long chains efficiently by mixing local and slithering-snake moves.

Remark. The efficiency of the slithering-snake algorithm (with or without local moves) or of the extended reptation algorithm deteriorates considerably as ϕ approaches 1, since there is not sufficient space for the growth step. If one is interested in these high den-

sities,^g an alternative simulation method may be provided by reptation moves including a “walker”.^{77,78} A “walker” is defined as an isolated monomer (or as a small group of monomers). In the MC move, the “walker” attaches to a chain in its neighborhood, which then releases a monomer somewhere along its backbone, yielding a new “walker” at a different position than the original one. Since the walker can be created by cleaving a monomer from a chain, the algorithm works even at $\phi = 1$. In order to preserve monodispersity the update of the configuration is finished if the “walker” attaches again to the chain it was originally cleaved from.

5.3 Non-Local Moves: The Pivot Algorithm

A non-local N -conserving move attempts to update a chain portion of order N . If successful, it drastically modifies the chain dimension. Global properties, such as the end-to-end distance, should therefore relax within a few (N -independent) steps so that the critical slowing-down is largely moderated –if not removed.^h

This appealing feature makes the search for appropriate non-local moves very attractive. However, not every conceivable move turns out to be efficient. This is mainly due to two reasons:

1. A drastic modification of the chain configuration is much more likely to violate the excluded-volume constraint than local or bilocal moves do. So, we expect the acceptance rate of non-local updates to decrease with N . The challenge consists in inventing moves whose acceptance rate does not rapidly vanish as $N \rightarrow \infty$.
2. A non-local move typically requires a CPU time of order N (checking self-avoidance, updating the configuration if accepted), in contrast to order 1 for a local or a bilocal move. The extra factor N must be compensated by a very efficient decorrelation of the configurations to justify the use of non-local moves.

A paradigm for a non-local algorithm, which satisfies these criteria, is the “pivot algorithm”.ⁱ

Pivot Algorithm for the SAW. The elementary move of the pivot algorithm works as follows (Fig. 13): First, one randomly chooses a monomer i and a symmetry operation (e.g. a rotation, a reflection, etc.). The monomer serves as a “pivot point” for the symmetry operation which turns the chain portion comprising the monomers $i + 1, \dots, N$ to a new position, while the other piece of the chain (monomers $1, \dots, i$) remains unchanged. The

^gNote that $\phi \approx 1$ means that the polymer melt has zero compressibility κ_T . By contrast, real polymer melts are compressible, with $k_B T \rho \kappa_T$ being of the order 10^{-1} for temperatures above the glass transition or crystallization temperatures.⁸⁶ For the BFM²⁴ at $\phi \approx 0.5$, $k_B T \rho \kappa_T \approx 0.2$ and for the bead-spring models of Sec. 4.2 at $\rho \approx 0.9$, $10^{-2} < k_B T \rho \kappa_T < 10^{-1}$ (see Refs. 24, 87). Thus, it appears that the limit $\phi \rightarrow 1$ is not needed to model dense polymer melts.

^hThis idea is related to that of cluster algorithms^{35,88} employed in MC simulations of spin systems near criticality. Close to the phase transition, the spins are strongly correlated. They form cluster of size ξ ($\equiv$ correlation length, corresponding to R_e in the polymer problem, see Sec. 1). A cluster algorithm finds one cluster (Wolff algorithm) or all of them (Swendsen-Wang algorithm) and updates all spins of the cluster(s) at once. This strongly reduces or even eliminates, in favorable cases, the critical slowing-down.^{35,88}

ⁱThe pivot algorithm was invented by Lal in 1969. A comprehensive discussion of the algorithm may be found in Refs. 34, 89.

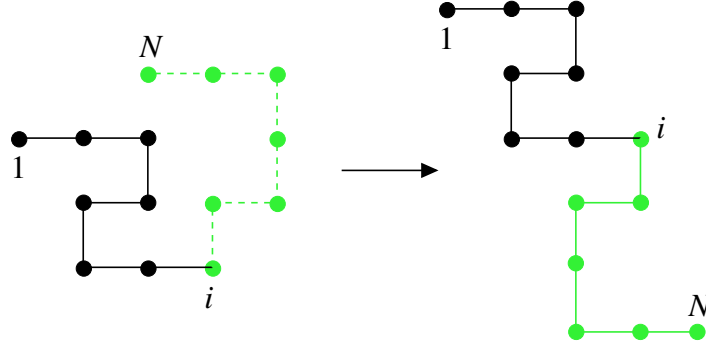


Figure 13. Illustration of the pivot algorithm. A monomer i ($=$ pivot point) is chosen at random. It divides the chain into two pieces: The monomers $1, \dots, i$ remain fixed, while the monomers $i+1, \dots, N$ are translated to a new position via a randomly chosen symmetry operation (rotation, reflection, etc.). In the example of the figure, a 180° rotation around the monomer i is shown.

proposed move is accepted if the resulting configuration is self-avoiding. Otherwise, it is rejected and the old configuration is recounted.

Qualitatively, the pivot move resembles an attempt to construct a SAW of $\sim N$ monomers by joining two SAW's of $\sim N/2$ monomers at the pivot point. The probability for the result to be self-avoiding should scale as $\mathcal{Z}(N)/\mathcal{Z}^2(N/2) \sim N^{-(\gamma-1)}$ [see Eq. (8)]. This heuristic argument suggests that the acceptance rate vanishes as $N^{\approx-0.344}$ in 2D ($\gamma = 43/32$)²¹ and as $N^{\approx-0.158}$ in 3D ($\gamma \simeq 1.158$)²¹. Although these estimates are quantitatively not very accurate, they correctly predict the qualitative trends: The acceptance rate decreases with increasing chain length as a power law $N^{-\alpha}$ and the exponent α is larger in 2D than in 3D (2D: $\alpha \approx 0.19$, 3D: $\alpha \approx 0.11$).⁸⁹

Relaxation Time and Computational Complexity. Fortunately, the numerical value of α is small, implying that even for long chains, e.g. $N = 10^5$, every $N^{0.11} \approx 3.5$ th move is accepted. Since a successful move implies a huge modification of the conformation, one can expect global properties to relax after a few steps. So, the relaxation time τ_N scales as $\tau_N \sim N^\alpha$. This increase is distinctly slower than that of all algorithms discussed so far. Due to this property and due to the fact that the pivot algorithm is known to be ergodic⁸⁹ it has become very popular (see e.g. the compilation of references in Ref. 90). Currently, the pivot algorithm is considered to be the most efficient algorithm for studying configurational properties^j of isolated SAW's.^{21, 89, 92}

The pivot algorithm quickly decorrelates global quantities, such as the end-to-end distance. However, it is not as efficient for local properties: The conformation of a specific monomer is only altered if this monomer is selected as a pivot point and if the move is successful. A successful move takes a time of order N^α and, as the chain consists of N monomers, the decorrelation time of a local observable should scale as $\tau_{\text{loc}} \sim N^{1+\alpha}$. This

^jBy configurational properties we mean quantities characterizing the chain dimension, such as R_e , R_g , etc. A very accurate estimate of the critical exponent ν in 3D ($\nu = 0.5877 \pm 0.0006$) was obtained by the pivot algorithm.⁹¹ By contrast, it appears difficult to measure precisely the partition function $\mathcal{Z}(N)$, and so the exponent γ with the pivot algorithm. For this purpose, other algorithms, such as the “join-and-cut” algorithm²¹ or chain-growth algorithms⁹², are better suited. The current best estimate for γ in 3D is $\gamma = 0.1575 \pm 0.0006$.²¹

extra factor N is felt if one starts from an arbitrary initial configuration. Full equilibration on all length scales is required before large-scale equilibrium properties may be sampled. The equilibration time must be longer than the longest relaxation time in the system, i.e., than τ_{loc} .

For a non-local algorithm the computational complexity is a particularly important quantity because inefficient implementations may ruin the advantage gained by fast decorrelation. The most naive check for self-avoidance would take a time of order N^2 so that $\tau_{\text{cc}} = N^2 \tau_N \sim N^{2+\alpha}$, comparable to the slithering-snake algorithm [Eq. (40)]. Obviously, a faster check is called for. In Ref. 89 it was argued that, by starting at the pivot point and working outwards, self-intersections may be detected in a time of order $N^{1+\alpha}$. This procedure must be repeated $\sim N^\alpha$ times to obtain one accepted pivot. So, the time required per accepted pivot scales as $\sim N^k$.^k Once the pivot is accepted, we still have to update the monomer positions which also takes a time of order N . So, in total we find $\tau_{\text{cc}} \sim N \tau_N \sim N^{1+\alpha}$ for global properties and $\tau_{\text{cc}}^{\text{loc}} \sim N \tau_{\text{loc}} \sim N^{2+\alpha}$ for local properties. The estimate for $\tau_{\text{cc}}^{\text{loc}}$ is again comparable to the slithering-snake algorithm [Eq. (40)]. Therefore, one could also use slithering-snake moves to engender an initial, equilibrated configuration for the pivot algorithm.

5.4 Non-Local Moves in the Melt: The Double-Pivot Algorithm

Due to its efficiency in decorrelating SAW configurations it is tempting to apply the pivot algorithm also to other situations, such as the collapse transition, SAW's in confined geometry, or dense polymer melts. However, in these cases, the algorithm becomes inefficient. The non-local moves either lead to large energy differences (collapse transition) or violate the excluded-volume condition (SAW's in confined geometry, dense melts) so that they are rejected.

Should one therefore give up the idea of using pivot-like moves, say, for dense melts? Recent work suggests that this conclusion might be wrong. Instead of pivoting a piece of one chain to a new position the MC move can involve two chains. Such a move was termed “double-pivot (DP)” algorithm.^l The basic idea of the algorithm is to cleave simultaneously a bond in a chain and in one of its neighbor chains, and to reconnect the monomers such that the chains remain monodisperse.^m The algorithm works as follows (Fig. 14):

1. A monomer, say monomer i in chain a , is chosen at random. Around this monomer the neighborhood is inspected to find bridgeable neighbors on other chains b . A bridgeable neighbor is defined by the following requirements:
 - (a) It must be possible to connect the neighbor to i by a bond vector. This imposes a restriction on the intermolecular distance between i and its neighbor. In the ex-

^kThe pivot algorithm may be implemented so that the time required to obtain an accepted move is of order N^q with $q < 1$.⁹⁰

^lThe double-pivot algorithm has been proposed recently in Ref. 93. In part, this work was motivated by a novel chain-bridging algorithm which was successfully employed in atomistic simulations of long polyethylene chains (see Ref. 94). Our presentation of the DP algorithm is inspired by the discussion of Ref. 94. Due to the newness of the algorithm the implementation that we propose might turn out not to be the most efficient one.

^mIn general, connectivity-altering moves between arbitrary monomer pairs of neighboring chains lead to a distribution of chain lengths, i.e., to polydispersity. The width of the distribution can be controlled by introducing a chemical potential. See Sec. 3.3 of Ref. 43 for further discussion in the context of lattice models and e.g. Ref. 94 for an application to atomistic MC simulations of polyethylene.

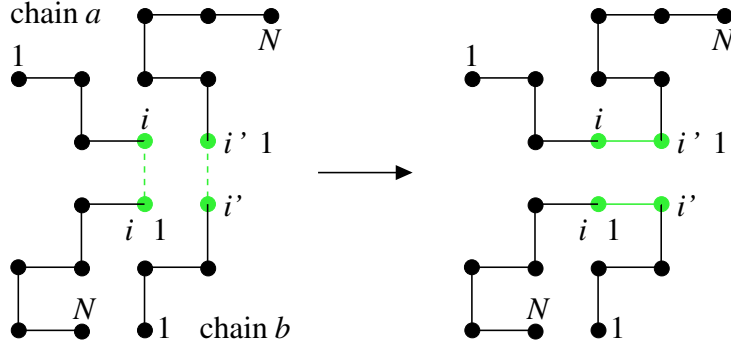


Figure 14. Illustration of a double-pivot (DP) move on a square lattice. A DP move flips bonds between two adjacent chains a and b . To this end, the neighborhood of monomer i in chain a is inspected to find potential bridging sites on chain b . To preserve monodispersity a potential bridging site has to satisfy the following conditions: (a) The $(i + 1)$ th monomer of chain b , denoted by $i' + 1$ in the figure, must be separated by a distance of the bond length from monomer i . (b) The same condition must also hold for the distance of the i th monomer of chain b , denoted by i' , from monomer $i + 1$ of chain a . Between these four monomers a connectivity-altering move is attempted. The bonds from i to $i + 1$ and from i' to $i' + 1$ are broken, and new bonds, i with $i' + 1$ and i' with $i + 1$, are created. The proposed move is accepted or rejected according to Eq. (43). From the figure it is clear that the DP algorithm can only be carried out if there are matching monomers on a neighbor chain within the distance of a bond. This is the more likely, the higher the concentration of the solution. Therefore, the algorithm (presumably) works best in concentrated solutions or melts. If successful, the move entails a drastic change of the chain configuration.

ample of Fig. 14 it must coincide with the lattice constant. For the BFM it must be among the set of allowed bond vectors [Eq. (33)], whereas, for a continuum model, the bond energy resulting from taking the intermolecular distance as a bond vector should not be so large that the proposed bond would never occur in equilibrium. In the latter case, it might be necessary to reduce the strength of the bond potential, e.g., the force constant of Eq. (34).⁹³

- (b) To maintain monodispersity the neighbor must be either the $(i - 1)$ th or the $(i + 1)$ th monomer of the chain b . We distinguish the monomers of chain b by a “prime”, e.g. i' , from those of chain a .
- (c) If it is monomer $i' \pm 1$, monomer i' must be separated from monomer $i \pm 1$ of chain a by a distance which satisfies condition (a).

Using these three criteria we determine the total number of bridgeable neighbors of monomer i , $N_{\text{DP}}(i, \mathbf{x}')$. If $N_{\text{DP}}(i, \mathbf{x}') = 0$ for all i , the configuration $\mathbf{x}'$ must be updated by local (or bilocal) moves to bring the monomers in more favorable positions for bridging.^{93,94}

2. A double-pivot move is initiated by randomly selecting one of the bridgeable neighbors, say $i' + 1$ in chain b , from $N_{\text{DP}}(i, \mathbf{x}')$. Then, the bonds between i and $i + 1$ and between i' and $i' + 1$ are cleaved, and one attempts to create new bonds between i and $i' + 1$ and between i' and $i + 1$. This move just switches the connectivity while

preserving the chain length, and is proposed with probability

$$P_{\text{pro}}(\mathbf{x}' \rightarrow \mathbf{x}) = \frac{1}{N_{\text{DP}}(i, \mathbf{x}')} . \quad (41)$$

3. To satisfy detailed balance we have to determine $P_{\text{pro}}(\mathbf{x} \rightarrow \mathbf{x}')$ of the reverse move. As the forward move only alters the connectivity between the chains, but does not displace the monomers, the number of bridgeable neighbors of a specific monomer remains unchanged. That is, $N_{\text{DP}}(i, \mathbf{x}) = N_{\text{DP}}(i, \mathbf{x}')$. To reverse the forward move, we have to select monomer $i' + 1$ on chain b and its bridgeable neighbor i on chain a . This occurs with probability

$$P_{\text{pro}}(\mathbf{x} \rightarrow \mathbf{x}') = \frac{1}{N_{\text{DP}}(i' + 1, \mathbf{x}')} \quad (42)$$

so that the Metropolis criterion reads

$$\text{acc}(\mathbf{x}' \rightarrow \mathbf{x}) = \min \left(1, \frac{N_{\text{DP}}(i, \mathbf{x}')}{N_{\text{DP}}(i' + 1, \mathbf{x}')} e^{-\beta[U(\mathbf{x}) - U(\mathbf{x}')] } \right) . \quad (43)$$

The difference $U(\mathbf{x}) - U(\mathbf{x}')$ is the local change in energy due to the switching of the bonds between chains a and b .

The steps 1.–3. may be repeated several times. However, the number of iterations should not be too large. Otherwise it is likely that an accepted move annihilates one of its predecessors by performing the transition between two chains in the reverse direction. To avoid this inefficiency it is important to mix up the local configuration of the system. This may be achieved by e.g. local MC moves or by combining⁹³ the DP algorithm with MD simulations.

6 Monte Carlo Methods for Polymers: Rosenbluth Sampling and its Modern Variants

The first MC method to simulate a SAW was “simple sampling”.⁴³ This static method (Sec. 3) works as follows:

1. Place the first monomer at the origin, randomly choose a bond vector, and append it to the monomer.
2. Choose the next bond vector, again randomly, connect it to the second monomer, and check the self-avoidance (Fig. 15).
3. If the chain is self-avoiding, the random growth process may be continued. If not, the self-avoiding piece of the SAW, obtained up to this point, must be discarded, and we have to start from scratch at the first step again.
4. The steps 1. to 3. are repeated until a SAW of the desired length N is obtained. Then, data analysis may be done.
5. Repeat steps 1. to 4. to gather sufficient statistics.

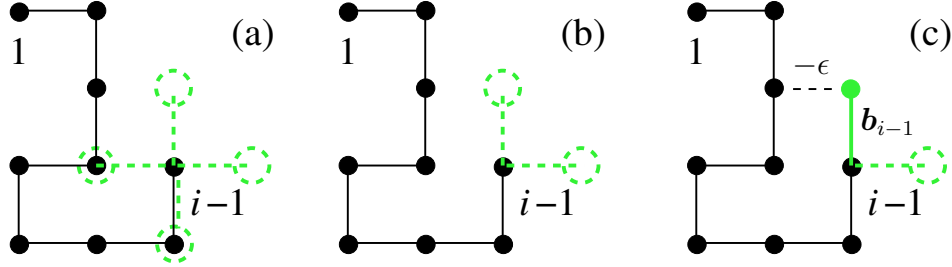


Figure 15. Illustration of the simple-sampling (a) and the Rosenbluth-Rosenbluth (RR) methods [(b) and (c)] on a square lattice. The coordination number of the lattice $z (= 4)$ defines the number of possible bond vectors $\{\mathbf{b}\}$. In simple sampling, all of these vectors have the same a priori probability. Thus, the bond vector $\mathbf{b}_{i-1}$ ($= \mathbf{r}_i - \mathbf{r}_{i-1}$) from monomer $i-1$ to the new monomer i can also point in the direction of already occupied lattice sites. This leads to the attrition problem. The RR method strongly reduces the attrition by taking $\mathbf{b}_{i-1}$ only from the open directions k_i ($= 2$ in the example of the figure). The new bond vector can be chosen either uniformly from the k_i possibilities, as indicated by the dashed lines and dashed circles in panel (b), or according to the local Boltzmann factor [Eq. (48)], if e.g. an attractive monomer-monomer interaction is present [filled grey circle in panel (c)].

Apparently, completed SAW's are independent of one another and occur with the same probability $P_s(\mathbf{x}) = P_{\text{sa}}(N)$. This is the main advantage of a static MC method. The main disadvantage of simple sampling is that $P_{\text{sa}}(N)$ becomes exponentially small for large N . To see that, let us calculate $P_{\text{sa}}(N)$ for a SAW on a hypercubic lattice. The hypercubic lattice has the coordination number $z = 2d$. Thus, the number of random walks (RW's) starting at the origin and having $N - 1$ steps, is $\mathcal{Z}_{\text{RW}} = z^{N-1}$. $\mathcal{Z}_{\text{RW}}$ is the partition function of the RW. Out of these z^{N-1} random-walk configurations simple sampling selects those which are self-avoiding. As there are $\mathcal{Z} [\sim \mu^N N^{\gamma-1}$, Eq. (8)] such configurations, $P_{\text{sa}}(N)$ is given by

$$P_{\text{sa}}(N) \sim \left(\frac{\mu}{z}\right)^N N^{\gamma-1} = N^{\gamma-1} e^{-\lambda N} \sim e^{-\lambda N} \quad (N \text{ large}), \quad (44)$$

where $\lambda = \ln(z/\mu)$ is called “attrition constant”. The attrition constant is $\lambda > 0$ (in 2D and 3D),ⁿ reflecting that the monomer partition function, i.e., the number of ways to place a monomer on the lattice, of a RW ($= z$) is, due to the neglect of self-avoidance, larger than that of a SAW ($= \mu$).

Equation (44) illustrates that simple sampling is not an efficient simulation method for SAW's. On average, $e^{\lambda N}$ random walks have to be constructed to obtain one SAW (“attrition problem”). As $\lambda = 0.416$ and 0.248 for the 2D and 3D hypercubic lattices, this number of constructions is prohibitively large already for $N \gtrsim 50$. Even if one modifies the construction method by avoiding immediate backfolds, thereby replacing z by $z - 1$ (“Non-Reversal Random Walk (NRRW)”), the exponential attrition remains. Generation of SAW's with $N > 10^2$ is still unfeasible. Therefore, all alternative simulation techniques have to alleviate this attrition problem.

ⁿFor a compilation of attrition constants see Refs. 34, 43. If $d \rightarrow \infty$, λ goes to zero as $\lambda \rightarrow 1/2d$ for the hypercubic lattice.

6.1 Inversely Restricted Sampling: Rosenbluth-Rosenbluth Method

The attrition problem arises because simple sampling chooses blindly from the nearest-neighbor sites to place a new monomer. A more clever algorithm could scan the local environment around the last monomer and exclude those trial directions which lead to self-intersections. The position of the next monomer i can then be chosen with equal probability from the remaining k_i open directions (Fig. 15).^o This method, known as “inversely restricted sampling” or “Rosenbluth–Rosenbluth (RR) algorithm”,^p strongly reduces the attrition^q at the expense of introducing a bias. A SAW is not generated with uniform probability, but with probability

$$P_s(\mathbf{x}) = \prod_{i=1}^N \frac{1}{k_i} \quad (k_1 = 1) . \quad (45)$$

Equation (45) shows that configurations with small k_i ’s have a higher probability of occurring. This bias toward dense configurations in the production of a SAW must be corrected in its analysis by the weight $W(\mathbf{x}) \propto 1/P_s(\mathbf{x})$ when calculating observables [see Eq. (18)].

The RR algorithm is a static MC method. As such, it has the advantage that successively generated SAW’s are independent of each other. All problems of decorrelating configurations, discussed in Sec. 5, are absent by construction. On the other hand, Eq. (45) also points to the major difficulty of the method. The RR method favors dense configurations which are not representative of long SAW’s. Thus, $P_s(\mathbf{x})$ differs from $P_{eq}(\mathbf{x})$. As N increases, the difference becomes more pronounced. To compensate the discrepancy between the two probabilities the distribution of weights must become broad: Dense configurations have small weights and open chains have in general larger weights. A detailed analysis of this problem was undertaken by Batoulis and Kremer.⁹⁵ They showed that the distribution of weights, obtained from M repetitions of the RR method ($\{W(\mathbf{x}_m), m = 1, \dots, M\}$), is dominated by few configurations having the largest weights. The most relevant SAW configurations have, however, smaller weights. To sample this portion of the weight distribution sufficiently, M has to become very large [see the discussion of Eq. (22)]. This problem makes the RR algorithm not suitable for the simulation of long SAW’s.^r

However, the RR algorithm should be well suited if the bias introduced by the sampling engenders configurations which are close to the physical ones. That is, if the equilibrium configurations are less swollen than those of a SAW. As this is the case close to Θ -point in 3D (see Sec. 1), one might expect the RR algorithm to be more efficient for $T \approx T_\Theta$. In fact, this expectation is nicely borne out. The biased sampling of the RR method produces

^oHere, we assume $0 < k_i \leq z - 1$. If all neighbor sites are blocked ($k_i = 0$), the configuration obtained until then is trapped. It must be discarded, and the construction resumes from the beginning. The first monomer can be placed anywhere on the lattice. If the construction always starts, say, from the origin, we set $k_1 = 1$ in Eq. (45).

^pSee Ref. 95 for a detailed statistical analysis of the RR algorithm⁹⁶ in the context of simulating SAW’s. Chapter 11 of Ref. 31 gives a discussion in the larger context of free energy calculations with applications to discrete and continuous chain models.

^qIn the RR algorithm, attrition occurs because the growing chain may be trapped, i.e., $k_i = 0$ for some i . In practice, this does not appear to be a serious problem in 3D for high-coordination lattices.^{95,97} It was suggested that effects of trapping should become visible only for $N \gtrsim 10^4$.⁹⁷ This implies that there is still exponential attrition, as $N \rightarrow \infty$, but at a much lower rate than in Eq. (44).

^rIn the simulations on the (high-coordination) FCC lattice of Ref. 95 the systematic error due to the weights rendered a precise determination of R_g impossible for $N \gtrsim 200$.

an effective attraction between the monomers which closely resembles that of Θ -chains. An important consequence of this special property is that the weights are nearly compensated by the Boltzmann factor.⁹⁷ Thus, the RR method was employed to study properties near the Θ -point (see e.g. Refs. 10,11 of Ref. 97), and it also represents the core of a modern algorithm, the “Pruned-Enriched Rosenbluth Method (PERM)⁴⁷”.

6.2 Pruned-Enriched Rosenbluth Method (PERM)

The simulation of a polymer chain close to the Θ -point requires the introduction of an attraction between the monomers to compensate their mutual repulsion. Typically, these thermal interactions are modeled by a short range inter-monomer potential.^s As a chain is grown according to the Rosenbluth scheme, the presence of the potential implies that the internal energy of a chain changes:

$$u_i(\mathbf{b}_{i-1}) = U(\mathbf{r}_1, \dots, \mathbf{r}_{i-1} + \mathbf{b}_{i-1}) - U(\mathbf{r}_1, \dots, \mathbf{r}_{i-1}) \quad [u_1(\mathbf{b}_0) := U(\mathbf{r}_1)] . \quad (46)$$

Here, $U(\mathbf{r}_1, \dots, \mathbf{r}_{i-1})$ is the potential energy of chain having $(i-1)$ monomers and $\mathbf{b}_{i-1}$ is the bond vector from the $(i-1)$ th to the i th monomer. A priori, there are different ways to incorporate $u_i(\mathbf{b}_{i-1})$ in the construction. Two possible choices are the following:⁴⁷

1. The first method is the classical Rosenbluth scheme. The position of the i th monomer is chosen from the free neighbors with uniform probability [see Eq. (45)]. Thus, the weight of the new chain configuration $\mathbf{x} (= \mathbf{r}_1, \dots, \mathbf{r}_N)$ is given by^t

$$W(\mathbf{x}) = \frac{e^{-\beta U(\mathbf{x})}}{P_s(\mathbf{x})} = \prod_{i=1}^N \left[k_i e^{-\beta u_i(\mathbf{b}_{i-1})} \right] . \quad (47)$$

2. An alternative consists in including the Boltzmann factor in the probability $P_{s,i}$ for placing the i th monomer. Let $\{\mathbf{b}\}$ denote the ensemble of possible bond vectors. For the hypercubic lattice, $\{\mathbf{b}\}$ coincides with the number of lattice directions z , for the BFM it is given by Eq. (33). Then, we may write for $P_{s,i}$ (Fig. 15)

$$P_{s,i} = \frac{e^{-\beta u_i(\mathbf{b}_{i-1})}}{\sum_{\{\mathbf{b}\}} e^{-\beta u_i(\mathbf{b})}} = \frac{e^{-\beta u_i(\mathbf{b}_{i-1})}}{w_i} \xrightarrow{\text{SAW}} \frac{1}{k_i} , \quad (48)$$

where the normalization w_i reduces to k_i if only excluded-volume interactions are taken into account (SAW limit). This implies that the weight $W(\mathbf{x})$ is given by

$$W(\mathbf{x}) = \frac{e^{-\beta U(\mathbf{x})}}{P_s(\mathbf{x})} = \prod_{i=1}^N \left[\frac{e^{-\beta u_i(\mathbf{b}_{i-1})}}{P_{s,i}} \right] = \prod_{i=1}^N w_i . \quad (49)$$

Both methods have been used to simulate Θ -polymers via the Rosenbluth algorithm.⁴⁷

Despite the fact that, precisely at T_Θ , the RR configurations more or less coincide with the equilibrium configurations, the accuracy of the method deteriorates for $N > N_c$.^u

^sFor the SAW on the hypercubic lattice the attraction is usually implemented between non-bonded nearest neighbors.⁴⁷ In a simulation of the Θ -transition with the BFM, a square-well potential of range $\sqrt{6}$ (in units of the lattice constant) was used.⁶⁰ This choice ensures that the first peak of the pair-distribution function is encompassed by the range of the potential. A more complicated choice was made in Ref. 98. In the continuum, a Gaussian chain model with a non-truncated LJ-interaction was extensively studied (see Refs. 64,99 and references therein).

^tNote that, contrary to the definition of $W(\mathbf{x})$ in Sec. 3, Eq. (47) does not include the factor of the (unknown) partition function.

^uThe critical chain length N_c depends on the model. For a simple cubic lattice it is $N_c \approx 10^3$.⁴⁷

Again, the reason is that $P_s(\mathbf{x})$ does not perfectly agree with $P_{\text{eq}}(\mathbf{x})$. As a result, the weight distribution becomes so broad that chains with the biggest weights dominate the sample (see Fig. 2 of Ref. 47). Due to Eq. (18) this leads to a large variance of the computed observables.

To improve the accuracy one has to reduce the variance. Grassberger proposed a clever way to achieve this.⁴⁷ Assume that we have constructed a chain up to monomer i ($1 \leq i \leq N$) via the RR method. This chain has the weight W_i which we want to prevent from fluctuating too much. That is, if W_i exceeds a lower or an upper bound, we interfere in the following way:

1. If $W_i < W_i^-$, we “prune” the sample: If W_i decreases below the threshold W_i^- , a random number $0 \leq \zeta \leq 1$ is uniformly drawn. If $\zeta < 1/2$, the chain is removed. Otherwise, it is kept, its weight is doubled ($W_i \rightarrow 2W_i$), and the step-by-step growth continues.
2. If $W_i > W_i^+$, we “enrich^v” the sample: If W_i exceeds the upper bound W_i^+ , c copies, typically 2,¹⁰⁰ of the configuration are made, each of which is given the new weight $W_i \rightarrow W_i/c$. These copies are then grown independently of each other.

This control of the weight distribution within the RR algorithm was termed “Pruned-Enriched Rosenbluth Method (PERM)”.⁴⁷

Of course, the question arises of how to choose the bounds $W_i^\pm$. Here, it is important to note that neither the pruning nor the enrichment step introduces any bias. In the calculation of the sums in Eq. (18) the increase of W_i by pruning is compensated by the probability $1/2$ with which the configuration is retained, and the decrease of W_i in the enrichment is compensated by the number of copies c . Thus, we are free to choose the bounds $W_i^\pm$. Bad choices can “only” render the method inefficient, but not incorrect. In order to determine optimum values for $W_i^\pm$ the following procedure was proposed (for temperatures that are not too low):^{47,100}

- First, one chooses $W_i^- = 0$ and W_i^+ very large. That is, one performs a simulation via the original RR method. This simulation yields the weights W_i for $i = 1, \dots, N$. First estimates for the bounds $W_i^\pm$ are then determined by $W_i^- = C^- W_i$ and $W_i^+ = C^+ W_i$ with $C^+/C^- \approx \text{O}(1)\text{--}\text{O}(10)$.
- These estimates are refined “on the fly”. Imagine that we have obtained M_i configurations of chain length i from the simulation. Then, we first calculate the partition function by $[\mathbf{x} = (\mathbf{r}_1, \dots, \mathbf{r}_i)]$, see Eq. (18)]

$$\mathcal{Z}_i = \int d\mathbf{x} e^{-\beta U_i(\mathbf{x})} \approx \frac{1}{M_i} \sum_{m=1}^{M_i} \frac{e^{-\beta U_i(\mathbf{x}_m)}}{P_s(\mathbf{x}_m)} = \frac{1}{M_i} \sum_{m=1}^{M_i} W_i(\mathbf{x}_m), \quad (50)$$

and from that, we determine the new bounds by $W_i^\pm = C^\pm \mathcal{Z}_i$.

^vEnrichment is a classical technique for simulating SAW’s. Briefly, it works as follows: If a chain survives the s step, c copies of its configuration are made, which serve as independent starting points for further growth. The method may be implemented in a “breadth-first” or a “depth-first” fashion. The former implies that all copies are first grown to size $2s$ before the entire sample is copied again. By contrast, the latter method tries to complete the construction of one copy up to chain length N before passing to next one. The pros and cons of the two implementations are discussed in the context of the PERM in Ref. 47. More details about enrichment may be found in Refs. 34, 43.

Applications of the PERM. The PERM was invented to simulate the transition from an excluded-volume to a collapsed chain at the Θ -temperature T_Θ . Theoretically, the transition is usually identified with a tricritical point in the limit $N \rightarrow \infty$.¹⁶ A tricritical point exhibits mean-field behavior in 3D. Thus, one expects that $\nu = 1/2$ and $\gamma = 1$ [Eq. (8)] at T_Θ for $N \rightarrow \infty$. This asymptotic large- N behavior is supplemented by (universal) corrections of order $1/\ln N$ for finite chain length (see e.g. Chap. 21 of Ref. 16 for a good discussion). A significant test of the theory therefore requires to study very long chains.

In Ref. 47 such a test was attempted. Grassberger performed a comparative study of various models employed in the literature: a SAW on a simple cubic lattice with attractive nearest-neighbor interactions, the BFM with two versions for the attractive monomer-monomer interactions,^{60,98} and a LJ bead-spring model⁶⁴. Using the PERM these models could be simulated with high precision and, partly, with much longer chain lengths than studied before. These simulations yielded refined estimates of T_Θ for the various models, confirmed the mean-field-like asymptotic character of the Θ -point, but also showed that the leading-order logarithmic corrections cannot explain the finite- N behavior found, even for $N = 10^4$.^w

On the technical side, it was found that the selection of the position of the next monomer i with uniform probability from the k_i open directions is only the best choice for the SAW on the simple cubic lattice (first method of page 118). As alluded to at the end of Sec. 6.1, this is due to a near cancellation of the Rosenbluth weight and the Boltzmann factor: Many nearest-neighbor contacts lead to a low Rosenbluth weight, but to a large Boltzmann factor, and vice versa. For more long-range or more complicated interactions the degree of cancellation need not necessarily be the same. In fact, for the BFM it was found in Ref. 47 that a selection of the next monomer position according to the Boltzmann factor [Eq. (48)] is more efficient. Similar approaches were also used to study e.g. simple models of proteins.

The preceding discussion appears to suggest that the PERM is a single-chain technique. This is not true. We just quote two recent examples. The PERM was utilized to simulate the denaturation transition of a simple model for double-stranded DNA (two SAW's).¹⁰² A truly multi-chain system was studied in Ref. 103. This work is concerned with the phase diagram of semidilute polymer solutions for $T \leq \Theta$ (see Fig. 3). For a review of these and other applications see Ref. 100.

6.3 Configurational-Bias Monte Carlo and Recoil-Growth Algorithm

An alternative multi-chain MC scheme, incorporating the RR method, is the Configurational-Bias Monte Carlo (CBMC) algorithm.^x Contrary to the PERM, CBMC builds up

^wThese findings elicited further theoretical¹⁰¹ and numerical work.^{99,101} On the theoretical side, subleading corrections of order $\ln(\ln N)/(\ln N)^2$ were calculated and found to be as large as the leading $1/\ln N$ term. On the numerical side, MC simulations¹⁰¹ for $N \leq 10^4$ of a NRRW, including weakly attractive two-body, but repulsive three-body interactions, were performed. This model shows logarithmic corrections which are much weaker than those found in Ref. 47 and are roughly compatible with the theoretical predictions. However, Ref. 101 stresses a problem in the analysis of the Θ -point. To estimate T_Θ , an infinite- N property, precisely from the simulations one has to rely on the theoretical predictions for the finite- N corrections to extrapolate to $N \rightarrow \infty$. To this end, the simulated chains must be long enough for these corrections to apply. This regime appears to be very hard to attain, even for $N \sim 10^6$.

^xCBMC was introduced by Siepmann and Frenkel in the 1990s. It can be applied to lattice and off-lattice models. The initial off-lattice applications have demonstrated the power of the algorithm for the study of a large

a new chain step-by-step without controlling the weights w_i “on the go”. It is only after a successful construction that the resultant bias is removed: The new chain is accepted according to the Metropolis criterion with a probability dependent on the total Rosenbluth weights of the new and the old chain configurations. This additional test warrants sampling from the Boltzmann distribution.

A recent extension of CBMC is the “Recoil-Growth (RG) algorithm”.^y Contrary to the RR method, which only looks ahead one step while constructing a chain, the RG algorithm uses a more sophisticated growth procedure. It places a long retractable feeler at the head of a growing chain. The feeler spys out the environment to find favorable pathways for the chain construction. The efficiency of the method resides in the fact that the growth does not terminate if the feeler encounters a trap. It merely recoils back from the trap and pursues its search in a different direction. After the construction is completed, the new chain replaces the old one, just as in CBMC, with a probability determined by their respective weights according to the Metropolis criterion.

6.3.1 Configurational-Bias Monte Carlo (CBMC)

To illustrate the CBMC scheme in more detail we consider a solution of SAW’s on a lattice (Fig. 16):

1. Given the initial configuration $\mathbf{x}'$ of the system we randomly select a chain and one of its monomers. Let this be the monomer n' of chain b ($1 \leq n' \leq N$). The configuration of the chain portion $i = n', \dots, N$ is characterized by the sequence of bonds $\mathbf{b}_{n'-1}, \dots, \mathbf{b}_{N-1}$. Each bond represents a specific choice from the set of all possible bond vectors $\{\mathbf{b}\}$. So, we can write the Rosenbluth weight of the monomers $n', \dots, N$ of chain b as

$$W^b(\mathbf{x}') = \prod_{i=n'}^N w_i^b(\mathbf{x}') , \quad w_i^b(\mathbf{x}') = e^{-\beta u_i^b(\mathbf{b}_{i-1})} + \sum_{\mathbf{b} \neq \mathbf{b}_{i-1}} e^{-\beta u_i^b(\mathbf{b})} . \quad (51)$$

Here, $u_{n'}^b(\mathbf{b}_{n'-1})$ is the energy of monomer n' at its actual position in the chain. It includes the interactions with the monomers of all other chains and with the monomers $i = 1, \dots, n' - 1$ of its own chain [Eq. (46)]. The monomers $n' + 1$ to N have to be omitted because one thinks of the chain b as being (re-) constructed step-by-step via the RR procedure with probability [Eq.(48)]

$$P_{s,i}^b(\mathbf{x}') = \frac{e^{-\beta u_i^b(\mathbf{b}_{i-1})}}{w_i^b(\mathbf{x}')} . \quad (52)$$

Thus, the potential energy of the chain portion $n', \dots, N$ is given by

$$U^b(\mathbf{x}') = \sum_{i=n'}^N u_i^b(\mathbf{b}_{i-1}) , \quad (53)$$

variety of problems in polymer physics. Therefore, CBMC has become an important and widely used simulation technique. A comprehensive and very pedagogical account of the method, including flowcharts of the algorithm and examples, is given in the textbook by Frenkel and Smit.³¹

^yThe RG algorithm was introduced in Refs. 104, 105. A detailed description may be found in Chap. 13.7 of the textbook by Frenkel and Smit.³¹ For practical applications, it is very helpful that the FORTRAN codes of the “Case Studies” may be downloaded from http://molsim.chem.uva.nl/frenkel_smit.

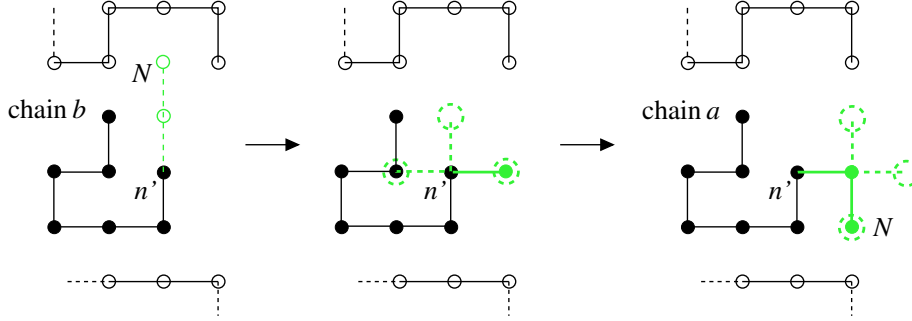


Figure 16. Schematic of a CBMC move executed in a polymer solution. The initial configuration $\mathbf{x}'$ of the solution is shown in the left figure. From all chains of $\mathbf{x}'$ the chain b and one of its monomers, n' , are chosen. Here, $n' = N - 2$. The chain portion $n', \dots, N$ is deleted (dashed grey lines connecting the open grey circles) and reconstructed step-by-step. Upon completion we obtain a new chain configuration, “chain a ”, and so a new configuration, $\mathbf{x}$, of the solution. In the construction, the new bond $\mathbf{b}_{i-1}$ from monomer $i - 1$ to monomer i ($= N - 1, N$) is selected from the set of bond vectors $\{\mathbf{b}\}$ according to the local Boltzmann factor $P_{s,i}^a(\mathbf{x})$ [Eq. (55)]. For lattice models this set, which defines the number of trial directions at each step, is finite and closely related to lattice structure: E.g. for the hypercubic lattice, it coincides with the number of lattice directions z (or $z - 1$ if backfolds shall be excluded a priori). For the BFM the trial directions may be taken as the ensemble of allowed bonds [Eq. (33)]. For a continuum model there is a priori an infinite number of possible directions, from which a suitable finite number of trial directions must be selected so that new chain configurations are efficiently generated (see Chap. 13.3 of Ref. 31 for details). If only excluded volume interactions are present, as assumed in the figure, $P_{s,i}^a(\mathbf{x}) = 1/2$ for $i = N - 1$ and $P_{s,i}^a(\mathbf{x}) = 1/3$ for the last monomer N . Thus, the total weight $W^a(\mathbf{x})$ of chain a is 6, whereas that of the chain portion $n' = N - 1, N$ of the old chain b is $W^b(\mathbf{x}') = 4$. Thus, $W^a(\mathbf{x})/W^b(\mathbf{x}') = 1.5$, and chain a will be accepted according to Eq. (58).

and the probability to propose the chain portion may be expressed as

$$P_s^b(\mathbf{x}') = \prod_{i=n'}^N P_{s,i}^b(\mathbf{x}') = \frac{e^{-\beta U^b(\mathbf{x}')}}{W^b(\mathbf{x}')} . \quad (54)$$

2. The monomers n' to N of chain b are deleted. This corresponds to a “shrinkage-growth” implementation of the algorithm, which is (presumably) more efficient than a “growth-shrinkage” procedure (see Sec. 5.2).
3. To obtain a new configuration $\mathbf{x}$ of the system the chain b is fully ($n' = 1$) or only in part ($n' > 1$) reconstructed:
 - If $n' = 1$, we start building a new chain by randomly placing the first monomer somewhere in system. Let this new chain be labeled “chain a ”. At the position, where the monomer is inserted, it interacts with the other chains of the system. It has an energy $u_{n=1}^a(\mathbf{b}_0)$ [$= U(\mathbf{r}_1)$, see Eq. (46)], giving rise to the weight $w_1 = \exp[-\beta u_1^a(\mathbf{b}_0)]$.
 - If $n' > 1$, we rebuild the chain portion $i = n', \dots, N$ monomer-by-monomer via the RR method. This growing chain, comprising the initial portion of chain b ($i = 1, \dots, n' - 1$) and the newly attached monomers ($i = n, \dots, N$) will also be called “chain a ”. In the following we suppress the prime ($'$) to indicate that a new chain conformation is obtained, even if the first $n-1$ monomers are identical

with those of chain b . To add the n th monomer we proceed as described in the next step.

4. A monomer may be attached to the growing chain via a bond from the set of possible bond vectors $\{\mathbf{b}\}$. Out of these trial directions we choose one according to Eq. (48). This implies for the monomer i ($= n, \dots, N$)

$$P_{s,i}^a(\mathbf{x}) = \frac{e^{-\beta u_i^a(\mathbf{b}_{i-1})}}{\sum_{\{\mathbf{b}\}} e^{-\beta u_i^a(\mathbf{b})}} = \frac{e^{-\beta u_i^a(\mathbf{b}_{i-1})}}{w_i^a(\mathbf{x})}, \quad (55)$$

where $u_i^a(\mathbf{b}_{i-1})$ is the change in potential energy of the system due to the addition of the new bond $\mathbf{b}_{i-1}$ from monomer $i-1$ to monomer i [Eq. (46)].

5. The preceding step is repeated until the construction of the chain is completed. Thus, the new chain configuration occurs with probability [Eq. (49)]

$$P_s^a(\mathbf{x}) = \prod_{i=n}^N P_{s,i}^a(\mathbf{x}) = \frac{e^{-\beta U^a(\mathbf{x})}}{W^a(\mathbf{x})}, \quad (56)$$

where $W^a(\mathbf{x})$ and $U^a(\mathbf{x})$ are defined analogously to Eqs. (51,53).

6. Now, we recognize that $P_s^a(\mathbf{x})$ may be interpreted as the probability of proposing a transition from the old configuration $\mathbf{x}'$ to the new configuration $\mathbf{x}$. Correspondingly, $P_s^b(\mathbf{x}')$ is the probability for the reverse step. That is,

$$P_s^a(\mathbf{x}) = P_{\text{pro}}(\mathbf{x}' \rightarrow \mathbf{x}) \quad \text{and} \quad P_s^b(\mathbf{x}') = P_{\text{pro}}(\mathbf{x} \rightarrow \mathbf{x}'). \quad (57)$$

Finally, we insert Eq. (57) into Eq. (29) to obtain the acceptance probability for the new configuration

$$\begin{aligned} \text{acc}(\mathbf{x}' \rightarrow \mathbf{x}) &= \min \left(1, \frac{P_{\text{pro}}(\mathbf{x} \rightarrow \mathbf{x}')}{P_{\text{pro}}(\mathbf{x}' \rightarrow \mathbf{x})} e^{-\beta[U^a(\mathbf{x}) - U^b(\mathbf{x}')] } \right) \\ &= \min \left(1, \frac{W^a(\mathbf{x})}{W^b(\mathbf{x}')} \right). \end{aligned} \quad (58)$$

Discussion. An important feature of the CBMC method is its non-local character: A successful CBMC construction implies a large-scale configurational change. Either a new chain is inserted somewhere in the system –this may be employed very efficiently to study phase equilibria of polymer solutions in the bulk¹⁰⁶ and in thin films¹⁰⁷– or part^z of a chain is regrown. This leads to a rapid decorrelation of chain configurations and to efficient sampling, provided the system has a low or moderate density and the chains are not too long. When dealing with long chains and/or dense melts the following problems occur:

^zThis part need not necessarily start at monomer $n' + 1$ and terminate at the chain end, as assumed in our previous discussion. It can also comprise an inner portion of the chain, say the monomers $n' + 1, \dots, n' + m' - 1 < N$. In this case, the reconstruction has to satisfy the additional condition that it starts at monomer n' and ends at the position of monomer $n' + m'$ (see Chap. 13.4 of Ref. 31). This variant of CBMC may be used to relax e.g. ring polymers, which have no free ends so that the method described above could not be applied.

- Chain construction in CBMC is based on the Rosenbluth method, yielding a distribution of configurations that differs from the Boltzmann distribution. This difference becomes more pronounced with increasing N , even for an isolated chain at the Θ -point (see Secs. 6.1,6.2). As CBMC does not control the weights while synthesizing the chain, contrary to the PERM, the acceptance rate for chain reconstructions falls exponentially in N for large chain lengths.
- Another problem results from the “shortsightedness” of the RR algorithm. By looking only one step ahead, the chain construction can run into traps. For isolated chains this trapping implies that there is still an exponential attrition in N for long chains, albeit with an attrition constant much smaller than λ (see Eq. (44) and footnote on page 117).
- In a dense system, in addition to trapping, a further problem occurs. If a (part of a) new chain is inserted, it is fairly likely to be constructed just in the space originally occupied by the (part of the) old chain which was removed. This can lead to strong correlations between the new and old chain configurations so that sampling becomes inefficient.⁷⁵ In this situation, it might be best to combine CBMC with the slithering-snake algorithm (Sec. 5.2). That is, to try to regrow just the terminal bond at the other chain end according to Eq. (58). In a dense melt, it thus appears as if CBMC cannot be expected to decorrelate chain configurations more efficiently than the slithering-snake algorithm.

6.3.2 Recoil-Growth (RG) Algorithm

The RG algorithm was suggested as an alternative to CBMC,^a exhibiting two major changes:

1. Instead of looking one step ahead the RG algorithm scans the environment via a retractable “feeler”. The feeler consists of a self-avoiding chain portion having at most N_{recoil} monomers.^b The ability of the feeler to shrink and to grow helps to circumvent dense regions. This allows for the search of suitable pathways to complete the construction of the chain.
2. Contrary to CBMC, the incremental weights w_i for each newly added monomer are not calculated “on the fly”, but only after a new chain has been successfully constructed. Thus, the computation of the weights is carried out only once. In CBMC, it can happen that a lot of time is spent to calculate the weights of a partially grown

^aThe RG algorithm was introduced for SAW’s on a cubic lattice in Ref. 104. Reference 105 extends this study to continuum models. A comprehensive discussion of these works may be found in the textbook by Frenkel and Smit (Chap. 13.7).³¹

^bOf course, in general $N_{\text{recoil}} > 2$, implying that the feeler is longer than one bond. Otherwise, the “shortsightedness” of CBMC is not removed. The idea to improve the RR method by looking several steps ahead is not new. It is embodied e.g. in the “scanning method” of Meirovitch.¹⁰⁸ This method still uses a one-step growth, but chooses a new bond $\mathbf{b}_i$ according to the probability that a SAW of N_{scan} monomers can be constructed in direction of $\mathbf{b}_i$. As this implies an enumeration of all possible SAW’s of length N_{scan} starting at some monomer i , the scanning parameter N_{scan} is usually much smaller than N . Thus, trapping cannot be avoided completely. This would only be the case if $N_{\text{scan}} = N - i$, i.e., if one scanned all possible ways to complete the chain up to monomer N in direction of $\mathbf{b}_i$. For a comparative discussion of the RG algorithm and the scanning method see Ref. 104.

chain which must then be discarded because the construction has run into a trap before completion.

Description of the RG Algorithm. These differences show that the RG algorithm comprises two independent steps: a construction step using the retractable feeler and an acceptance step including the weight calculation. We discuss these steps separately. In our discussion we assume that a whole chain is inserted in the system after a randomly chosen old chain has been removed (“shrinkage-growth procedure”). As before, the new chain is called “chain a ”, the old “chain b ”, and the respective new and old configurations of the system are denoted by $\mathbf{x}$ and $\mathbf{x}'$.

Construction step (Fig. 17):

1. We place the first monomer of chain a at the random trial position $\mathbf{r}_1$ and determine its energy $U(\mathbf{r}_1) = u_1^a(\mathbf{b}_0)$. To decide whether the position is accessible (“open”) we accept it with probability (see e.g. Ref. 31 for further discussion of this point)

$$P_1^{\text{open}}(\mathbf{x}) = \min \left(1, e^{-\beta u_1^a(\mathbf{b}_0)} \right) \xrightarrow{\text{SAW}} \begin{cases} 1 & \text{if } \mathbf{r}_1 \text{ is unoccupied,} \\ 0 & \text{otherwise,} \end{cases} \quad (59)$$

where “SAW” means that there are only hard-core interactions.^c If the position is “open”, we continue with step 2. Otherwise, step 1 is repeated.

2. Assume that the chain construction has arrived at monomer i . We randomly choose a new bond $\mathbf{b}_i$ from monomer i to monomer $i + 1$, compute $P_{i+1}^{\text{open}}(\mathbf{x})$ for the trial position of $i + 1$, and decide whether it is open or not according to Eq. (59). If it is closed, we keep choosing new bonds up to a maximum number of k_{RG} trial directions.^d As soon as the first open direction is found, we proceed to step 3.

Otherwise, a recoil step is performed. That is, the chain moves back to monomer $i - 1$ where it searches for an open direction, using the $k_{\text{rest}} = k_{\text{RG}} - k_{\text{checked}}$ previously unchecked directions. The first open direction found is used to place (again) monomer i . If the chain fails to find an open direction from the k_{rest} remaining ones at monomer $i - 1$, it recoils to $i - 2$ and checks the previously unused directions for a possibility to grow from there. In difficult situations the chain keeps falling back until a maximum number of N_{recoil} recoil steps has been performed. If this number is exceeded, the construction of chain has to resume from step 1.

This shows that the RG algorithm is characterized by two parameters: k_{RG} and N_{recoil} . Tuning of these parameters is important to optimize the efficiency of the method.

3. If the construction has successfully placed the n th monomer, we attach monomer $n - N_{\text{recoil}} + 1$ permanently to the chain, as recoiling can only fall back to $n - N_{\text{recoil}} + 1$.

^cIn the SAW-limit, the decision of whether the position is open is not probabilistic, but deterministic. One just checks on overlaps with other monomers: If no overlap occurs, the position is open. For continuous potentials, however, the decision becomes probabilistic. We compare $P_1^{\text{open}}(\mathbf{x})$ to a random number ζ , uniformly distributed between 0 and 1. The position is open if $\zeta < P_1^{\text{open}}(\mathbf{x})$. Otherwise, it is closed. This means that we accept the position with probability $P_1^{\text{open}}(\mathbf{x})$.

^dIn practical implementations, the number of trial directions need not be the same for every monomer i . It may depend on i . For instance, one could choose for half of the monomers 2 trial directions and for the other half 3 directions so that $k_{\text{RG}} = 2.5$ on average. Thus, k_{RG} need not be integer.

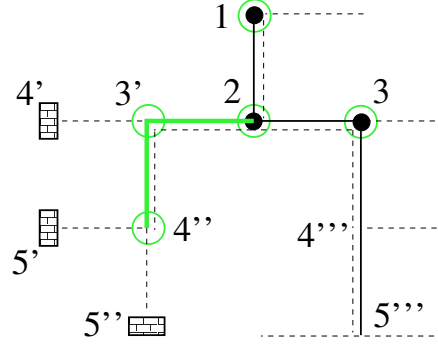


Figure 17. Illustration of the recoil-growth procedure for a SAW on a square lattice, using $k_{\text{RG}} = 2$ and $N_{\text{recoil}} = 3$. The first monomer is placed at an empty (“open”) lattice site. At each site that the growing chain visits, it has k_{RG} trial directions for the next step (indicated by the dashed lines: there are 2 for each monomer). These directions are chosen at random from the $z - 1$ forward lattice directions. Starting at monomer 1 the chain first tries the path 1 2 3'. At that point, the first monomer ($3 - N_{\text{recoil}} + 1$) becomes fixed (indicated by $\circ \rightarrow \bullet$ in the figure). Then, the chain attempts to grow to position 4' and finds it blocked. Thus, it recoils to monomer 3' and tries the remaining direction leading to monomer 4''. At that point, the monomer $4 - N_{\text{recoil}} + 1 = 2$ becomes fixed ($\circ \rightarrow \bullet$), as the “feeler” 2 3' 4'' has attained its full length $N_{\text{recoil}} = 3$ (grey thick solid line). However, the next two trial directions, $4' \rightarrow 5'$ and $4' \rightarrow 5''$, are found blocked. So, the feeler first recoils to 3', where it realizes that k_{RG} trial directions have been exhausted so that it must fall back its maximum length N_{recoil} to monomer 2. From there, it finds the open path 2 3 4''' 5''' so that monomer 3 becomes permanently attached to the chain. In this way, the construction goes on up to monomer $N - N_{\text{recoil}} + 1$. If this monomer was fixed, the construction stops, as a feeler to monomer N exists. Figure adapted from Ref. 104.

If this occurs and still no open direction is found at monomer $n - N_{\text{recoil}} + 1$, the growth process terminates and we must go back to step 1. Thus, the N_{recoil} monomers, $i = n - N_{\text{recoil}} + 1, \dots, n$, may be regarded as “retractable feeler” of maximum length N_{recoil} which probes the territory ahead of monomer $n - N_{\text{recoil}}$.

4. Repeat steps 2 and 3 until the monomer $N - N_{\text{recoil}} + 1$ was attached to the chain. This implies that the feeler attained the chain end. A complete chain of length N has thus been constructed.

Acceptance step: If chain a has been successfully constructed, we have to determine its weight $W^a(\mathbf{x})$ and that of the old chain b , $W^b(\mathbf{x}')$, which chain a attempts to replace. This may be done as follows:

1. In the construction step each monomer i disposes of (at most) k_{RG} trial directions into which the growth of a feeler of maximum length N_{recoil} may be attempted. Assume that we have checked $k = 1, \dots, k_{\text{checked}}$ of these directions and found that the k_{checked} th direction is the first allowing us to complete the construction of the feeler up to length N_{recoil} . We now test the remaining $k_{\text{rest}} = k_{\text{RG}} - k_{\text{checked}}$ directions to find out how many “feelers” of length N_{recoil} can be grown. (For the last monomers, $N - N_{\text{recoil}} + 1 < i \leq N$, the length of the feeler is shortened by one bond, as i advances step-by-step toward N .) For monomer i let $m_i(\mathbf{x})$ denote the number of directions where a feeler of full length can be grown. We have $1 \leq m_i(\mathbf{x}) \leq 1 + k_{\text{rest}}$.

As monomer i is only irrevocably added to the chain, if its position was initially open [Eq. (59)] and if at least one feeler of length N_{recoil} may be grown from it, the probability with which we propose to place monomer i is given by [see Eq. (55)]

$$P_{s,i}^a(\mathbf{x}) = P_i^{\text{open}}(\mathbf{x}) \cdot \frac{1}{m_i(\mathbf{x})} = \frac{1}{w_i^a(\mathbf{x})} . \quad (60)$$

With respect to CBMC $P_{s,i}^a(\mathbf{x})$ differs by the absence of the Boltzmann factor $\exp[-\beta u_i^a(\mathbf{b}_{i-1})]$, since only excluded-volume interactions are taken into account during the construction. So, the new chain a is proposed with probability

$$P_s^a(\mathbf{x}) = \prod_{i=1}^N \frac{1}{w_i^a(\mathbf{x})} = \frac{1}{W^a(\mathbf{x})} \quad (m_N(\mathbf{x}) = 1) . \quad (61)$$

2. The weight calculation of the old chain b has two ingredients. First, we calculate $P_i^{\text{open}}(\mathbf{x}')$ for each monomer i according to Eq. (59).^e Second, we attempt to grow feelers in $k_{\text{RG}} - 1$ randomly chosen directions and we count the number of feelers that attain length N_{recoil} . This number plus 1 yields $m_i(\mathbf{x}')$ (“plus 1”, since one feeler of length N_{recoil} already exists along the backbone of chain b). In analogy to chain a , we thus write

$$P_s^b(\mathbf{x}') = \prod_{i=1}^N \frac{P_i^{\text{open}}(\mathbf{x}')}{m_i(\mathbf{x}')} = \prod_{i=1}^N \frac{1}{w_i^b(\mathbf{x}')} = \frac{1}{W^b(\mathbf{x}')} \quad (m_N(\mathbf{x}') = 1) . \quad (62)$$

3. Finally, the total potential energy of chains a and b , $U^a(\mathbf{x})$ and $U^b(\mathbf{x}')$, must be computed before the new chain can be accepted with probability [see Eqs. (57,58)]

$$\text{acc}(\mathbf{x}' \rightarrow \mathbf{x}) = \min \left(1, \frac{W^a(\mathbf{x})}{W^b(\mathbf{x}')} e^{-\beta[U^a(\mathbf{x}) - U^b(\mathbf{x}')] } \right) . \quad (63)$$

The RG algorithm has two adjustable parameters: k_{RG} , the (average) number of trial directions of a monomer, and N_{recoil} , the length of the retractable feeler. Intuitively, one expects that N_{recoil} should be large, whereas k_{RG} should be small. The main advantage of the RG algorithm is its ability to avoid traps by probing the environment with a feeler. A short feeler will strongly reduce this ability and thus the rate of successful chain construction. On the other hand, the value of k_{RG} should not be too large because $1 \leq m_i(\mathbf{x}) \leq k_{\text{RG}}$. (Remember that $m_i(\mathbf{x})$ denotes the number of feelers of length N_{recoil} , which may be grown from monomer i .) A large k_{RG} allows for many different values of $m_i(\mathbf{x})$, as we go along the chain to calculate the weight $W^a(\mathbf{x})$. The spread in $m_i(\mathbf{x})$ leads to a wide distribution of $W^a(\mathbf{x})$, which will strongly deviate from the Boltzmann distribution (see Sec. 6.1).

^eFor the old chain we only calculate $P_i^{\text{open}}(\mathbf{x}')$. This is different from the construction step, where we decide whether the position is open or closed according to probability $P_i^{\text{open}}(\mathbf{x})$. That is, we first calculate $P_i^{\text{open}}(\mathbf{x})$ and then compare it to a random number ζ uniformly distributed between 0 and 1. If $\zeta < P_i^{\text{open}}(\mathbf{x})$, the position is open. Otherwise, it is closed [see Eq. (59)].

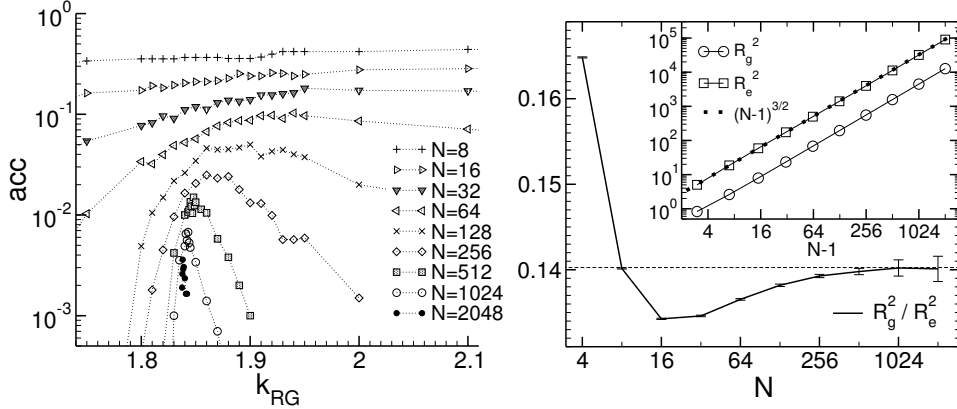


Figure 18. Left: Acceptance probability versus k_{RG} for $8 \leq N \leq 2048$ obtained from simulations of isolated 2D chains (“Kremer-Grest model”) via the RG algorithm. For all cases, $N_{\text{recoil}} = N/2$. Note the sharp peak of $\text{acc}(k_{RG})$ for long chains. The value of k_{RG} at the peak yields the optimum choice $k_{RG}^{\text{opt}}(N)$ for this chain length. For k_{RG}^{opt} we find: $\text{acc} \propto 1/N$. Right: Plot of R_g^2/R_e^2 versus N (main figure) and of R_g^2 and R_e^2 versus the number of bonds $N - 1$ (inset). The dashed horizontal line in the main figure represents the large- N limit of R_g^2/R_e^2 [Eq. (64)]. The dotted line in the inset shows the theoretically predicted power law $R_e^2 \propto R_g^2 \sim N^{2\nu}$ with $\nu = 3/4$.

Applications of the RG Algorithm. These expectations are confirmed by applications of the RG algorithm to 3D lattice¹⁰⁴ and 3D off-lattice¹⁰⁵ models of polymer solutions. For the chain lengths studied ($N \leq 100$) good choices are $2 \lesssim k_{RG} \lesssim 3$ and $3 \lesssim N_{\text{recoil}} \lesssim 10$, with the need to have larger values for both parameters if the density of the solution increases. Furthermore, it was observed that, for high densities and long chains, the RG algorithm may be an order of magnitude more efficient than CBMC.

Our preliminary results¹⁰⁹ on 2D polymer solutions, simulated with the Kremer-Grest model (see Sec. 4.2), appear to confirm the trends observed in 3D. Presumably due to the fact that the risk of trapping is more severe in 2D than in 3D,^{95,110} we found that, even for an isolated chain, chain lengths $N \gtrsim 100$ are very difficult to simulate via CBMC. Here, the RG algorithm provides a powerful alternative. While the performance of the algorithm depends only weakly on the choice of N_{recoil} , provided a long feeler is employed, e.g. $N_{\text{recoil}} = N/2$, k_{RG} must be optimized carefully. Figure 18 shows that the acceptance probability develops a pronounced peak for long chains. While for $N \lesssim 64$ the acceptance probability is fairly insensitive to the precise choice of k_{RG} , if $k_{RG} \approx 2$, it has to be adjusted to 4 significant digits for $N = 2048$. In this case, the optimum value is $k_{RG}^{\text{opt}} = 1.838$. This means that on average there are four monomers with two trial directions, followed by one monomer with only one trial direction.^f

To exemplify that the RG algorithm produces correct statistical properties Fig. 18 also shows the radius of gyration R_g^2 , the end-to-end distance R_e^2 , and the ratio R_g^2/R_e^2 . For isolated, two-dimensional chains one expects to find the critical exponent $\nu = 3/4$ ²² and

^fWe arrive at that conclusion in the following way. Suppose we allow the monomers to have either 1 or 2 trial directions. Let p denote the probability that a monomer has two trial directions. Then, we pose $k_{RG} = 2 \cdot p + 1 \cdot (1 - p)$ so that $p = 0.838$ with $k_{RG}^{\text{opt}} = 1.838$ for $N = 2048$. Thus, out of 5 monomers, roughly four monomers have two trial directions, leaving one trial direction for the remaining monomer.

the ratio¹¹¹

$$\lim_{N \rightarrow \infty} \frac{R_g^2}{R_e^2} = 0.14026 \pm 0.00011 . \quad (64)$$

The figure illustrates that both predictions are well borne out by the simulation data. Apparently, the 2D Kremer-Grest model quickly converges to the large- N limit, i.e., to $R_e^2 \propto R_g^2 \sim N^{3/2}$ and to Eq. (64). Deviations are only visible when investigating the ratio R_g^2/R_e^2 for $N \lesssim 512$. For longer chains corrections-to-scaling are small. In this respect, the Kremer-Grest model agrees with the results obtained for the SAW on a square lattice.¹¹¹

7 Conclusions

Behind the title of this chapter “Monte Carlo Simulation of Polymers: Coarse-Grained Models” a topic of a large breadth is hidden. So, a selection is necessary. We have employed several criteria in this selection.

First, we concentrate on simple generic models retaining only basic features of a polymer chain (chain connectivity, excluded volume, etc.; see Sec. 4). Coarse-grained models derived from specific polymers are only touched upon briefly (Appendix 7), although this is an important current research topic.^{41,42}

Second, we define a generic model as one consisting of monomers with the simplest imaginable structure. They are identified with sites on a lattice or with Lennard-Jones spheres in the continuum. The monomers in the chain are all the same and uncharged (“neutral homopolymers”). Their interaction is either purely repulsive or consists of a short-range repulsion supplemented by an attractive potential at larger distances (see Sec. 4). Thus, nor did we consider specific interactions, such as electrostatic interactions, H-bonds, interactions between different species of monomers (e.g. binary mixtures, block-copolymers), etc. –this will be done in other chapters of this book– neither did we discuss current coarse-graining approaches which do not model a chain as a concatenation of monomers, but represent the whole polymer as a soft ellipsoidal^{112,113} or spherical¹¹⁴ particle.

Within the scope of these generic models we presented various algorithms. What is the upshot of this discussion for applications? Here are some suggestions:

- Dilute solutions:
 - For isolated chains, exempt of strong monomer-monomer interactions, the pivot algorithm is currently considered as being the most efficient method to study global properties related to the chain extension (R_g , R_e , the exponent ν , etc.). To study properties related to the partition function (e.g. the exponent γ) other algorithms are better suited (see footnote on page 112).
 - To initialize the simulation via the pivot algorithm or for the analysis of local properties the slithering-snake algorithm represents an attractive alternative (see discussion on page 113).
 - If attractive monomer-monomer interactions are present, as it is the case close to the Θ -point, it appears as if the pruned-enriched Rosenbluth method (PERM) is currently the most efficient algorithm (Sec. 6.2).

- From dilute solutions to dense melts:
 - As the density of the solution increases, the large-scale pivot moves become quickly inefficient. Here, either bilocal updating schemes, such as the slithering-snake algorithm, or non-local chain reconstructions via configurational-bias Monte Carlo (CBMC) are better suited. CBMC has become a well established tool, particularly in its grand-canonical formulation for the study of phase diagrams. It transpires that its range of applicability may be extended by the recently proposed recoil-growth (RG) algorithm (see Sec. 6.3). When the solution becomes more and more concentrated, the probability to renew the configuration of large chain portions becomes small. In this limit, the CBMC and RG methods reduce to the slithering-snake algorithm.
 - In dense melts consisting of long chain ($N \gtrsim 10^3$) also the slithering-snake algorithm faces serious problems to equilibrate the system (Sec. 5.2). Here, it is better to use connectivity-altering moves instead of attempting to regrow (parts of) a chain. An example is the recently proposed double-pivot algorithm which appears to be very efficient in equilibrating dense melts (see Sec. 5.3).

The previous points only refer to the study of conformational and structural properties of polymer systems as well as to an efficient equilibration of the system. If the focus of interest are dynamic properties, local moves should be employed because they mimic best the physical dynamics (if solvent-mediated hydrodynamic interactions are absent; see Ref. 14).

The suggestions made above are a result of the discussion given in this chapter. Certainly, our discussion suffers from omissions. Personally, we feel that the most serious one are generalized ensemble techniques,^{115,116} such as simulated or parallel tempering (see Chap. 14.1 of Ref. 31 for an introduction). In the context of polymer physics, these methods have been applied e.g. to determine the chemical potential of polymer chains (simulated tempering)¹¹⁷, to accelerate the equilibration of dense polymer melts (parallel tempering)¹¹⁸ or to the simulation of phase transitions¹¹⁹.

We apologize for this review-like end of our report, but hope that its content, together with the bibliography, will be helpful for those interested in MC simulations of polymer models.

Appendix

Realistic Models and Coarse-Graining of Real Polymers

The presentation of this chapter concentrated on generic coarse-grained models (Sec. 4). They are particularly useful for the study of the large-scale behavior of polymers. For these generic models several simulation algorithms were discussed. If one is now interested in properties of a specific molecule or material, the presented methods can still be applied. But their efficiency must be tested for the particular application.⁷⁵ The simulation of real materials must use models which reflect the underlying chemical structure of the polymer, even at the coarse-grained level. So, some mapping between a detailed, chemically realistic and a coarse-grained model must be established.^{41,42,120} The following paragraphs explain

the need and the main steps of such a coarse-graining procedure and briefly describe an application to modeling a melt of poly(vinyl alcohol).

From Realistic to Coarse-Grained Models. An ideal simulation approach would consist in studying systems of long chains ($N \gtrsim 10^3$) with potentials being calculated from the simultaneous motion of all nuclei and electrons (via the Car-Parrinello method¹²¹). However, already for an isolated chain such an approach is not feasible due to the large spread of subatomic and molecular time scales.^{4,5} For $N \sim 10^3$ the relaxation time of a polymer in dilute solution is about 1 μ s, whereas the inclusion of the electrons requires a time step of $\sim 10^{-17}$ s in a MD simulation. This disparity of 11 orders of magnitude cannot be covered in present day simulations. A further restriction results from the system sizes that may be simulated. Typically, the total number of particles (electrons and nuclei) is limited to $\sim 10^3$.

Thus, simplifications are necessary.⁹ There are several levels to this.¹²⁰ In a first step, the “on the fly” calculated quantum-mechanical potentials can be replaced by empirical potentials for the bond lengths, the bond angles, the torsional angles and the non-bonded interactions between distant monomers along the backbone of the chain (“quantum level $\rightarrow$ atomistic level”). After careful optimization of the parameters of these potentials quantitative agreement between simulation and experimental results can be obtained.^{41,15} If the system is to stay in thermal equilibrium on both the local monomeric scale and the global scale of the chain, such comparisons between simulation and experiments are limited to high temperatures up to now. Extensions to low temperatures where crystallization or sluggish glass-like relaxation occur still represent a great challenge for simulations on atomistic scale.^h To tackle these problems the models must be further simplified. In a second step of simplification, fast degrees of freedom (bond length and bond angle vibrations, etc.) may therefore be eliminated by a coarse-graining procedure (“atomistic level $\rightarrow$ coarse-grained level”). This procedure leads either to generic models, as those which were discussed in the previous sections, or to a coarse-grained model for a specific polymer (see Sec. 4). Recent approaches to the latter case have been reviewed in Refs. 41, 42, 120.

The first step, from the quantum to the atomistic level, may be interpreted as an ab-initio approach or a bottom-up construction of the model. When pushing the simplifications of the model beyond the atomistic to the coarse-grained level it becomes more and more apparent that the empirical potentials are just simulation parameters. So, instead of the bottom-up construction, one may also choose a more pragmatic top-down procedure. That is, given some (measured) properties, what are the potentials which can reproduce them?ⁱ In practice, one often proceeds as follows: By the bottom-up method a first guess of coarse-grained potentials is constructed. In a second step, these potentials are optimized with respect to some experimentally or theoretically known properties. With the final potentials

⁹Here, we assume that the whole chain is simulated at the same level of modeling. Of course, it is also possible to treat some part of the chain in atomistic detail (e.g. via Car-Parrinello), whereas others are modeled on a coarse-grained level. A recent example of such a multiscale approach is the selective adsorption of polycarbonate on a nickel surface.¹²²

^hAn important contribution to this field are the “amorphous cell” simulations pioneered by Theodorou and Suter^{123,124}. Here, the idea is to fold an infinitely long chain in the simulation box such that the monomer density is close to the experimental one and the resulting chain structure is reasonable. With this approach mechanical properties of the amorphous, glassy melt have been studied successfully.

ⁱOne such method for the solution of the inverse problem, known as “reverse Monte Carlo”, is actually used to interpret neutron and x-ray scattering data.^{125,126}

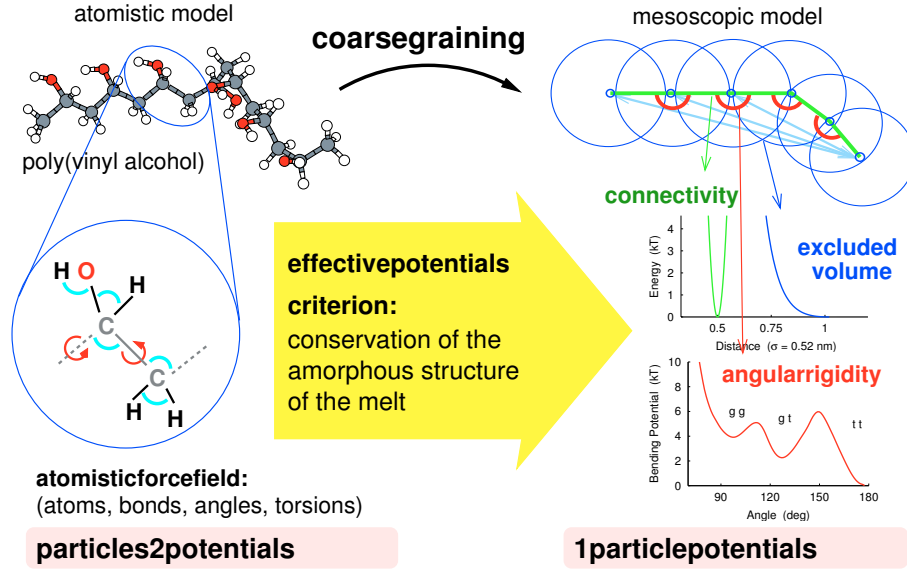


Figure 19. Schematic illustration of the coarse-graining procedure developed in Ref. 127. Starting from an atomistically detailed model of poly(vinyl alcohol) a bead-spring model is constructed. A monomer of the atomistic model is represented by a sphere on the coarse-grained level. The spheres interact by a harmonic bond potential and a purely repulsive LJ-like potential [Eq. (65)]. The parameters of these potentials are optimized in the following way. In an atomistic simulation at high temperature, the structure of a melt of poly(vinyl alcohol) and the local conformation of a polymer are determined: the structure by the pair-distribution function $g(r)$ ⁵⁹ and the conformation by the distribution of the angle between two vectors formed by connecting the C-atom of the CHOH-group of the $(i - 1)$ th monomer with that of the i th monomer (first vector) and that of the i th monomer with that of the $(i + 1)$ th monomer (second vector). From the angular distribution a potential of mean force⁵⁹ is calculated. In this case, it could directly be used as the bending potential for the bond angle in the coarse-grained model. Contrary to that, the parameters of the LJ-potentials are iteratively optimized until $g(r)$ of the poly(vinyl alcohol) melt coincides with that of the coarse-grained model.

one may then (try to) predict quantities which are not easily measurable. In the following we will sketch one such pragmatic approach.

An Example of How to Coarse-Grain Real Polymers. The idea and the realization of the coarse-graining are illustrated by the example of poly(vinyl alcohol) in Fig. 7.¹²⁷

On the coarse-grained level, the chemical structure of the monomers is suppressed. They are represented by spheres, bound to each other by a harmonic potential. The orientation of successive bonds along the chain is determined by a bond angle potential, whereas other non-bonded monomers interact by the repulsive part of a LJ-like potential

$$U_{\text{LJ}}^{\text{rep}}(r) = \begin{cases} \epsilon_n(\sigma/r)^n - \epsilon_6(\sigma/r)^6 + \epsilon_0 & \text{for } r \leq r_{\text{min}} , \\ 0 & \text{else ,} \end{cases} \quad (65)$$

where r_{min} is the minimum of the potential and ϵ_0 is chosen such that the minimum is shifted to zero to have a continuous force. When fitting the potential to real polymers it

turned out⁶⁷ that for the repulsive term, an exponent smaller than 12 may be better adapted because the coarse-grained units have to be somewhat softer than atoms.^j By comparing high-temperature simulations of an atomistically detailed and a coarse-grained model, the potential parameters of the latter are adjusted in such a way that it faithfully reproduces the backbone rigidity of poly(vinyl alcohol) and the local packing of its monomers in the liquid state.

The construction of the model was done at one state point. It is a priori not guaranteed how far from this reference conditions the model still matches the behavior of the original one. It turns out that in the present case, the crystallization at lower temperatures is well reproduced.¹²⁹

Acknowledgments

The results reported here, as far as it concerns our own work, were obtained in fruitful collaboration with M. Aichele, J.-L. Barrat, O. Biermann, K. Binder, R. Faller, B. Lobe, E. Luitjen, M. Müller, F. Müller-Plathe, L. Mattioni, W. Paul, D. Reith, and M. Wolfgardt. The authors' research is supported in part by the ESF Supernet programme and by the LEA "Macromolecules in nanoscopically structured media". J.B. is indebted to the IUF for financial support. Furthermore, we gratefully acknowledge generous grants of computing time by the IDRIS (Orsay, France).

References

1. T. P. Lodge and M. Muthukumar, *Physical Chemistry of Polymers: Entropy, Interactions, and Dynamics*, J. Phys. Chem. **100**, 13275 (1996).
2. H. M. Jaeger, S. R. Nagel, and R. P. Behringer, *Granular solids, liquids, and gases*, Rev. Mod. Phys. **68**, 1259 (1996).
3. R. A. L. Jones, *Soft Condensed Matter* (Oxford University Press, New York, 2002).
4. K. Binder, *Introduction: General Aspects of Computer Simulation Techniques and their Applications in Polymer Science*. In: *Monte Carlo and Molecular Dynamics Simulations in Polymer Science* (Oxford University Press, New York, 1995), pp. 3–46.
5. K. Kremer, *Computer Simulations for Macromolecular Science*, Macromol. Chem. Phys. **204**, 257 (2003).
6. P. G. de Gennes, *Scaling Concepts in Polymer Physics* (Cornell University Press, Ithaca, 1996).
7. M. Doi and S. F. Edwards, *The Theory of Polymer Dynamics* (Oxford University Press, Oxford, 1986).
8. K. Binder and W. Paul, *Monte Carlo Simulations of Polymer Dynamics: Recent Advances*, J. Poly. Sci. B **35**, 1 (1997).
9. T. C. B. McLeish, *Tube theory of entangled polymer dynamics*, Adv. Phys. **51**, 1379 (2002).

^jBy doing more radical coarse-graining steps lumping many monomers¹²⁸ or even whole chains into one particle¹¹⁴, one will end up with soft overlapping potentials, as used e.g. in "Dissipative Particle Dynamics" (see Ref. 14).

10. M. Allen, *Introduction to molecular dynamics simulation*. In: *Computational Soft Matter: From Synthetic Polymers to Proteins* (NIC Series, 2004), pp. 1–28.
11. M. Müller, *Phase behavior and chain conformations in polymer blends: Monte Carlo simulation vs mean field theory*. In: *Computational Soft Matter: From Synthetic Polymers to Proteins* (NIC Series, 2004), pp. 237–274.
12. C. Holm, *Efficient methods for long range interactions in periodic geometries plus one application*. In: *Computational Soft Matter: From Synthetic Polymers to Proteins* (NIC Series, 2004), pp. 195–236.
13. A. Mark and S.-J. Marrink, *Self-organization of biomolecular systems: Simulating the folding and aggregation of peptides, proteins and lipids*. In: *Computational Soft Matter: From Synthetic Polymers to Proteins* (NIC Series, 2004), pp. 321–322.
14. B. Dünweg, *Advanced simulations for hydrodynamic problems: Lattice Boltzmann and dissipative particle dynamics*. In: *Computational Soft Matter: From Synthetic Polymers to Proteins* (NIC Series, 2004), pp. 61–82.
15. G. D. Smith, W. Paul, M. Monkenbusch, and D. Richter, *A comparison of neutron scattering studies and computer simulations of polymer melts*, Chem. Phys. **261**, 61 (2000).
16. L. Schäfer, *Excluded Volume Effects in Polymer Solutions* (Springer, Berlin–Heidelberg, 1999).
17. P. J. Flory, *Statistical Mechanics of Chain Molecules* (Wiley, New York, 1969).
18. J.-Ph. Bouchaud and A. Georges, *Anomalous Diffusion in Disordered Media: Statistical Mechanisms, Models and Physical Applications*, Phys. Rep. **195**, 127 (1990).
19. M. E. Fisher, *The renormalization group in the theory of critical behavior*, Rev. Mod. Phys. **46**, 597 (1974).
20. C. Domb, *The Critical Point* (Taylor & Francis, London, 1996).
21. S. Caracciolo, M. S. Causo, and A. Pelissetto, *High-precision determination of the critical exponent γ for self-avoiding walks*, Phys. Rev. E **57**, R1215 (1998).
22. A. Pelissetto and E. Vicari, *Critical phenomena and renormalization-group theory*, Phys. Rep. **368**, 549 (2002).
23. A. Y. Grosberg and A. R. Khokhlov, *Statistical Physics of Macromolecules* (AIP Press, New York, 1994).
24. K. Binder, J. Baschnagel, and W. Paul, *Glass Transition of Polymer Melts: Test of Theoretical Concepts by Computer Simulation*, Progr. Poly. Sci. **28**, 115 (2003).
25. J.-U. Sommer and G. Reiter (Eds.), *Polymer Crystallization: Observations, Concepts and Interpretations*, Lecture Notes in Physics **606** (Springer, Berlin–Heidelberg, 2003).
26. S. F. Edwards, *The size of a polymer molecule in a strong solution*, J. Phys. A: Math. Gen. **8**, 1670 (1975).
27. W. Paul, K. Binder, D. W. Heermann, and K. Kremer, *Crossover scaling in semidilute polymer solutions: a Monte Carlo test*, J. Phys. II **1**, 37 (1991); *Dynamics of polymer solutions and melts. Reptation predictions and scaling of relaxation times*, J. Chem. Phys. **95**, 7726 (1991).
28. S.-H. Chong and M. Fuchs, *Mode-Coupling Theory for Structural and Conformational Dynamics of Polymer Melts*, Phys. Rev. Lett. **88**, 185702 (2002).
29. T. P. Lodge, N. A. Rotstein, and S. Prager, *Dynamics of Entangled Polymer Liquids: Do Linear Chains Reptate?* Adv. Chem. Phys. **79**, 1 (1990).

30. H. Tao, T. P. Lodge, and E. D. von Meerwall, *Diffusivity and Viscosity of Concentrated Hydrogenated Polybutadiene Solutions*, *Macromolecules* **33**, 1747 (2000).
31. D. Frenkel and B. Smit, *Understanding Molecular Simulation – From Algorithms to Applications* (Academic Press, San Diego, 2nd ed. 2002).
32. K. Binder and D. W. Heermann, *Monte Carlo Simulation in Statistical Physics* (Springer, Berlin–Heidelberg, 2002).
33. D. P. Landau and K. Binder, *A Guide to Monte Carlo Simulations in Statistical Physics* (Cambridge University Press, Cambridge, 2000).
34. A. D. Sokal, *Monte Carlo Methods for the Self-Avoiding Walk*. In: *Monte Carlo and Molecular Dynamics Simulations in Polymer Science* (Oxford University Press, New York, 1995), pp. 47–124.
35. A. D. Sokal, *Monte Carlo Methods in Statistical Mechanics: Foundations and New Algorithms*. In: *Quantum Fields on the Computer* (World Scientific, Singapore, 1992).
36. D. Frenkel, *Introduction to Monte Carlo methods*. In: *Computational Soft Matter: From Synthetic Polymers to Proteins* (NIC Series, 2004), pp. 29–60.
37. N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. N. Teller, and E. Teller, *Equation of state calculations by fast computing machines*, *J. Chem. Phys.* **21**, 1087 (1953).
38. O. Narayan and A. P. Young, *Convergence of Monte Carlo simulations to equilibrium*, *Phys. Rev. E* **64**, 021104 (2001).
39. W. Feller, *An Introduction to Probability Theory and its Applications*, Vol. 1 and Vol. 2 (Wiley, New York, 1966).
40. V. I. Manousiouthakis and M. W. Deem, *Strict detailed balance is unnecessary in Monte Carlo simulation*, *J. Chem. Phys.* **110**, 2753 (1999).
41. F. Müller-Plathe, *Coarse-Graining in Polymer Simulation: From the Atomistic to the Mesoscopic Scale and Back*, *ChemPhysChem* **3**, 754 (2002).
42. J. Baschnagel, K. Binder, P. Doruker, A. A. Gusev, O. Hahn, K. Kremer, W. L. Matice, F. Müller-Plathe, M. Murat, W. Paul, S. Santos, U. W. Suter, V. Tries, *Bridging the Gap Between Atomistic and Coarse-Grained Models of Polymers: Status and Perspectives*, *Adv. Polym. Sci.* **152**, 41 (2000).
43. K. Kremer and K. Binder, *Monte Carlo Simulations of Lattice Models for Macromolecules*, *Comp. Phys. Rep.* **7**, 259 (1988).
44. K. Binder, *Monte Carlo and Molecular Dynamics Simulations of Amorphous Polymers*. In: *Computational Modeling of Polymers* (Marcel Dekker, New York, 1992), pp. 221–295.
45. H.-P. Wittmann and K. Kremer, *Vectorized version of the bond fluctuation method for lattice polymers*, *Comp. Phys. Commun.* **61**, 309 (1990); *Erratum notice*, *ibid.* **71**, 343 (1992).
46. W. J. C. Orr, *Trans. Faraday Soc.* **43**, 12 (1947); E. W. Montroll, *Markoff Chains and Excluded Volume Effects in Polymer Chains*, *J. Chem. Phys.* **18**, 734 (1950).
47. P. Grassberger, *Pruned-enriched Rosenbluth method: Simulations of θ polymers of chain length up to 1 000 000*, *Phys. Rev. E* **56**, 3682 (1997).
48. P. H. Verdier and W. H. Stockmayer, *Monte Carlo Calculations on the Dynamics of Polymers in Dilute Solution*, *J. Chem. Phys.* **36**, 227 (1962).
49. N. Madras and A. D. Sokal, *Nonergodicity of local, length-conserving Monte Carlo*

- algorithms for the self-avoiding walk*, J. Stat. Phys. **47**, 573 (1987).
50. S. Caracciolo, M. S. Causo, G. Ferraro, M. Papinutto, and A. Pelissetto, *Bilocal Dynamics for Self-Avoiding Walks*, J. Stat. Phys. **100**, 1111 (2000).
 51. S. Caracciolo, M. Papinutto, and A. Pelissetto, *Dynamic critical behavior of an extended reptation dynamics for self-avoiding walks*, Phys. Rev. E **65**, 031106 (2002).
 52. I. Carmesin and K. Kremer, *The Bond Fluctuation Method: A New Effective Algorithm for the Dynamics of Polymers in All Spatial Dimensions*, Macromolecules **21**, 2819 (1988).
 53. I. Carmesin and K. Kremer, *Static and dynamic properties of two-dimensional polymer melts*, J. Physique **51**, 915 (1990).
 54. H.-P. Deutsch and K. Binder, *Interdiffusion and self-diffusion in polymer mixtures: a Monte Carlo study*, J. Chem. Phys. **94**, 2294 (1991).
 55. B. Lobe and J. Baschnagel, *Monte Carlo simulation of the glass transition in two- and three-dimensional polymer melts: Influence of the spatial dimension*, J. Chem. Phys. **101**, 1616 (1994).
 56. A. Z. Panagiotopoulos, *On the equivalence of continuum and lattice models for fluids*, J. Chem. Phys. **112**, 7132 (2000).
 57. H.-P. Deutsch and R. Dickman, *Equation of state for athermal lattice chains in a 3d fluctuating bond model*, J. Chem. Phys. **93**, 8983 (1990).
 58. M. Wittkop, S. Kreitmeier, and D. Göritz, *Monte Carlo simulations of a single polymer chain under extension above and below the Θ temperature*, Phys. Rev. E **53**, 838 (1996).
 59. J.-P. Hansen and I. R. McDonald, *Theory of Simple Liquids* (Academic Press, London, 1986).
 60. N. B. Wilding, M. Müller, and K. Binder, *Chain length dependence of the polymer-solvent critical point parameters*, J. Chem. Phys. **105**, 802 (1996).
 61. G. S. Grest and K. Kremer, *Molecular dynamics simulation of polymers in the presence of a heat bath*, Phys. Rev. A **33**, 3628 (1986); K. Kremer and G. S. Grest, *Dynamics of entangled linear polymer melts: A molecular-dynamics simulation*, J. Chem. Phys. **92**, 5057 (1990).
 62. M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids* (Oxford University Press, Oxford, 1994).
 63. B. Smit, *Phase diagram of Lennard-Jones fluids*, J. Chem. Phys. **96**, 8639 (1992).
 64. C. W. Yong, J. H. R. Clarke, J. J. Freire, and M. Bishop, *The theta condition for linear polymer chains in continuous space and three dimensions*, J. Chem. Phys. **105**, 9666 (1996).
 65. C. Bennemann, W. Paul, K. Binder, and B. Dünweg, *Molecular-dynamics simulations of the thermal glass transition in polymer melts: α -relaxation behavior*, Phys. Rev. E **57**, 843 (1998).
 66. M. Müller and L. Gonzalez MacDowell, *Interface and Surface Properties of Short Polymers in Solution: Monte Carlo Simulations and Self-Consistent Field Theory*, Macromolecules **33**, 3902 (2000).
 67. H. Meyer, O. Biermann, R. Faller, D. Reith, and F. Müller-Plathe, *Coarse graining of nonbonded inter-particle potentials using automatic simplex optimization to fit structural properties*, J. Chem. Phys. **113**, 6264 (2000).
 68. G. S. Grest and M. Murat, *Computer Simulations of Tethered Chains*. In: *Monte Carlo*

and *Molecular Dynamics Simulations in Polymer Science* (Oxford University Press, New York, 1995), pp. 476–578.

69. C. F. Abrams and K. Kremer, *The effect of bond length on the structure of dense bead-spring polymer melts*, J. Chem. Phys. **115**, 2776 (2001); *Effects of excluded volume and bond length on the dynamics of dense bead-spring polymer melts*, *ibid.* **116**, 3162 (2002).
70. L. Mattioni *et al.*, in preparation.
71. H. E. A. Huitema and J. P. van der Eerden, *Can Monte Carlo simulation describe dynamics? A test on Lennard-Jones systems*, J. Chem. Phys. **110**, 3267 (1999).
72. W. Paul, *Anomalous diffusion in polymer melts*, Chem. Phys. **284**, 59 (2001).
73. See e.g. <http://hostingworks.com/support/dict.phtml?foldoc=complexity>.
74. F. T. Wall and F. Mandel, *Macromolecular dimensions obtained by an efficient Monte Carlo method without sample attrition*, J. Chem. Phys. **63**, 4592 (1975); F. T. Mandel, *Macromolecular dimensions obtained by an efficient Monte Carlo method: The mean square end-to-end separation*, J. Chem. Phys. **70**, 3984 (1979).
75. E. Leontidis, B. M. Forrest, A. H. Widmann, and U. W. Suter, *Monte Carlo Algorithms for the Atomistic Simulation of Condensed Polymer Phases*, Faraday Trans. **91**, 2355 (1995).
76. J. Reiter, *Monte Carlo Simulations of Linear and Cyclic Chains on Cubic and Quadratic Lattices*, Macromolecules **23**, 3811 (1990).
77. J. Reiter and E. R. Duering, *Monte Carlo simulation of freely jointed chains with variable bond lengths*, Macromol. Theory Simul. **4**, 667 (1995).
78. J. Reiter, T. Edling, and T. Pakula, *Monte Carlo simulation of lattice models for macromolecules at high densities*, J. Chem. Phys. **93**, 837 (1990).
79. J. M. Deutsch, *The dynamics of entangled polymers*, J. Physique **48**, 141 (1987).
80. L. Mattioni, J. P. Wittmer, J. Baschnagel, J.-L. Barrat, E. Luijten, *Dynamical Properties of the Slithering Snake Algorithm: A numerical test of the activated reptation hypothesis*, Eur. Phys. J. E **10**, 369 (2003).
81. H. Weber, *Monte-Carlo-Simulationen der Gasdiffusion in Polymermatrizen*, Dissertation (Universität Mainz, 1997), unpublished.
82. A. N. Semenov, *The Effect of Polymer Density Fluctuations on the Dynamics of High Molecular-Weight Polymers*, In: *Theoretical Challenges in the Dynamics of Complex Fluids* (Kluwer, Dordrecht, 1997), pp. 77–86.
83. A. N. Semenov and M. Rubinstein, *Dynamics of strongly entangled polymer systems: activated reptation*, Eur. Phys. J. B **1**, 47 (1998).
84. M. Wolfgang, J. Baschnagel, and K. Binder, *Monte-Carlo Simulation of 3-Dimensional Glassy Polymer Melts: Reptation versus Single Monomer Dynamics*, J. Phys. II **5**, 1035 (1995).
85. V. Tries, W. Paul, J. Baschnagel, and K. Binder, *Modeling polyethylene with the bond fluctuation model*, J. Chem. Phys. **106**, 738 (1997).
86. S. W. Sides, J. Curro, G. S. Grest, M. J. Stevens, T. Soddemann, A. Habenschuss, and J. D. Londono, *Structure of Poly(dimethylsiloxane) Melts: Theory, Simulation, and Experiment*, Macromolecules **35**, 6455 (2002); J. G. Curro, J. D. Weinhold, J. J. Rajasekaran, A. Habenschuss, J. D. Londono, and J. D. Honeycutt, *Intermolecular Packing in Stereoregular Polypropylene Liquids: Comparison between Theory and*

- X-ray Scattering Experiments*, *Macromolecules* **30**, 6264 (1997).
87. J. G. Curro, E. B. Webb III, G. S. Grest, J. D. Weinhold, M. Pütz, and J. D. McCoy, *Comparisons between integral equation theory and molecular dynamics simulations for realistic models of polyethylene liquids*, *J. Chem. Phys.* **111**, 9073 (1999).
 88. W. Janke, *Monte Carlo Simulations of Spin Systems*. In: *Computational Physics* (Springer, Berlin-Heidelberg, 1996), pp. 11–43.
 89. N. Madras and A. D. Sokal, *The pivot algorithm: a highly efficient Monte Carlo method for the self-avoiding walk*, *J. Stat. Phys.* **50**, 109 (1988).
 90. T. Kennedy, *A Faster Implementation of the Pivot Algorithm for Self-Avoiding Walks*, *J. Stat. Phys.* **106**, 407 (2002).
 91. B. Li, N. Madras, and A. D. Sokal, *Critical Exponents, Hyperscaling, and Universal Amplitude Ratios for Two- and Three-dimensional Self-Avoiding Walks*, *J. Stat. Phys.* **80**, 661 (1995).
 92. P. Grassberger, P. Sutter, and L. Schäfer, *Field theoretic and Monte Carlo analysis of the Domb-Joyce model*, *J. Phys. A* **30**, 7039 (1997).
 93. R. Auhl, R. Everaers, G. S. Grest, K. Kremer, and S. J. Plimpton, *Equilibration of Long Chain Polymer Melts in Computer Simulations*, preprint (cond-mat/0306026v1).
 94. N. Ch. Karayiannis, A. E. Giannousaki, V. G. Mavrantzas, and D. N. Theodorou, *Atomistic Monte Carlo simulation of strictly monodisperse long polyethylene melts through a generalized chain bridging algorithm*, *J. Chem. Phys.* **117**, 5465 (2002); N. Ch. Karayiannis, V. G. Mavrantzas, and D. N. Theodorou, *A Novel Monte Carlo Scheme for the Rapid Equilibration of Atomistic Model Polymer Systems of Precisely Defined Molecular Architecture*, *Phys. Rev. Lett.* **88**, 105503 (2002).
 95. J. Batoulis and K. Kremer, *Statistical properties of biased sampling methods for long polymer chains*, *J. Phys. A: Math. Gen.* **21**, 127 (1988).
 96. M. N. Rosenbluth and A. W. Rosenbluth, *Monte Carlo simulations of the average extension of molecular chains*, *J. Chem. Phys.* **23**, 356 (1955).
 97. C. M. Guttman, *Relationship of the Unweighted Rosenbluth and Rosenbluth Walk to a Polymer Chain at the Θ Point*, *Macromolecules* **19**, 833 (1986).
 98. M. Wittkop, S. Kreitmeier, and D. Göritz, *The collapse transition of a single polymer chain in two and three dimensions: A Monte Carlo study*, *J. Chem. Phys.* **104**, 3373 (1996).
 99. A. M. Rubio, J. J. Freire, C. W. Yong, and J. H. R. Clarke, *Conformation properties of polymer chains in the theta region*, *J. Chem. Phys.* **111**, 1302 (1999).
 100. P. Grassberger, *Go with the winners: a general Monte Carlo strategy*, *Comp. Phys. Comm.* **147**, 64 (2002).
 101. J. Hager and L. Schäfer, *Θ -point behavior of diluted polymer solutions: Can one observe the universal logarithmic corrections predicted by field theory?*, *Phys. Rev. E* **60**, 2071 (1999).
 102. M. S. Causo, B. Coluzzi, and P. Grassberger, *Simple model for the DNA denaturation transition*, *Phys. Rev. E* **62**, 3958 (2000).
 103. H. Frauenkron and P. Grassberger, *Critical unmixing of polymer solutions*, *J. Chem. Phys.* **107**, 9599 (1997).
 104. S. Consta, N. B. Wilding, D. Frenkel, and Z. Alexandrowicz, *Recoil growth: An efficient simulation method for multi-polymer systems*, *J. Chem. Phys.* **110**, 3220 (1999).
 105. S. Consta, T. J. H. Vlucht, J. W. Hoeth, B. Smit, and D. Frenkel, *Recoil growth algo-*

- rithm for chain molecules with continuous interactions*, Mol. Phys. **97**, 1243 (1999).
106. J. J. de Pablo, Q. Yan, and F. A. Escobedo, *Simulation of Phase Transitions in Fluids*, Annu. Rev. Phys. Chem. **50**, 377 (1999).
 107. M. Müller, K. Binder, and E. V. Albano, *Phase equilibria in thin polymer films*, Int. J. Mod. Phys. B **15**, 1867 (2001).
 108. See e.g. H. Meirovitch, *On the simulation of the entropy of macromolecules with different flexibilities*, J. Chem. Phys. **114**, 3859 (2001); *Entropy, pressure, and chemical potential of multiple chain systems from computer simulation. I. Application of the scanning method*, ibid. **97**, 5803 (1992).
 109. M. Aichele *et al.*, in preparation; M. Aichele, *Simulation Studies of Correlation Functions and Relaxation in Polymeric Systems*, Dissertation (Universität Mainz, 2003), unpublished.
 110. P. C. Hemmer and S. Hemmer, *Trapping of genuine self-avoiding walks*, Phys. Rev. A **34**, 3304 (1986).
 111. S. Caracciolo, A. Pelissetto, and A. D. Sokal, *Universal distance ratios for two-dimensional self-avoiding walks: corrected conformal-invariance predictions*, J. Phys. A: Math. Gen. **23**, L969 (1990).
 112. M. Murat and K. Kremer, *From many monomers to many polymers: Soft ellipsoid model for polymer melts and mixtures*, J. Chem. Phys. **108**, 4340 (1998).
 113. F. Eurich and P. Maass, *Soft ellipsoid model for Gaussian polymer chains*, J. Chem. Phys. **114**, 7655 (2001).
 114. A. A. Louis, P. G. Bolhuis, R. Finken, V. Krakoviack, E. J. Meijer, and J. P. Hansen, *Coarse-graining polymers as soft colloids*, Physica A **306**, 251 (2002).
 115. B. A. Berg, *Generalized ensemble simulations for complex systems*, Comp. Phys. Comm. **147**, 52 (2002).
 116. A. D. Bruce and N. B. Wilding, *Computational strategies for mapping equilibrium phase diagrams*, preprint (cond-mat/0210457v1).
 117. N. B. Wilding and M. Müller, *Accurate measurements of the chemical potential of polymer systems by Monte Carlo simulation*, J. Chem. Phys. **101**, 4324 (1994).
 118. A. Bunker and B. Dünweg, *Parallel excluded volume tempering for polymer melts*, Phys. Rev. E **63**, 016701 (2000).
 119. Q. Yan and J. J. de Pablo, *Hyperparallel tempering Monte Carlo simulation of polymeric systems*, J. Chem. Phys. **113**, 1276 (2000).
 120. F. Müller-Plathe, *Scale Hopping in Computer Simulations of Polymers*, Soft Materials **1**, 1 (2003).
 121. G. Galli and M. Pasquarello, *First-Principles Molecular Dynamics*. In: *Computer Simulation in Chemical Physics* (Dordrecht, Kluwer, 1993), pp. 261–313; D. Marx and J. Hutter, *Ab initio Molecular Dynamics: Theory and Implementation*. In: *Modern Methods and Algorithms of Quantum Chemistry* (NIC Series Vol. **3**, 2000), pp. 329–477 (available from <http://www.fz-juelich.de/nic-series>).
 122. L. Delle Site, C. F. Abrams, A. Alavi, and K. Kremer, *Polymers near metal surfaces: Selective adsorption and global conformations*, Phys. Rev. Lett. **89**, 156103 (2002).
 123. D. N. Theodorou and U. W. Suter, *Detailed Molecular Structure of a Vinyl Polymer Glass*, Macromolecules **18**, 1467 (1985).
 124. T. A. Weber and E. Helfand, *Molecular dynamics simulation of polymers. I. Structure*, J. Chem. Phys. **71**, 4760 (1979).

125. R. L. McGreevy and L. Pusztai, *Reverse Monte Carlo simulation: A new technique for the determination of disordered structure*, Mol. Sim. **1**, 359 (1988).
126. A. K. Soper, *Empirical Monte Carlo simulation of fluid structure*, Chem. Phys. **202**, 295 (1996); *The quest for the structure of water and aqueous solutions*, J. Phys.: Condens. Matter **9**, 2717 (1997).
127. D. Reith, H. Meyer, and F. Müller-Plathe, *Mapping atomistic to coarse-grained polymer models using automatic simplex optimization to fit structural properties*, Macromolecules **34**, 2335 (2001).
128. J. T. Padding and W. J. Briels, *Time and length scales of polymer melts studied by coarse-grained molecular dynamics simulations*, J. Chem. Phys. **117**, 925 (2002).
129. H. Meyer and F. Müller-Plathe, *Formation of Chain-Folded Structures in Supercooled Polymer Melts Examined by MD Simulations*, Macromolecules **35**, 1241 (2002); H. Meyer, *Structure Formation and Chain-Folding in Supercooled Polymer Melts. Some Ideas from MD Simulations with a Coarse-Grained Model*, chapter 10 of Ref. 25.

Entangled Polymers: From Universal Aspects to Structure Property Relations

Kurt Kremer

Max-Planck-Institut für Polymerforschung,
55021 Mainz, Germany
E-mail: kremer@mpip-mainz.mpg.de

Simulations are very versatile tools to study the relaxations 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 paper discusses some basic concepts and conditions and ways to study such problems by computer simulations.

1 Introduction

In the previous lectures of this school the basic concepts to describe polymer melts or dense solutions were introduced and discussed (cf. chapters by J. Baschnagel et al, B. Dünweg and W. Paul). Among others random walks, excluded volume screening, the Rouse model as well as the reptation concept have been included. In the present chapter I mainly focus on the consequences of the fact that chains cannot pass through but only along each other, which eventually leads to the reptation or tube model. To discuss this I will shortly review the background coming from the 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. 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 research¹⁻³.

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 disordered liquids and 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. This question has been examined in detail by the contribution of W. Paul, where simulations of atomistically detailed melts are discussed. When we 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 properties⁴. It is the main focus of the present contribution to discuss generic aspects common to all polymers and then later on 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 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. As shown in previous chapters, the overall extension of the chains is well characterized by the properties of random walks⁷⁻⁹. With ℓ being the average bond length we then have (for $N \gg 1$) for the mean square end to end distance

$$\langle R^2(N) \rangle = \ell_K \ell (N - 1) \approx \ell_K \ell N \quad (1)$$

and $\langle R_G^2(N) \rangle = \frac{1}{6} \langle R^2(N) \rangle$ for the radius of gyration respectively. ℓ_K is the Kuhn length and a measure for the stiffness of the chain. This gives an average volume per chain of

$$V \propto \langle R^2(N) \rangle^{3/2} \sim N^{3/2} \quad (2)$$

leading to a vanishing self density of the chains in a melt. In order to pack beads to the monomer density ρ , $0(N^{1/2})$ other chains share the volume of the very same chain. 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 general property is firmly established by experiment and many simulations¹⁰.

On very large scales 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. A detailed discussion of hydrodynamic effects is given by B. Dünweg in this school. The random 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 exactly¹. This was first done in a seminal paper by Rouse¹¹ and about the same time in a similar fashion by Bueche¹². 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 e.g. $M \leq 20\,000$ for polystyrene [PS] or $M \leq 2000$ for polyethylene [PE], both qualitatively and quantitatively almost perfectly, though the reason is not well understood. Only recently some deviations have been observed¹³. The effects are rather subtle and would require a detailed discussion beyond the scope of this lecture. 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 in the viscosity towards $\eta \sim N^{3.4}$ ¹. The time-dependent 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 in the same way at a chemistry- and temperature-dependent chain length N_e or molecular weight M_e , 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. This is what I am going to discuss in the subsequent chapters.

2 Polymer Dynamics and Network Elasticity

The plateau modulus $G(t)$ can be derived from the restoring force of a polymer melt or network after a step strain. Experimentally usually an oscillatory shear is applied. What one finds then is illustrated in Fig. (1).

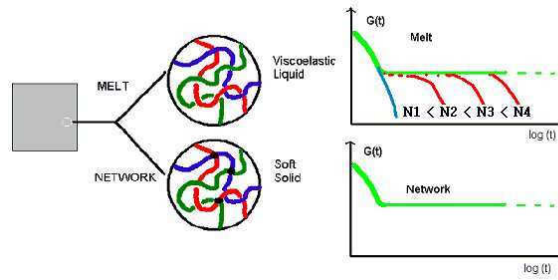


Figure 1. Cartoon of the characteristic structure and response of a polymer melt (top) and a polymer network (bottom) after a step strain. For short chains (length N_1) the restoring force decays to zero very fast, while for the longer ones with increasing length N , as indicated, a plateau in the time dependent modulus occurs, which is independent N

After a drastic fast initial decay, if the chains are long enough, the modulus $G(t)$ stays almost constant at a value G_N^o for a long time until $G(t)$ eventually decays to zero. Many related experimental findings can be found in Ferry's book from 1980¹⁵. Fig. (1) 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 Edwards¹⁶ and deGennes¹⁷.

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. Fig. (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 line and a rather crude sketch of the surrounding. Edwards discussed how the black center chain could move around subject to obstacles created by all the other chains which in this case are part of the network. He noted that due to the 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 called the entanglement length N_e and a sphere of the diameter d_T of the tube typically contains $d_T^{1/\nu} = N_e$ monomers, where $\nu = 1/2$ is the random walk exponent. Within this picture the 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 end. 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. This gives $D \sim N^{-2}$, $\tau_d \sim N^3$ 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.

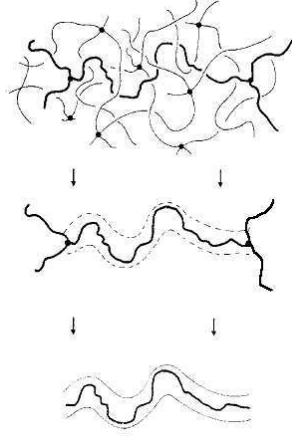


Figure 2. Sketch of the historical development of the tube constraint and reptation concept. Starting from a network Edwards in 1967 defined the confinement to the tube, while deGennes in 1971 realized that for long chains the ends only play a small role for intermediate times.

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 regime². This is a little different for the diffusion of constant D , where the most complete and careful collection of data finds $D \propto N^{-2.4}$ instead of N^{-2} . Here it is not yet settled whether this is the “same” crossover effect²¹ or whether this is mostly due to the so called correlation hole^a effect. A detailed discussion however is beyond the scope of the present contribution.

There have also been a variety of nonreptation/tube phenomenological approaches, which only treat the interactions between the chains either in an averaged mean field approximation or develop a memory function formalism. Though they can reproduce experimental data to some degree, they fail when it comes to the microscopic motions of the polymers and the confinement due to the surrounding strands. Here I do not discuss these approaches any further. Rather comprehensive reviews are given by McLeish² and on the simulation aspects by Kremer and Grest³(and references therein). In a similar way our understanding of networks has improved over the last years. It has been known for a long time, especially also due to simulations, that the noncrossability of the chains plays an important role for the elasticity and related phenomena. Experiments on networks with the

^aThe 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 these spheres 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.

same average strand length but cross-linked at different initial concentrations directly prove this point. However, when it comes to the question of identifying the different contributions (cross-links and entanglements) still conceptual problems exist. Also, the comparison of scattering experiments to theory/simulation only very recently made some significant progress. There the relaxation phenomena are only partially the same as for melts, since the cross-links at the chain ends introduce boundary conditions for the tube contour which do not exist for uncross-linked melts^{22,23}.

From the above 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. Pulsed large field gradient spin-echo NMR is able to address the appropriate time and distance scales as well. However, 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

The Rouse and reptation models are also shortly discussed in the contributions by W. Paul, J. Baschnagel, and B. Dünweg. I now first review some more of the background, restricting myself mostly to quantities which can be investigated directly by simulation. In a melt of homopolymers, the excluded volume interaction is effectively screened. There is no tendency for a chain to swell beyond the ideal random-walk dimension. Only the prefactor, or more precisely the Kuhn length l_K ($\ell_K = \ell_{c\infty}$ in Flory’s terminology²⁴), is governed by the local monomer-monomer interactions.

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, though it actually works very well for short chains in a melt. For chains in solution we refer to the chapter by Dünweg in this volume. Here I follow essentially the book of Doi and Edwards¹ and a recent review by McLeish². The polymer is modelled as a freely jointed chain of N beads connected by $N - 1$ springs, immersed in a Newtonian continuum. Hydrodynamic interactions are neglected. Each bead experiences a friction, with friction coefficient ζ . The beads are connected by a Hookean spring with a force constant $k = 3k_B T/b^2$, where $b^2 = \ell \ell_K$. Each bead-spring unit is intended to model a subchain of the real molecule, not a monomer. The equation of motion of the beads is given by a Langevin equation. For monomer i ($i \neq 1, N$) it reads,

$$\zeta \dot{\mathbf{r}}_i = k(2\mathbf{r}_i - \mathbf{r}_{i-1} - \mathbf{r}_{i+1}) + \mathbf{f}_i \quad (3)$$

Usually the model is solved for a ring with no free ends. If the chain ends are free, as for all linear chains, the first and the last monomer have to be treated differently. For $i = 1$, the first term on the right hand side is $-k(\mathbf{r}_1 - \mathbf{r}_2)$ and similarly for $i = N$. The distribution of random forces $\mathbf{f}_i$ is Gaussian with zero mean and the second moment:

$$\langle \mathbf{f}_i(t) \cdot \mathbf{f}_j(t') \rangle = 6\zeta k_B T \delta_{ij} \delta(t - t'). \quad (4)$$

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 a discrete monomer chain these are given by⁵

$$\mathbf{X}_p(t) = \frac{1}{N-1} \sum_{i=1}^{N-1} \mathbf{r}_i(t) \cos \frac{p\pi(i+1/2)}{N-1} \quad (5)$$

and $p = 0, 1, 2, \dots, N-1$. Equation (3) can be rewritten as

$$\zeta_p \dot{\mathbf{X}}_p = k_p \mathbf{X}_p + \mathbf{f}_p \quad (6)$$

where $\zeta_0 = N\zeta$ and $\zeta_p = 2N\zeta$ for $p \geq 1$ and

$$k_p = 8Nk \sin^2 \frac{p\pi}{2(N-1)}. \quad (7)$$

For small p/N , one recovers the usual result

$$k_p = 2\pi^2 k_p^2 / N = \frac{6\pi^2 k_B T}{N\ell\ell_K} p^2 \quad (8)$$

with $Nb^2 = \langle R^2 \rangle$. Since the random forces $\mathbf{f}_p$ are not correlated, the $\mathbf{X}_p$ decouple and the motion of the polymer can be decomposed into independent modes.

For chains in a melt, the Rouse modes are eigenmodes of the chains. This has been verified by MD for melts of short chains. The time correlation functions of the normal modes, $p \geq 1$, are

$$R_p(t) = \frac{\langle \mathbf{X}_p(t) \cdot \mathbf{X}_p(0) \rangle}{\langle X_p^2(0) \rangle} = \exp(-t/\tau_p), \quad \tau_p = \frac{\zeta(\ell\ell_K)^2 N N}{3\pi^2 k_B T p^2}, \quad (9)$$

where we have used the small p/N approximation for k_p . The longest relaxation time is $\tau_R = \tau_1 \sim N^2$. For long chains in a melt, this equation is expected to only describe the relaxation of high p modes with $N/p \leq N_e$. The relaxation modulus of the melt is given by

$$G(t) = \frac{\rho k_b T}{N} \sum_p R_p(2t). \quad (10)$$

where ρ is the monomer density. This assumes that the single chain Rouse modes can be taken as eigenmodes of the whole melt. This certainly is an assumption, as briefly discussed in the introduction. For the Rouse model this gives

$$G(t) = \frac{pK_B T}{N} \sum_P \exp(-2tp^2/\tau_R) \quad (11)$$

and the viscosity η reads

$$\eta = \int_0^\infty G(t)dt = \frac{\rho k_B T}{2N} \sum_p \tau_p = \frac{\pi^2 \rho k_B T}{12N} \tau_R = \frac{\rho \zeta}{36} \langle R^2 \rangle \sim N \zeta \quad (12)$$

The self-diffusion constant D can be determined from the mean-square displacement of $\mathbf{X}_0 = \mathbf{r}_{cm}$, the center-of mass of the chain,

$$g_3(t) = \langle (\mathbf{r}_{cm}(t) - \mathbf{r}_{cm}(0))^2 \rangle \quad (13)$$

Within the Rouse model $g_3(t) \sim t$ for all times and the diffusion constant $D(N) = \lim_{t \rightarrow \infty} g_3(t)/6t$ is expected to reach the asymptotic value

$$D = \frac{k_B T}{N \zeta} \quad (14)$$

for relatively short times. In simulations often the mean-square displacement of a monomer $g_1(t)$ as a function of time t

$$g_1(t) = \frac{1}{N} \sum_i^N = 1 \langle [\mathbf{r}_i(t) - \mathbf{r}_i(0)]^2 \rangle. \quad (15)$$

is studied. Using the fact that the chain structure is that of a random walk, it is easy to show that

$$g_1(t) \sim \begin{cases} t^1, & t < \tau_0, & g_1(t) < \ell \ell_K \\ t^{1/2}, & \tau_0 < t < \tau_R, & g_1(t) \lesssim \langle R^2 \rangle \\ t^1, & t > \tau_R, & g_1(t) \gtrsim \langle R^2 \rangle \end{cases} \quad (16)$$

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. Within the worm-like chain model in which the chain is a continuous flexible path, τ_0 is zero. 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 is observed free diffusion with a diffusion coefficient $D \sim N^{-1}$. 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 as well as NMR^{25,26} and

neutron spin-echo scattering experiments²⁷ which probe the motion of the monomers agree to that. Results for molecular dynamics simulations on short chains also agree surprisingly well. For short chains, it turns out that the noncrossability of the chains, as well as the chain nature and chemical structure of the surrounding of each monomer mostly affects 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. Clearest evidence for this slowing down comes from the diffusion constant $D^{14,28,29}$. For $N \geq N_e$,

$$D \sim N^{-2} \dots N^{-2.4} \quad (17)$$

Several forms for the prefactor of D have been discussed in the literature. Similarly the viscosity η increases so that

$$\eta \sim N^{3.4} \quad (18)$$

compared to N for short chains. In the reptation theory the motion of the chains is viewed as Rouse motion of chains in a tube of diameter d_T , which follows the coarse grained back bone of the chain. 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 the 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 chain ends 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 only¹.

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 $O(d_T^2 \equiv \langle R^2(N_e) \rangle)$ 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 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 $N \gg N_e$. Following this regime the overall diffusion along the tube gives a second $t^{1/2}$ regime for the motion in space. There the elemental step is the displacement of a polymer chain along its tube. The longest relaxation time is the mean lifetime of the tube. 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} . The theory predicts the following general power-law sequence

for the mean-square displacement in space, $g_1(t)$:

$$g_1(t) \sim \begin{cases} t^1, & t < \tau_0; \\ t^{1/2}, & \tau_0 < t < \tau_e \sim N_e^2; \\ t^{1/4}, & \tau_e < t < \tau_R \sim N^2; \\ t^{1/2}, & \tau_R < t < \tau_d \sim N^3/N_e; \\ t^1, & t > \tau_d \end{cases} \quad (19)$$

which is shown schematically in Fig. (3). For the motion of the center-of-mass $g_3(t)$ one expects

$$g_3(t) \sim \begin{cases} t^1, & t < \tau_e \sim N_e^2; \\ t^{1/2}, & \tau_e < t < \tau_R \sim N^2; \\ t^1, & \tau_R < t \end{cases} \quad (20)$$

These power laws are schematically indicated in Fig. 3. Direct experimental evidence for these intermediate time regimes has been seen NMR and diffusion of polymers at an interface, however simulations were the first to observe the crossover into the $t^{1/4}$ regime³. Indirectly the tube confinement is also seen by scattering experiments.

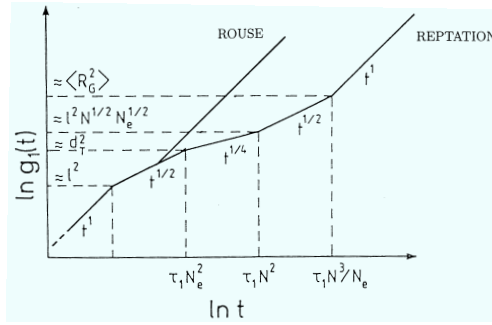


Figure 3. Schematic plot of the mean-square displacement for a monomer in the Rouse and the reptation model.

The reduced mobility also affects the relaxation of the long-wavelength modes $\mathbf{X}_p$. The relaxation time of mode p , with $N/p > N_e$ is enlarged by a factor of N/N_e , giving

$$\tau_{p,Rep} = \frac{N \langle R^2 \rangle}{p^2} \frac{\zeta}{\pi^2 k_B T} \frac{N}{N_e} = \tau_R \cdot \frac{N^3}{3p^2} \sim \frac{N^3}{p^2} \quad (21)$$

The plateau modulus G_N^0 which was at the beginning when the similarities between melts and networks have been discussed, is determined from rheological experiments. Within the reptation model, it is related to the tube diameter $d_T = \langle R^2(N_e) \rangle^{1/2}$ and thus N_e via $G_N^0 = \frac{4}{5} \frac{\rho k_B T}{N_e}$. The prefactor 4/5 arises from tube length fluctuations, which allow for a somewhat more efficient relaxation under deformation, which is not the case for networks. The typical asymptotic properties are again summarized in the table:

	Rouse Regime $N < N_e, M < M_e$	Reptation Regime $N \gg N_e, M \gg M_e$
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 ($t \ll \tau_d$)	$G_N^o = 0$	$G_N^o \propto \frac{kT}{N_e}$

It is one of the central current research topics on entangled polymer systems, how to determine N_e uniquely and how this actually quantitatively determines G_N^o . Before we go into that we discuss shortly the connection to classical network elasticity^{1,15,30}. For this we go back to the random walk statistics of the chains and the idea of the Rouse model that an individual chain experiences all its surrounding only as a viscous background not affecting the conformations at all. For this we start with the probability distribution of chain ends:

$$P(\mathbf{R}) = \left(\frac{3}{2\sigma L\ell_K} \right)^{3/2} \exp(-3R^2/2L\ell_K) \quad (22)$$

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 $\mathbf{R}$. The total number of possible conformations with a separation of $\mathbf{R}$ is the fraction $\Omega(\mathbf{R}) = \Omega_{tot}P(\mathbf{R})$ of all random walk conformations. This defines the entropy for a given $\mathbf{R}$ as

$$S = k_B \ell_K \Omega(\mathbf{R}) = const - 3k_B R^2/2L\ell_K \quad (23)$$

with the energy $U = 0$ (chain is viewed as a free random walk)

$$F = U - TS = \frac{3k_B R^2}{2L\ell_K} \quad (24)$$

for the conformation dependent contribution to the free energy of a chain with fixed end positions. This leads to an entropic force

$$\mathbf{f} = -\nabla F(\mathbf{R}) = -\frac{3k_B T}{L\ell_K} \mathbf{R} \quad (25)$$

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. No let us assume a perfect network with the crosslinks fixed in space. Under deformation $\underline{\underline{\lambda}}$ they move affinely with the macroscopic deformation of the box. Then the force contributed from each network strand is

$$\mathbf{f} = -\frac{3k_B T}{L\ell_K} (\underline{\underline{\lambda}} - \underline{\underline{1}}) \mathbf{R} \quad (26)$$

Each chain contributes with a force constant $\frac{3k_B T}{L\ell_K}$ to the restoring modulus of the system. Since there are ρ/N chains per unit volume present, we can use this to estimate

the stress in a given volume element of volume $V = B^3$. The probability that one chain cuts a plane with index j ($j = 1, 2, 3$) is just R_j/B . The force component transmitted through this plane contributes from this subchain to $3k_B T R_i R_j / L \ell_K B^3$. Summing over all strands passing through our test volume and replacing $R_i R_j$ by the ensemble average $\langle R_i R_j \rangle$ the microscopic stress tensor σ_{ij} reads

$$\sigma_{ij} = \frac{3k_B T \rho}{L \ell_K N} \langle R_i R_j \rangle \quad (27)$$

In analytic theory now usually a continuous path is taken for the chain and then taking the crossover to an infinitely small volume element, using $L = \ell \cdot N$ and for small N : $R_i/N \approx \partial R_i / \partial N$ one gets for the stress tensor

$$\sigma_{ij} = \frac{3k_B T \rho}{\ell \ell_K} \left\langle \frac{\partial R_i}{\partial n} \frac{\partial R_j}{\partial n} \right\rangle \quad (28)$$

Transforming from the bead index to the contour index s and taking the statistical segment length ($\ell \ell_K$) into account, this rewrites to

$$\sigma_{ij} = 3k_B T \frac{\rho}{N} \left\langle \frac{\partial r_i}{\partial s} \frac{\partial r_j}{\partial s} \right\rangle \quad (29)$$

for the microscopic formulation of the stress tensor in a polymer melt or network. In the affine picture this yields

$$G = k_B T \rho / N \quad (30)$$

In networks a variety of modifications are introduced. The above equation assumes an affine deformation of the whole system. This certainly is not the case! In an alternative approach the positions of the network beads at the box surface are fixed and the whole rest is only subject to connectivity constraints. In this phantom network model the modulus is somewhat reduced. In general, disregarding any entanglement contribution one can write

$$G = \frac{\rho k_B T}{N} [\nu - h\mu] \quad (31)$$

where ν is the number of active strands in a network (no closed loops ending in the same point, no free ends ...) and μ is the number of crosslinks. The quantity $0 \leq h \leq 1$ interpolates between the two models and is often taken as an adjustable parameter.

In the case of very long chains, $N \gg N_e$ so that the simple Rouse model assumption that the chains or strands act as non-interaction random walks a hierarchy of models trying to identify constraints has been introduced. They range from the constraint junction fluctuation approach of Flory all the way to the tube model. Ideas along these lines are employed to understand the spatial fluctuations in randomly crosslinked systems under strain. Taking all this one can view the strands in the Edwards tube as a continuation of entropic springs of length N_e , which leads to¹⁵

$$G_N^o = \frac{\rho k_B T}{N_e} \quad (32)$$

for networks and taking additional relaxation mechanisms into account

$$G_N^o = \frac{4}{5} \frac{\rho k_B T}{N_e} \quad (33)$$

for melts. This is the formula commonly used to determine N_e from measurements of the plateau modulus. A typical example is given in Fig. (4).

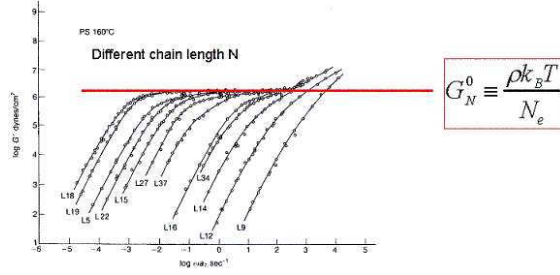


Figure 4. Experimental determination of the frequency dependent modulus $G'(\omega)$ of polystyrene melts for many different chain lengths, from¹

Considering the previous discussion simulations 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? Whatever are the links between different experimental observations?
- What are the effects of dilution and/or solvent quality in semidilute systems?
- What does the tube “look like”?
- Tube deformation and tube relaxation, quantitative and qualitative differences between networks and melts
- Swelling behavior with and without the influence of charges (polyelectrolyte gels).
- Polydispersity effects, 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?

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.

3.3 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^{10,31,5}. To study entanglement effects one can construct special model situations and study those. However more information applicable to experiment can be gained by studying “well equilibrated samples”. A variety of methods to generate those has been discussed in other chapters. The most direct way would be to set up a system in an arbitrary state and equilibrate by running it for a few relaxation times. This however would need CPU time which by itself is of at least of the order of the running time of the whole project. Employing standard MD one typically can reach around $(2 - 4) \times 10^5$ particle steps per second on a single processor. A typical (large) time step is about $10^{-2} \tau$ for a dense Lennard Jones system (see below). For that system with $\tau_e \approx 2000\tau$ (estimated from the mean square displacement of inner beads, taking the stress relaxation after a step strain τ_e is estimated close to 10000τ), one needs about $(2 - 5) \times 10^4$ steps or at least one second of CPU time per particle to reach τ_e ³². Thinking of about $M = 100$ chains of $N \approx 800$, about $10N_e$, a system just big enough to compete with typical modern scattering experiments, for the fully flexible bead spring system defined below, we get $\tau_D \geq \tau_e \cdot (N/N_e)^3 = 10^3 \tau_e$, equivalent to 10^3 sec per particle. Thus the above system on a single processor would require at least $M \cdot N \cdot 10^3 \text{ sec}$, more than 20000 hours. Modern parallel systems reduce this time significantly, however, more complex models, such as semiflexible chains or models which more closely resemble the atomistic structure easily increase the time by several orders of magnitude time. For typical all atom models one can estimate the increase in CPU time roughly by the following. Taking about the same number of monomers for one entanglement molecular weight N_e (which is typically not enough) one can multiply the above time by the number of atoms per monomer (e. g. 4 for PE and 16 for PS ...) and the average ratio of the integration time steps, which is between 10^3 and 10^4 . To reach the same relaxation, about $10^4 - 10^5$ times more CPU time is needed, which is beyond reach. This estimate did not take into account the typically higher complexity of the force calculations! These numbers illustrate that the computers of the nearer future also will not solve the problem. Also, the tendency to reach high computing power by increasing the number of processors is not helpful since our systems are moderate in size but need long time runs. To study in very detail the final dependency of the elastic response of a system, one would need roughly $20 N_e$ and run that system for a time $\tau \gg \tau_e$ and eventually close to τ_d . Fortunately, for equilibration there are ways out. However, this requires a solid knowledge of the average equilibrium conformation³³. 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 the 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. There are two typical options, internal distances or the chain form factor. For the internal distances $\langle \Delta r_{ij}^2 \rangle$ one finds

$$\langle \Delta r_{i-j}^2 \rangle \propto \begin{cases} |i-j|^2 & |i-j| \ell < \ell_k \\ |i-j| & |i-j| \ell \gg \ell_k \end{cases} \quad (34)$$

For distances $r \gg l$ beyond the local sphere or bead packing the crossover between the two regimes is approximated by the worm like chain model:

$$\langle \Delta r_{|i-j|=n}^2 \rangle = nl^2 \left(\frac{1 + \langle \cos \theta \rangle}{1 - \langle \cos \theta \rangle} - \frac{1}{n} \frac{2 \langle \cos \theta \rangle (1 - \langle \cos \theta \rangle^n)}{(1 - \langle \cos \theta \rangle)^2} \right) \quad (35)$$

where the asymptotic prefactor also is referred to as

$$c_\infty = \frac{1 + \langle \cos \theta \rangle}{1 - \langle \cos \theta \rangle}. \quad (36)$$

with $\langle \cos \theta \rangle$ the average bond angle “measured” between subsequent bonds. This function has to be determined in a complete simulation of shorter chains. Thus in a melt a system specific master plot gives $\langle \Delta r_{ij}^2 \rangle / l |i - j|$ vs $l |i - j|$ the corresponding contour length. A characteristic example is given in Fig. (5). This shows characteristic deviations from the theoretically expected curve for small distances, where the bead packing is dominant. For larger distances, the more flexible chains still show significant deviations. There the approximate formula of Eq. (5) is only useful as a general guide.

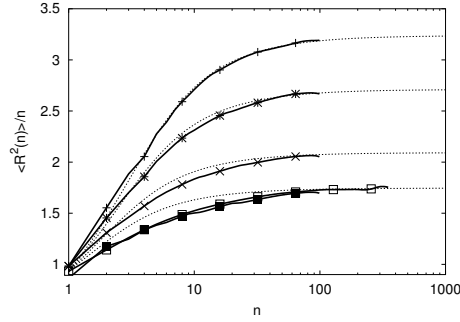


Figure 5. Normalized internal distances for different chain lengths for a standard bead spring LJ polymer, for different ℓ_K ranging from about $\ell_K = 1.7\sigma$ to 3.3τ . From³³

In a similar way one can apply the analysis of the chain form factor $S(k)$

$$S(k) = \frac{1}{N} \left\langle \left| \sum_{j=1}^N e^{i\mathbf{k} \cdot \mathbf{r}_j} \right|^2 \right\rangle_{|k|} \quad (37)$$

where the index $|k|$ denotes a spherical average. In a melt one finds

$$S(k) \propto \begin{cases} N(1 - \frac{1}{3}k^2 \langle R_G^2 \rangle) & \frac{2\pi}{k} \gg \langle R_G^2 \rangle \\ k^{-2} & \frac{2\pi}{\ell_k} \ll \frac{2\pi}{k} \ll \langle R_G^2 \rangle \\ k^{-1} & \frac{2\pi}{\ell} < \frac{2\pi}{k} < \frac{2\pi}{\ell_k} \\ O(1) & \frac{2\pi}{k} > \frac{2\pi}{\ell} \end{cases} \quad (38)$$

Besides the initial decay the form factor describes one characteristic, chain length independent curve. Fig. (6) shows a typical example.

Often the so-called Kratky plot, $k^2 S(k)$ vs k , is shown. Deviations from slope zero in the k^{-2} regime are easily to detect, even by eye. This helps for a quick check. Master plots of $S(k)$ or $\langle \Delta r_{ij}^2 \rangle$ can be used to control the equilibration of melts.

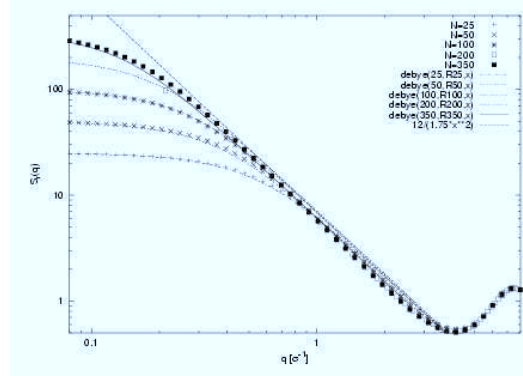


Figure 6. Chain form factor $S(q)$ vs q for fully flexible LJ chains in a melt. Chain lengths range from $N = 25$ to $N = 350$. For fits the Debye function was used. The straight line indicates a slope of q^{-2} , from³⁴

3.4 Preparing an Equilibrated Melt or Network, 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 very appropriate. For dense continuum systems as well as systems with realistic chemical details such an approach fails. The strategy we meanwhile follow is as follows³³:

- 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. Fig. (6))
- Create non reversal random walks of the correct bond length, which match the target function closely, especially at the longer distances. 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 interaction, taking care that in the beginning the chains can easily cross each other (see below)
- Run until the excluded volume potential is completely introduced. Control internal distances permanently to check for possible overshoots, deviations.
- Eventually support long range relaxation by so called end bridging^{35,36} or double pivot moves

Independent on the details of the procedures, it is important to continuously compare the actual structure to the master or target curves. Ref.³³ also demonstrates some typical deviations from the master curve as they can occur during the set up.

As an example I now discuss the above steps along a recent publication³³ on simple bead spring chains^{37,38}. An extension to other chains is straight forward. We consider bead with a unit mass. All beads interact via a purely repulsive LJ potential, to model the excluded volume interaction.

$$U_{LJ}^r = \begin{cases} 4\epsilon \left\{ (\sigma/r)^{12} - (\sigma/r)^6 + \frac{1}{4} \right\} & r \leq r_c \\ 0 & r \geq r_c \end{cases} \quad (39)$$

with a cutoff $r_c = 2^{1/6}\sigma$. The beads are connected by a finite extensible non-linear elastic potential (FENE)

$$U_{FENE}^{(r)} = \begin{cases} -0.5R_o^2 k \ell_n (1 - (r/R_o)^2) & r \leq R_o \\ \infty & r > R_o \end{cases} \quad (40)$$

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 $\tau = \sigma(m/\epsilon)^{1/2}$. Volume and temperature are kept constant and the bead density is $\rho = 0.85\sigma^{-3}$. The temperature was kept constant by coupling the motion to a Langevin thermostat. An alternative, which does not screen hydrodynamics is the DPD thermostat (cf. contribution by B. Dünweg). The average bond length with these parameters is $\langle \ell^2 \rangle^{1/2} = 0.97\sigma$. Chains of N beads each are considered.^b

For this model in a melt of density $\rho = 0.85\sigma^{-3}$ the bond length is $\langle \ell^2 \rangle^{1/2} = 0.97\sigma$ with $c_\infty \ell^2 = ll_k = 1.7\sigma^2$. Since, in our case, there are no torsional barriers, the monomer packing relaxes quickly. It locally depends very sensitively on the ratio of bond length to effective excluded volume of the beads. Both do not allow **any** conclusion for the overall relaxation of the melt. 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 from equilibrium at the large length scales on the order of the chain extension. For this system chains of $N = 350$ beads (which was not really needed in that case) were equilibrated by brute force simulations running in the background. Fig (5) also includes some data for stiffer chains. From these data we get the target function:

$$\langle R^2(|i-j|, N) \rangle / |i-j| \propto \ell \ell_k, |i-j| \ell \gg \ell_k \quad (41)$$

Having this reference state, we can proceed as mentioned before. The crucial steps are the last two. To introduce the excluded volume we proceed as follows:

1. Use a force-capped-Lennard-Jones-potential
2. increase the push-off time to make the procedure “quasi static”.

^bIn 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 $\hat{\mathbf{r}}_{i,i-1} = (\mathbf{r}_i - \mathbf{r}_{i-1}) / |\mathbf{r}_i - \mathbf{r}_{i-1}|$ and accordingly $\hat{\mathbf{r}}_{i,i+1}$ 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.

The force capped potential we use reads :

$$U_{\text{FCLJ}}(r) = \begin{cases} (r - r_{fc}) * U_{\text{LJ}}'(r_{fc}) + U_{\text{LJ}}(r_{fc}) & r < r_{fc} \\ U_{\text{LJ}}(r) & r \geq r_{fc} \end{cases} \quad (42)$$

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).

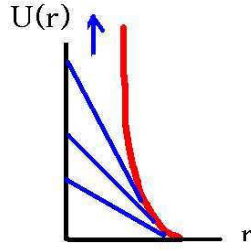


Figure 7. Illustration of the “force capped” Lennard Jones interaction used to introduce step by step the full excluded volume of the chains.

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

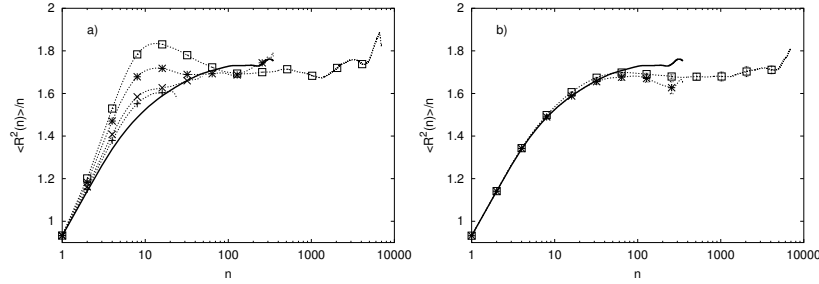


Figure 8. Mean square internal distances for chains of length $N = 25$ (+), 50 (x), 350 (*) and 7000 (□) for a too fast pushoff procedure (a) and the slower version as described in the text (b), from³³.

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, called the double-pivot or bridging algorithms are used. Note that they very effectively relax and manipulate large

distances, however affect small scales only extremely slowly! 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 different methods.

The preparation of networks deserves another short comment. In the literature a number of different crosslinked systems has been studied extensively^{34,39-43}. A typical range of systems is illustrated in Fig. (9)

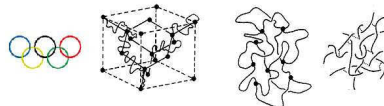


Figure 9. Typical crosslinked systems ranging from randomly crosslinked melts via various versions of endlinked polymer networks to the idealized lattice structure systems and the theorist's dream of an "olympic gel"

The randomly crosslinked system certainly is closer to experiments however encounters many difficulties³⁹. 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 role². In addition the disorder or fluctuations are 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.

3.5 Entanglement Analysis: Melts and Networks

In the following I will give a few examples of numerical experiments for melts and in some cases networks as well. The first almost trivial question is, whether the tube exists, the chain motion is confined and whether the noncrossability clearly makes a difference. 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. To do so one simply can plot the initial conformation of the chain and on top of that subsequent conformations. This is shown in Fig. (10) for network chains.

From other investigations we know that $35 < N_e$ and $100 > N_e$, which is nicely confirmed by the pictures. 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 fits 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 in the number of constraints, this for shorter chains, is more difficult to visualize. Thus one uses a small trick. What we now plot is not the bare conformation, but the back bone of the statically averaged chain contour, namely

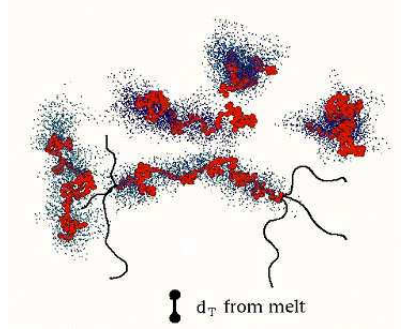


Figure 10. Time evaluation of the conformation of chains in an endlinked network of functionality $f = 4, N = 100$, where all network strands have the same length. The chains are, for clarity shown without the network. The indicated tube diameter was determined in an independent melt simulation, from⁴⁰.

$$\mathbf{R}_i = \frac{1}{n+1} \sum_{j=1-n/2}^{i+n/2} \mathbf{r}_i \quad (43)$$

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 can then be plotted vs time and compared to a different system, where the chains can easily cross each other. For the full LJ interaction the barrier height is around $70kT$. One can 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

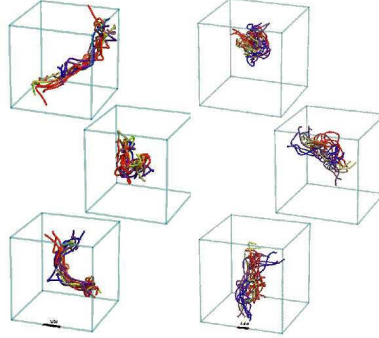


Figure 11. Time evolution of the coarse grained backbone of the chains for a melt of $N = 350$ bead chains for times up to the Rouse time τ_R . The left panel shows chains with the full excluded volume, while for the right panel the crossing barrier is only a few kT , from³⁴.

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 these networks with diamond lattice

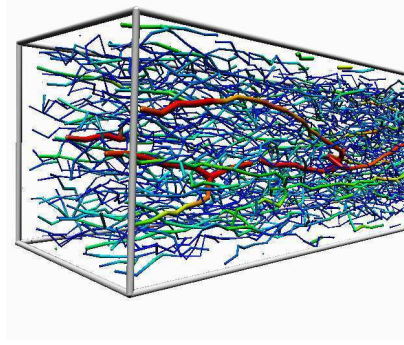


Figure 12. Elongated diamond lattice networks, where the only source of disorder are random links between network loops. The strands are colored due to their stretching (similar to the stress they carry) from small stretching (blue) to strong stretching (red). From³

connectivity were prepared. 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.

After these intriguing pictures we can move to the data analysis which demonstrates the confinement not that directly, however is more closely related to possible experiments. The first to look at is the mean square displacement of the middle beads and the chains. Many studies on this have been performed^{38,44–47}, for an overview see³ and computer simulations actually were the first to observe the slowing down of the motion clearly^{38,48}. Fig. (13) shows results from a system of fully flexible polymers. More complicated models have

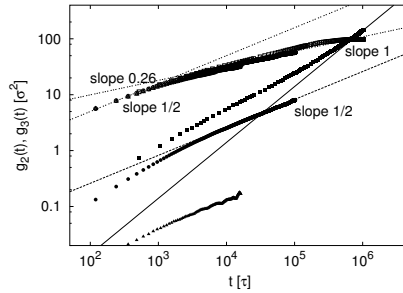


Figure 13. Mean square displacements $g_2(t)$ (open symbols) and $g_3(t)$ (closed symbols) for chain length $N = 350$ ($\square$), 700 ($\circ$) and 10000 ($\triangle$). The straight lines show power laws as guide to the eye. The local reptation power laws $g_2(t) \propto t^{1/4}$ and $g_3(t) \propto t^{1/2}$ are verified with remarkable clarity³².

been studied by Faller and Müller-Plathe^{49,50} and⁴⁵ for instance. Fig. (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 $N_e \approx 30$, 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 analyse the chain motion directly and display the expected power laws. They however 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 $k_B T / \zeta$ but via conformational changes also N_e ! Here simulations, despite of their many shortcomings are still superior.

Another important experimental technique is the neutron spin echo (NSE) method²⁷. In NSE experiments the motion of the segments can be obtained by measuring the expectation value of the time-dependent single-chain structure function

$$S(\mathbf{k}, t) = \frac{1}{N} \sum_{i,j} \langle \exp(i\mathbf{k} \cdot (\mathbf{r}_i(t) - \mathbf{r}_j(0))) \rangle. \quad (44)$$

For reptating chains deGennes calculated an expression for $S(k, t)$ neglecting the initial Rouse-like motion⁵¹. In recent simulations a slightly modified version is applied^{52,53} 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):

$$\begin{aligned} \frac{S(k, t)}{S(k, 0)} = & \left\{ [1 - \exp(-(kd/6)^2)] \cdot f\left(k^2 b^2 \sqrt{12Wt/\pi}\right) \right. \\ & \left. + \exp(-(kd/6)^2) \right\} \times \frac{8}{\pi^2} \sum_{p=1, \text{odd}}^{\infty} \frac{\exp(-tp^2/\tau_d)}{p^2}, \end{aligned} \quad (45)$$

where $f(u) = \exp(u^2/36)\text{erfc}(u/6)$. Note, that the derivation of deGennes does not take into account chain end effects and thus the above formula may only be applied in a narrow range of k -space $\frac{2\pi}{R_G} \lesssim k \lesssim \frac{2\pi}{d_T}$. Also since the initial Rouse-motion is not explicitly described, this formula only applies for $t > \tau_e$ (The short time motion only enters through the inverse friction coefficient $W = k^T / \zeta \ell \ell_K$).

What one finds there is 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. Though these are, to my knowledge, the best data available, there might be some problems with the statistics ($N = 10000$). 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 we index N_e determined from the plateau-modules with an additional index p) is by measuring the plateau modulus under oscillatory shear^{15,1}. 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 was σ_N was determined by the

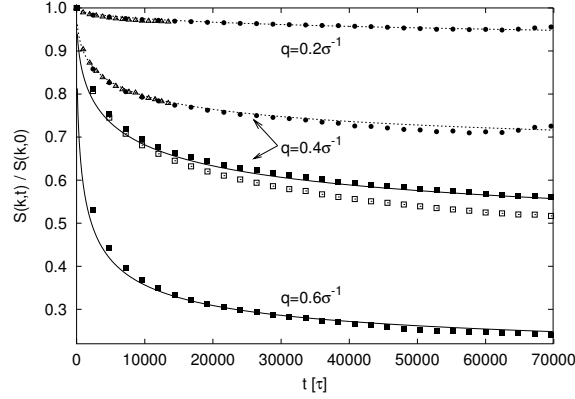


Figure 14. $S(k, t)$ for different chain lengths: $N = 10000(\Delta)$, $N = 2000(*)$, and $N = 700(\square)$ and the centered subchain of length 550 of the same $N = 700$ chains ($\blacksquare$). Continuous curves are simultaneous fits to the $N = 2000$ data corresponding to $d_T = 9.6\sigma$. The dotted curve is a simultaneous fit to the $N = 700, 550$ monomer subchain with a tube diameter of $d_t = 12.9\sigma$.^{32,54}

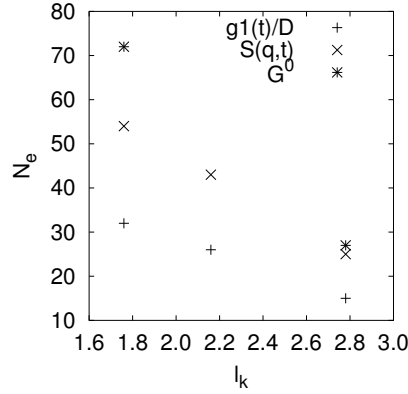


Figure 15. N_e as a function of the Kuhn length of the chains ℓ_k determined from the mean square displacements, the scattering function and the modulus.

microscopic virial-tensor. The plateau-values of the stress were fitted to the stress-strain formulas for classical rubber elasticity (CRE)⁵⁵ $\sigma_N = G \left(\lambda^2 - \frac{1}{\lambda} \right)$ and to the Mooney-Rivlin (MR) formula⁵⁶ $\sigma_N = 2G_1 \left(\lambda^2 - \frac{1}{\lambda} \right) + 2G_2 \left(\lambda - \frac{1}{\lambda^2} \right)$ to determine G_N^0 . MR gives $G_N^0 = 0.0105k_B T \sigma^{-3}$ while CRE gives $G_N^0 = 0.008k_B T \sigma^{-3}$. It is known experimentally that MR slightly overestimates the modulus while CRE always underestimates it. The standard formula¹ to calculate $N_{e,p}$,

$$G_N^0 = \frac{4}{5} \frac{\rho k_B T}{N_{e,p}}, \quad (46)$$

gives $N_{e,p} \approx 65$ for the MR fit and $N_{e,p} \approx 80$ for the classical formula. Both values

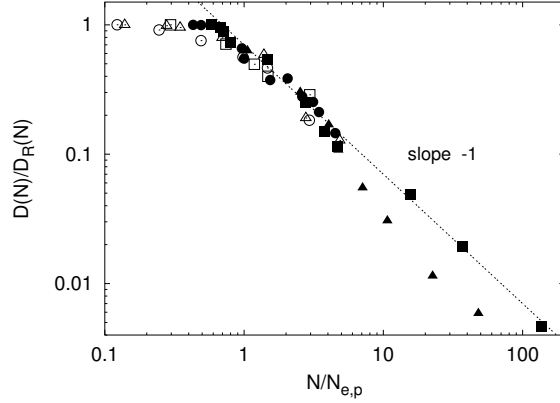


Figure 16. Scaled diffusion constant $D(N)/D_R(N)$ vs. scaled chain length $N/N_{e,p}$ for polystyrene ($\bullet$) ($M_{e,p} = 14600$, $T = 485K$), polyethylene ($\blacksquare$) ($M_{e,p} = 870$, $T = 448K$), PEB2 ($\blacktriangle$) ($M_{e,p} = 990$, $T = 448K$), the present bead spring model ($\triangle$) ($N_{e,p} = 72$), 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 $2.25N_e$ from $g_1(t)$. From³²

are much higher than our previous estimate, $N_e = 32$. From a theoretical point of view this 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 difference is relevant.

At this point we can conclude that we are able to clearly observe and demonstrate entanglement effects. However the different values originating from different experiments prohibit to give an overall consistent picture. Fig. (15) shows the outcome of such investigations as we plot the entanglement molecular weight as a function of stiffness for different “experimental” tests. This illustrates that one has to decide on one definition of N_e . Taking the original discussion on the chain confinements, the results taken from the modulus be taken as a reference state. If done like this, one can use this to scale other measurements. Fig. (15) shows the scaled diffusion constants for a variety of experiments. The actual diffusion constant $D(N)$ is normalized by the hypothetical Rouse regime diffusion constant $D_R(n)$ as it can be extrapolated from short chain simulations or experiments. However, it should be noted that a similar plot employing e. g. $S(k, t)$ will lead to different fits for N_e for different polymers or models systems.

3.6 Structure and Property Relations

Though a general picture emerges, there is still the problem of the structure property relation. Is there a way to predict N_e or at least one value for a specific experiment based on the conformations, meaning the chemistry alone?

We saw that with increasing stiffness, the value of N_e decreased, however, the ratio of the different values obtained from the mean square displacements, scattering and modulus did not remain constant. Also, the modulus always predicted the largest value. This seems to hold for simple bead spring chains, while there are chemical systems where this is not

the case. Polycarbonate as an example of a technically very important polymer (e. g. for compact discs) has an $N_e \approx 6$ from the modulus, while the mean square displacement suggests via a scaling from Fig. (15) a value of $N_e \approx 12^{57}$. Obtaining just one measurement obviously is not sufficient! Thus there is a significant need for a predictive theory or ansatz, which determines a value of N_e from the chain conformations, which then can be used to predict the other quantities, especially the modulus. How different N_e , or M_e respectively is shown in the table. A typical ambient temperatures (note that the M_e depends on temperature just as the overall chain extension does) one obtains the following values:

	PE (Polyethylene)	PS (Polystyrene)	PDMS (Polydimethyl-Siloxane)	BPA PC (Polycarbonate)
N_e	100	170	135	5 - 7
M_e	1400	18000	10000	≈ 1700

Even within one class of materials significant differences can be observed, as indicated in Fig. (17). There different modifications of polycarbonate are shown, which within experimental accuracy have about the same Kuhn length ℓ_K but rather different entanglement lengths, as indicated^{58,59}.

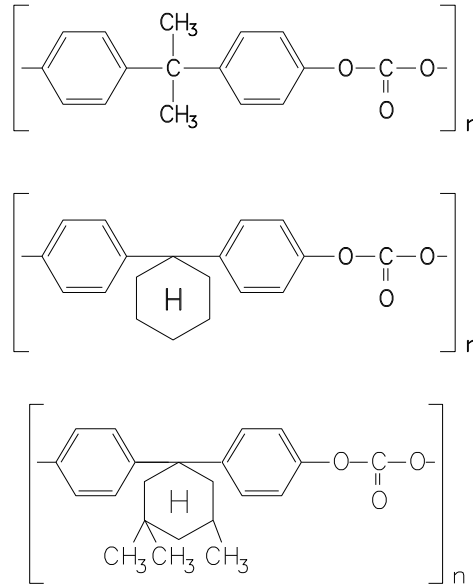


Figure 17. Structure of different polycarbonate modifications, the respective entanglement molecular weight are $N_e \approx 6, 10, 15$ from top to bottom.

These results demonstrate a key problem for a successful link between the more basic science, generic phenomena, oriented research and modern materials science. Our understanding has to become more quantitative. This holds for many fields in the science of microscopically or nanoscopically designed molecular structures as well as for more conventional bulk systems. Because of the huge relaxation times a factor of 2 in the entanglements change the viscosity also by a factor of $2^{1.4} \approx 2.6$ and similar reduces the

diffusion constant. Uncertainties of this order in predictions are relatively small, however, are of significant experimental and technological relevance.

Is there a way to understand these deviations as well as the sensitivity to chemical structure and conformation? Over the years several criteria have been developed to identify entanglements. One set of publications deals with an analysis of the topological structure of a network and related approximations for melts⁶⁰. For this the problem however is that multi chain / multi ring effects can 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. (18) for an example of three chains only^{42,61}.

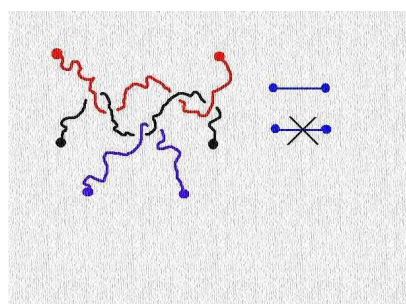


Figure 18. Constraints between two chains (red, black) depend on the presence of the 3rd chain.

In principle one would need for long chains an “infinite” hierarchy of such interactions. On the other hand, in order to observe N_e is dependent on N for $N \gg N_e$ a constant average density of constraints along the backbone of the chain is needed to provide the tube. Iwata and Edwards^{61,62} tried to overcome this by looking at the local values of the Gauss integral upon numerical integral. All this suggests that the local chain-chain packing plays a crucial role. Therefore relatively early first attempts were made to relate N_e to the number of different chains within the value d_T^3 ^{63,64}. All these attempts were not really satisfactory. More recently 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 $p = N/\rho\langle R \rangle$, which is a good measure for the spatial packing, in a way that $G \sim p^{-3}$. 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. Beyond that Everaers et al.²⁰ just recently performed a topological analysis of a number of different polymer melts and dense solutions. They studied simple bead spring chains of different stiffness as well as simplified model for polycarbonate. In all cases the data of G_N^0 derived from the topological determination of the primitive path and the analysis from analyzing normal stress differences after a step strain quantitatively agree. Thus it seems that (1) there is a way to relate purely static quantities to the modulus with reasonable accuracy and (2) the tube model can quantitatively be linked to a topological analysis of the state of the melt. With this we arrive at the question of the very beginning, namely to what extent does the fact that the chains do not pass through each other, influence the relaxational properties of these fascinating ultra soft solids or viscoelastic liquids. Based on recent

progress in various fields extensions to more complicated systems such as mixtures of linear and branched polymers or more stiff polymers, as they can be found in biological systems, are underway.

4 Summary

In the first part of this contribution certain aspects concerning the dynamics and stress relaxation in polymer melts or networks, complementary to the other contributions has been discussed. In addition typical “measurements” to study reptating polymer melts have been mentioned. From this a basic ingredient to study the dynamics of polymer melts and networks, namely the equilibration and preparation of suitable starting conformations has been presented in more detail. This is a crucial condition for the next step, namely to study both the dynamic development as well as the conformational properties of a system. Then a few typical examples of numerical tests performed over the last years illustrated, what can be done by simulations and dense model systems employing modern numerical methods and computers.

Acknowledgement

I would like to thank many collaborators over many years of research. The topics described above benefitted especially from longstanding fruitful interaction with G. S. Grest and R. Everaers.

References

1. M. Doi and S. F. Edwards. *The Theory of Polymer Dynamics*. Clarendon, Oxford, 1986.
2. T. C. B McLeish. Tube theory of entangled polymer dynamics. *Adv. in Phys.*, 5:1379-1527, 2002.
3. K. Kremer and G. S. Grest. In K. Binder, editor, *Monte Carlo and Molecular Dynamics Simulations in Polymer Science*, page 194. Oxford University Press, New York, 1995.
4. C. F. Abrams and K. Kremer. *J. Chem. Phys.*, 116:3162, 2002.
5. K. Kremer. *Soft and Fragile Matter, Nonequilibrium Dynamics, Metastability and Flow*. NATO ASI Workshop, St. Andrews, 2000.
6. K. Kremer C. F. Abrams, L. DelleSite. In M. Mareschal and G. Ciccotti, editors, *Bridging Time Scales: Molecular Simulations for the Next Decade*. Springer, Berlin, Heidelberg, 2002.
7. P. G. de Gennes. *Scaling Concepts in Polymer Physics*. Cornell University Press, Ithaca NY, 1979.
8. M. Rubinstein and R. Colby. *Polymer Physics*. Oxford, Oxford, 2003.
9. A. R. Khokhlov A. Yu Grosberg. *Statistical Physics of Macromolecules*. AIP Press, New York, 1994.
10. K. Binder. In K. Binder, editor, *Monte Carlo and Molecular Dynamics Simulations in Polymer Science*, page 356. Oxford University Press, New York, 1995.

11. P. E. Rouse. *J. Chem. Phys.*, 21:1272, 1953.
12. F. Bueche. *J. Chem. Phys.*, 22:603, 1954.
13. W. Paul, G. D. Smith, D. Y. Yoon, B. Farago, S. Rathgeber, A. Zirkel, L. Willner, and D. Richter. Chain motion in an unentangled polyethylene melt: A critical test of the rouse model by md simulatons and neutron spin echo spectroscopy. *Phys. Rev. Lett.*, 80:2346, 1998.
14. H. Tao, T. P. Lodge, and E. D. von Meerwall. *Macromolecules*, 33:1747, 2000.
15. J. D. Ferry. *Viscoelastic Properties of Polymers*. Wiley, New York, 1980.
16. S. F. Edwards. *Proc. Phys. Soc.*, 92:9, 1967.
17. P. G. de Gennes. *J. Chem. Phys.*, 55:572, 1971.
18. L. J. Fetters, D. J. Lohse, S. T. Milner, and W. W. Graessley. *Macromolecules*, 32:6847, 1999.
19. L. J. Fetters, D. J. Lohse, and W. W. Graessley. *J. Pol. Sc. B. Pol. Phys.*, 37:1023, 1999.
20. R. Everaers, S. K. Sukumaran, G. S. Grest, C. Svaneborg, A. Sivasubramanian, and K. Kremer. *preprint*, 2003.
21. S. T. Milner and T. C. B. McLeish. Star polymers and failure of time-temperature superposition. *Macromolecules*, 31:8623, 1998.
22. G. S. Grest, M. Pütz, R. Everaers, and K. Kremer. *J. Non-Crystal. Solids*, 274:139, 2000.
23. B. Mergell and R. Everaers. *Macromolecules*, 34:5675, 2001.
24. P. Flory. *Principles of Polymer Chemistry*. Cornell University Press, Ithaca, 1953.
25. G. Fleischer and F. Fajara. In R. Koesfeld and B. Blümich, editors, *NMR Basic Principles and Progress*. Springer-Verlag, Heidelberg, 1994.
26. H. W. Spiess. *Macromol. Symp.*, 174:111, 2001.
27. B. Ewen and D. Richter. *Adv. Pol. Sci.*, 134:3, 1997.
28. D. S. Pearson. *Rubber Chem. Tech.*, 60:439, 1987.
29. D. S. Pearson, L. J. Fetters, W. W. Graessley, G. ver Strate, and E. von Meerwall. *Macromolecules*, 27:711, 1994.
30. S. F. Edwards and T. A. Vilgis. *Rep. Prog. Phys.*, 51:243, 1988.
31. K. Kremer. *Computer simulation methods for polymer physics*. Intl. Phys. Soc., Rome, 1996.
32. M. Pütz, K. Kremer, and G. S. Grest. *Europhys. Lett.*, 49:735, 2000.
33. R. Auhl, R. Everaers, G. S. Grest, K. Kremer, and S. J. Plimpton. *J. Chem. Phys.*, 2003.
34. M. Pütz, *Dynamik von Polymerschmelzen und Quellverhalten ungeordneter Netzwerke*, Ph.D. Thesis, University of Mainz, 1999.
35. V. G. Mavrantzas, T. D. Boone, E. Zervopoulou, and D. N. Theodorou. *Macromolecules*, 32:5072, 1999.
36. A. Uhlherr, S. J. Leak, N. E. Adam, P. E. Nyberg, M. Doastakis, V. G. Mavrantzas, and D. N. Theodorou. Large scale atomistic polymer simulations using Monte Carlo methods for parallel vector processors. *Comp. Phys. Comm.*, 144:1, 2002.
37. G. S. Grest and K. Kremer. *Phys. Rev. A*, 33:3628, 1986.
38. K. Kremer and G. S. Grest. *J. Chem. Phys.*, 92:5057, 1990.
39. E. R. Duering, K. Kremer, and G. S. Grest. *Phys. Rev. Lett.*, 67:3531, 1991.
40. E. R. Duering, K. Kremer, and G. S. Grest. *J. Chem. Phys.*, 101:8169, 1994.

41. M. Pütz, R. Everaers, and K. Kremer. Self-similar chain conformations in polymer gels. *Phys. Rev. Lett.*, 84:298, 2000.
42. R. Everaers and K. Kremer. *Phys. Rev. E*, 53:R37, 1996.
43. R. Everaers. *New J. Phys.*, 1:12.1, 1999.
44. W. Paul, K. Binder, D. W. Heermann, and K. Kremer. *J. Chem. Phys.*, 95:7726, 1991.
45. W. Paul, K. Binder, K. Kremer, and D. W. Heermann. *Macromolecules*, 24:6323, 1991.
46. S. W. Smith, C. K. Hall, and B. D. Freeman. *J. Chem. Phys.*, 104:5616, 1996.
47. M. Kröger and H. Voigt. On a quantity describing the degree of entanglement in linear polymer systems. *Macromol. Theory Simul.*, 3:639, 1994.
48. K. Kremer, G. S. Grest, and I. Carmesin. *Phys. Rev. Lett.*, 61:566, 1988.
49. R. Faller, F. Müller-Plathe, and A. Heuer. *Macromolecules*, 33:6602, 2000.
50. R. Faller and F. Müller-Plathe. *Chem. Phys. Chem*, 2:180, 2001.
51. P. G. de Gennes. *J. Phys. (Paris)*, 42:735, 1981.
52. K. Kremer and K. Binder. *J. Chem. Phys.*, 81:6381, 1984.
53. M. Doi. *J. Phys. A*, 8:959, 1988.
54. M. Pütz, K. Kremer, and G. S. Grest. *Europhys. Lett.*, 52:721, 2000.
55. L. R. G. Treloar. *The Physics of Rubber Elasticity*. Clarendon Press, Oxford, 1986.
56. M. Mooney. *J. Appl. Phys.*, 11:582, 1940.
57. S. Leon and K. Kremer. *in preparation*.
58. K. Sommer, J. Batoulis, W. Ilge, 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. Ludovice. Correlation between primary chemical structure and property phenomena in polycondensates. *Advanced Materials*, 3:590, 1991.
59. 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, 2000.
60. W. Michalke, M. Lang, S. Kretmeier, and D. Gönitz. *Phys. Rev. E*, 64:012801, 2001.
61. K. Iwata and S. F. Edwards. *J. Chem. Phys.*, 90:4567, 1989.
62. K. Iwata. *Macromolecules*, 24:1107, 1991.
63. T. A. Kavassalis and J. Noolandi. *Phys. Rev. Lett.*, 59:2674, 1987.
64. T. A. Kavassalis and J. Noolandi. *Macromolecules*, 22:2709, 1989.

Chemically Realistic Computer Simulations of Polymer Melts: Equilibration Issues and Study of Relaxation Processes

W. Paul

Institut für Physik, Johannes Gutenberg-Universität,
55099 Mainz, Germany
E-mail: wolfgang.paul@uni-mainz.de

Introduction

Polymers are complex objects displaying non-trivial structure from the scale of a (typically) carbon-carbon bond ($1 \text{ \AA}$) to the radius of gyration of the coil ($R_g \approx 10 - 100 \text{ \AA}$). Here we will be concerned with melts of simple linear polymers where the polymer coils behave as if they were random walks ($R_G \propto N$ where N is the degree of polymerization of the chains). Connected with the spread in length scales is an even wider spread in time scales: from local bond-length and bond-angle vibrations ($10^{-15} - 10^{-13} \text{ s}$) over conformational transitions between isomeric states in the dihedral potential ($10^{-12} - 10^{-10} \text{ s}$) to the self-diffusion and overall configurational relaxation of the whole chain which for short chains scales as N^2 times the time-scale for conformational transitions and for long chains as $N^{3.4}$ times that time-scale. With N ranging from 100 (short chains) up to 10000 and more (long chains) these time scales cover 10^{-8} s up to 10 s.

The local dynamics is naturally strongly dependent on the exact chemical nature and structure of the polymer one studies. The large scale dynamics, however, is largely universal and is described with the Rouse model whereas for longer chains the tube model and reptation concept is employed to describe the chain dynamics² (see also the contribution of K. Kremer in this volume). It is easy to see that no single simulation method can capture the physics of polymer dynamics on all these length and time scales³. To study the large scale dynamics in polymer melts therefore, simulations of coarse-grained models - both Molecular Dynamics (MD) and Monte Carlo (MC) methods were used - have been successfully applied. On large length and time scales entropy dominates the physical behavior of polymer melts giving rise to universality and therefore every model of a linear polymer faithfully representing excluded volume between monomers as well as connectivity of the chains can be used in computer simulation studies.

On the local scale, however, the detailed chemistry and energetics dominates structure as well as relaxation behavior. For situations where we can ignore quantum effects (which can, however, be important in polymer crystals⁴) MD simulations of chemically realistic force fields are the method of choice to study local relaxation.

Besides the necessity for carefully optimized chemically realistic force fields there is always the question to address whether the simulation is able to equilibrate the model system at a given thermodynamic state point of temperature, density and chain length. This means that the simulations mostly are limited to polymer melts well above the glass transition temperature (or density) and to chains not exceeding the entanglement molecular

weight.

In the next section we will introduce a simple chemically realistic force field for 1,4-polybutadiene (PB) which we will use as our model system for the discussions in this contribution. As this polymer is structurally simple and since there are no relevant partial charges or specific interactions present in this polymer melt, this polymer can be quantitatively described by a relatively simple force field and efficiently simulated using MD algorithms.

In section 2 we will then discuss the question of equilibration of such a system and methods developed for efficient equilibration in a computer simulation. Section 3 will then introduce some techniques for studying local relaxation processes in polymer melts and describe how to implement these measurements in a computer simulation. From there we will go in section 4 to the large scale dynamics in polymer melts as observable in chemically realistic simulations and will finish with some conclusions in section 5.

1 A Simple Chemically Realistic Force Field for PB

Chemically realistic force fields for polymers in general have a complicated functional form

$$\begin{aligned}
 H(\{\mathbf{r}_i\}) = & \sum_i U(l_i) + \sum_j U(\vartheta_j) + \sum_k U(\phi_k) \\
 & + \frac{1}{2} \sum_{i,j} U_{\text{disp}}(r_{ij}) + \frac{1}{2} \sum_{i,j} U_{\text{Coul}}(r_{ij}) + U_{\text{Hbond}} \\
 & + \text{cross terms between all intramolecular potentials} + \dots
 \end{aligned} \tag{1}$$

The first line contains the intramolecular potentials for bond lengths (harmonic), bond angles (harmonic) and dihedral angles, the second line contains the non-bonded interactions due to dispersion forces, Coulomb interactions and hydrogen bonds and the third line contains possible cross terms between all the intramolecular degrees of freedom. There are force fields in the literature containing easily over 30 parameters, able to fit about the same amount of experimental information.

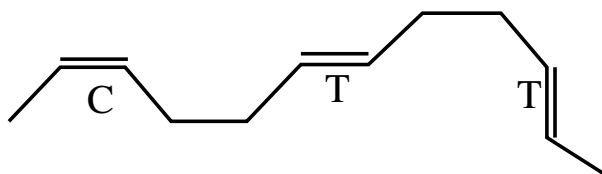


Figure 1. Schematic representation of the chemical micro structure of PB indicating two types of repeat units, cis and trans conformers.

To really understand the necessary ingredients of such a force-field, for instance to be able to reproduce experiments on polymer melts, one best starts with very simple polymers like polyethylene⁵⁻⁹, polyisoprene^{10,11}, polyisobutylene¹² or polybutadiene¹³, to give just a few selected examples. In the following we will focus on the force field used for PB¹³,

see Fig.1 for a schematic representation of the chemical micro structure of the chains. Since we are not interested in very high frequency vibrational motion, we will first of all collapse the hydrogen atoms onto the carbons they are bound to, defining a united atom model, and furthermore constrain the carbon-carbon bonds to their mechanical equilibrium values. The non-bonded interactions for this polymer contain no Coulomb contribution and no hydrogen bonding (making a united atom approach possible). The force-field we are employing therefore has the form

$$H(\{\mathbf{r}_i\}) = \sum_j k_j^\vartheta (\cos(\vartheta_j) - \cos(\vartheta_0))^2 + \sum_j \sum_{n=1}^6 k_{jn}^\phi (1 - \cos(n\phi_j)) + \frac{1}{2} \sum_{i,j} \varepsilon_{\alpha\beta} \left[\left(\frac{\sigma_{\alpha\beta}}{r_{ij}} \right)^{12} - 2 \left(\frac{\sigma_{\alpha\beta}}{r_{ij}} \right)^6 \right]. \quad (2)$$

The constants k_j^ϑ and k_{jn}^ϕ depend on the type of bond or torsion angle, and $\varepsilon_{\alpha\beta}$ and $\sigma_{\alpha\beta}$ depend on the types of interacting united atom group (CH, CH₂ or CH₃).

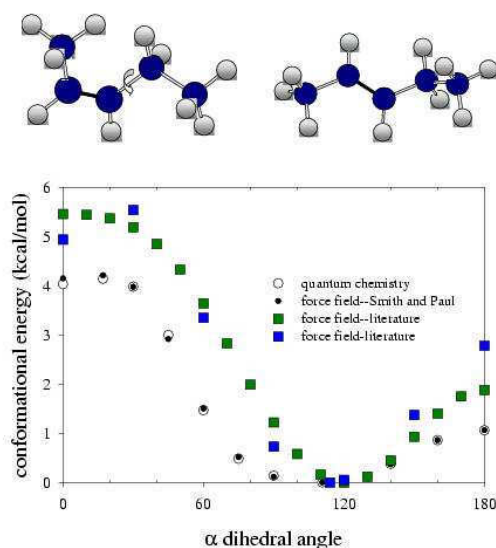


Figure 2. Effective potential (containing a non-bonded LJ contribution) for rotation around the alpha-bond next to the double bond (black bond in the cis-pentene molecule) in a cis group. The new force field is compared to literature force fields^{14,15}.

The parameters for the bond angle potential can be obtained from quantum chemistry and from spectroscopic information, for the accurate determination of the parameters of the dihedral potential, however, quantum chemistry is the only tool. Only quantum chemical calculations yield information on the barriers between different isomeric states in the dihedral potentials. This information is crucial for an accurate modeling of local conformational transitions between isomeric states, since their rate is exponentially sensitive to

the barriers¹³. And it is these transitions, as we will see in the following, which underly all local relaxation processes in a polymer melt and in general set the microscopic time scale (Rouse rate, effective segmental friction) for the large scale relaxation processes. Even more difficult to determine than the torsion force constants are the parameters of the dispersion forces. Equation of state and cohesive energy data are one source of experimental information these parameters can be gouged on, but much more sensitive are phase diagrams of small oligomers of the polymer under study⁷.

These chemically realistic models can be simulated using MC as well as MD techniques. Advanced MC techniques which address the equilibration of these model systems will be discussed in the next section. When we want to study local relaxation processes we are bound to use MD simulation techniques which faithfully capture the short time ballistic and vibrational motion influenced by inertia effects. The quantitative reproduction of local relaxation processes needs a realistic dynamics on the fs to ns time scales. Typically one uses Verlet integrators in the NVE ensemble or their second order Runge Kutta counterpart when one switches to the NVT ensemble by using for instance a Nosé-Hoover thermostat. Although it can not be in general established that this change of ensemble has no influence on the relaxation behavior, it was empirically observed to have no discernible quantitative effect on the local relaxation behavior. The effectiveness of the MD code for these simulations can be increased using multiple time-step techniques¹⁶.

2 Equilibration of Polymer Melts

In this section we want to address the question of equilibration of chemically realistic models of polymer melts. In addition to the methods discussed already in the contribution of K. Kremer on the study of reptation dynamics in polymer melts we will now discuss some advanced MC techniques developed for equilibration of dense polymer systems.

Basically, a discussion of the problems with the equilibration of polymer melt models can be divided into identifying the two main sources for slow relaxation in polymers, entanglement effects and the glass transition. The first one is entropic in origin, the second one - at least in chemically realistic polymer models - to a large degree enthalpic. We write the largest relaxation time in the melt as

$$\tau_1(T, N) = \tau_{\text{mes}}(T)N^x \quad (3)$$

where τ_{mes} is a mesoscopic time scale. The chain length dependence crosses over from $x = 2$ for Rouse behavior to $x = 3.4$ for reptating chains. Every simulation method performing configuration changes typical for the mesoscopic time scale τ_{mes} , i.e., local rearrangements, leads to a relaxation of the large scale structure of the polymer chains in the melt only after $O(N^x)$ such configuration changes, quickly limiting the range of chain lengths one can treat in thermodynamic equilibrium. To circumvent this problem one has to use advanced MC techniques implementing global configurational changes within a single Monte Carlo step.

A class of these advanced MC techniques is comprised of the so-called connectivity altering moves like the cooperative motion algorithm¹⁷ on the lattice and the end-bridging algorithm^{18,19} and its newest variant, the double-bridging algorithm²⁰. The latter two have been developed with chemically realistic polymer models in mind and we will now shortly discuss the idea and properties of these algorithms.

In the original end-bridging algorithm^{18,19} an end monomer i of one chain in the melt attacks a backbone atom j of another chain in sufficiently close proximity and tries to initiate a change in connectivity of the two involved chains by forming a trimer bridge to this backbone atom. In the event of a successful bridging the attacking chain grows by a part cut off the attacked chain and the attacked chain shrinks correspondingly. This description already exhibits the main drawbacks of the algorithm: it generates polydisperse polymer melts and it needs a sufficient amount of chain ends to be efficient. Empirically it was found that the efficiency of the algorithm dropped considerably as the stiffness of the chains was increased and in the presence of chain orientation. The fact that the algorithm generates polydisperse samples can be used to simulate the experimentally given molecular weight distribution through the so-called $NpT\mu^*$ ensemble where number of chains, number of monomers, pressure, temperature and the relative chemical potential of the different chain lengths are prescribed¹⁸. The latter thermodynamic constraints generate the desired molecular weight distribution. In this formulation the algorithm was successfully applied to, e.g., polyethylene melts²¹ and cis-1,4 polyisoprene melts¹⁰.

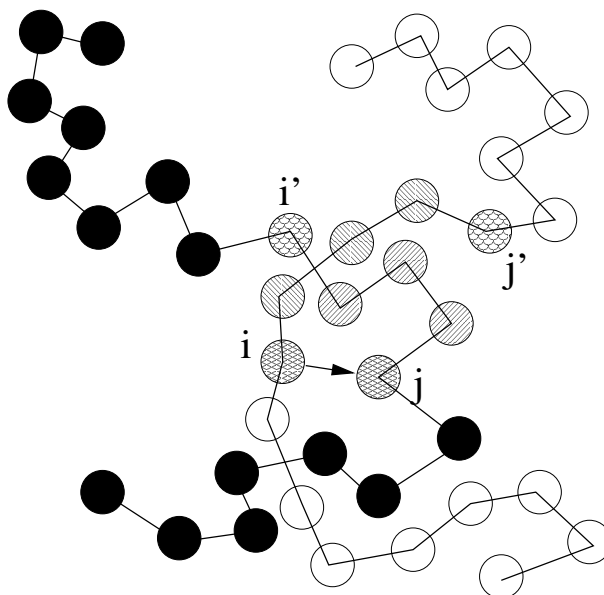


Figure 3. Monomer i of one chain attacks monomer j of a different chain. Only when the shaded trimers are cut off and bridges between i and j and i' and j' are formed, are the resulting chains of the same length as the original ones.

In the double-bridging algorithm an inner monomer of a chain attacks an inner monomer of another chain (or the same chain) and tries to form a trimer bridge (this is shown in Fig.3). Simultaneously another bridge is formed between two monomers which are 4 steps apart from the first two monomers along the two chains generating two new chains with exactly the same length as the original one. These requirements understandably put heavy geometric constraints on the configurations of the two chains for which

this type of move is feasible since only special choices of involved monomers (i, j) and excised trimers (on which side of i or j , respectively) conserve the chain lengths in the move. The trimer bridge itself is made of three monomers (atoms) connected by bonds of fixed lengths l making a fixed angle, which is chosen as the value at the maximum of the bond-angle distribution of the model. Whenever the monomers i and j have a distance less than the maximum bridgeable distance of $4l \cos((\pi - \theta_{\max})/2)$ this geometric problem is in principle solvable. Wu and Deem²² have shown that it is equivalent to finding the solutions of an algebraic equation of order 16. Since each trimer bridging problem of the double-bridging algorithm therefore has between 0 and 16 solutions, the double bridging problem has between 0 and 256 solutions. To reduce this number Karayiannis et al.²⁰ suggest to discard directly solutions leading to a too high torsional energy or overlap between segments when hard spheres of diameter $\sigma = 3.2\text{\AA}$ are placed at the positions of the bridging atoms. From the remaining s energetically viable solutions, one is then chosen according to a Boltzmann factor involving the torsional and non-bonded energies

$$P_{\text{select}}(\text{old} \rightarrow \text{new}) = \frac{\exp\{-\beta(V_{\text{tor}}(\text{new}) + V_{\text{LJ}}(\text{new}))\}}{\sum_{k=1}^{s_{\text{old} \rightarrow \text{new}}} \exp\{-\beta(V_{\text{tor}}(k) + V_{\text{LJ}}(k))\}}. \quad (4)$$

Here $\beta = 1/k_{\text{B}}T$. The reverse move is selected according to

$$P_{\text{select}}(\text{new} \rightarrow \text{old}) = \frac{\exp\{-\beta(V_{\text{tor}}(\text{old}) + V_{\text{LJ}}(\text{old}))\}}{\sum_{k'=1}^{s_{\text{new} \rightarrow \text{old}}} \exp\{-\beta(V_{\text{tor}}(k') + V_{\text{LJ}}(k'))\}}. \quad (5)$$

Furthermore, the selection of a given bridge site is done by choosing from a list of bridgeable sites from atom i , $N_{\text{DB}}(i)$, which is constantly updated in the simulation.

$$\begin{aligned} W(\text{old} \rightarrow \text{new}) &= \frac{1}{[N_{\text{DB}}(i)]_{\text{old}}} \\ W(\text{new} \rightarrow \text{old}) &= \frac{1}{[N_{\text{DB}}(j)]_{\text{new}}}. \end{aligned} \quad (6)$$

The Metropolis acceptance criterion for the double-bridging move is therefore

$$P_{\text{accept}}(\text{old} \rightarrow \text{new}) = \min \left[1, \frac{W(\text{new} \rightarrow \text{old})P_{\text{select}}(\text{new} \rightarrow \text{old})J(\text{new})\exp\{-\beta V(\text{new})\}}{W(\text{old} \rightarrow \text{new})P_{\text{select}}(\text{old} \rightarrow \text{new})J(\text{old})\exp\{-\beta V(\text{old})\}} \right] \quad (7)$$

where $J(\text{new})$ and $J(\text{old})$ are Jacobian determinants for the bridges degrees of freedom in the initial and final stage²⁰.

This algorithm was found to be especially efficient to relax the large scale conformation of the chains as observed typically through end-to-end vector decorrelation and center of mass diffusion. There are two at first glance counterintuitive observations connected with these relaxation functions in simulations with these connectivity changing algorithms: the center of mass of a chain diffuses faster than a single monomer and longer chains diffuse faster (have a faster end-to-end vector decorrelation) than shorter chains as is shown in Fig. 4. Both of them, however, can be understood by a simple random walk type argument as for instance presented in²³ or along slightly different lines in the following paragraph.

Consider for the double-bridging algorithm monomer n , $3 \leq n \leq N/2 - 3$ attacking another chain successfully. The change in center of mass position of the chain n belongs

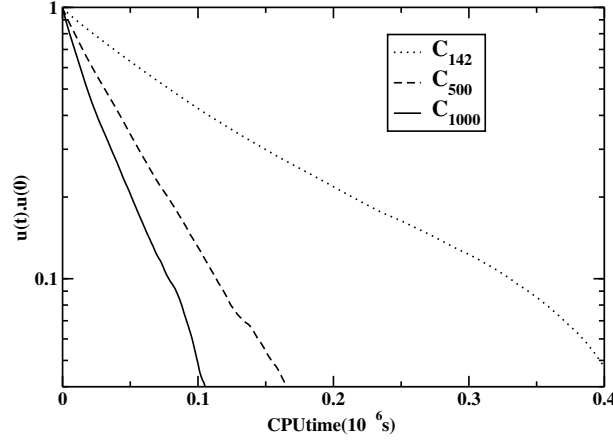


Figure 4. Autocorrelation function of the end-to-end unit vector for united atom model chains of alkanes of chain lengths $N = 142$ (dotted curve), $N = 500$ (dashed curve) and $N = 1000$ (full curve). One observes a clear decrease of the autocorrelation time scale with increasing chain length for the double-bridging algorithm. (I am grateful to the authors of²⁰ for supplying these data.)

to is given as

$$\Delta \mathbf{R}_{\text{cm}} = \frac{1}{N} \sum_{i=n+1}^N (\mathbf{r}_i^{\text{new}} - \mathbf{r}_i^{\text{old}}).$$

As $\mathbf{r}_i^{\text{new}}$ and $\mathbf{r}_i^{\text{old}}$ can be considered as random walks starting at the same position in space, we can write for the scaling of their difference

$$\Delta \mathbf{r}_i = \mathbf{r}_i^{\text{new}} - \mathbf{r}_i^{\text{old}} \approx \sqrt{i} \Delta \hat{\mathbf{r}}_i$$

where $\hat{\mathbf{r}}_i$ is indicating a unit vector. From this we get for a single successful double-bridging move

$$\langle \Delta \mathbf{R}_{\text{cm}}^2 \rangle_n \approx \frac{1}{N^2} \sum_{i,j=n+1}^N \sqrt{ij} \langle \Delta \hat{\mathbf{r}}_i \cdot \Delta \hat{\mathbf{r}}_j \rangle.$$

When we approximate $\Delta \hat{\mathbf{r}}_i$ and $\Delta \hat{\mathbf{r}}_j$ as being uncorrelated this reduces to

$$\langle \Delta \mathbf{R}_{\text{cm}}^2 \rangle_n \approx \frac{\langle \Delta \hat{\mathbf{r}}_i^2 \rangle}{2} \left(1 - \frac{n(n+1)}{N^2} \right).$$

Averaging this finally over the position of the attacking monomer yields

$$\langle \Delta \mathbf{R}_{\text{cm}}^2 \rangle = \Delta R_0^2 \left(1 - \frac{a_1}{N} + \frac{a_2}{N^2} \right). \quad (8)$$

The center of mass of the polymers therefore will perform a random walk with an average step length which increases up to a saturation value when the chain length N goes to infinity. In a similar fashion one can calculate the amount of decorrelation of the unit

vector along the end-to-end vector of a chain in a single successful double-bridging move and arrive at

$$\Delta\phi(t) = -c_0 \left(1 - \frac{c_1}{N}\right). \quad (9)$$

This behavior gives rise to an increase of the measured autocorrelation time for the end-to-end vector with decreasing chain length. Qualitatively this is observed in the simulations²⁰ as can be seen in Fig. 4.

$$\tau_N = \tau_\infty \left(1 + \frac{c_1}{N}\right). \quad (10)$$

When one performs an exponential fit to the short time decay observable in figure 4 and

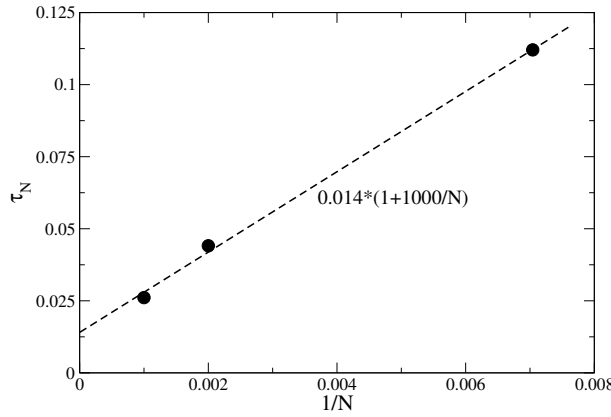


Figure 5. Autocorrelation time for the end-to-end vector reorientation in figure 4 versus inverse chain length. The predicted scaling is nicely compatible with the data.

then plots the simulation time per chain (note that for Fig. 4 the number of monomers was fixed) as a function of inverse chain length, the scaling of equ. (10) is clearly observed (see Fig. 5).

We already commented on the fact that the connectivity changing algorithms are especially efficient in decorrelating large scale structure in the melt. They are typically augmented by local moves and reptation moves to equilibrate the local structure. Before we turn to a detailed discussion of the local relaxation and the energetics involved in this for a chemically realistic polymer melt we want to mention an alternative method for overcoming the entropic slowing down in a polymer melt. This method has so far been tested on the structural relaxation of a collapsed polymer globule where the connectivity of the chains and the high density inside the globule lead to a dramatic increase in the structural relaxation time of the globule. The idea of this algorithm²⁴ is to turn part of the monomers into ghost particles (alternatively one can think of removing some particles into the fourth dimension similar to desorbing and readsorbing particles from a two dimensional film into the third dimension) and forcing them back into the three dimensional structure by applying

an external field. One can formulate this algorithm in the context of a parallel tempering simulation of an extended system Hamiltonian

$$H = H_0 + \sum_{i=1}^{N_{\text{part}}} h x_4(i),$$

where H_0 is the real Hamiltonian of the model system, the external field h is set to different values labeling the replicas of the real system, and the fourth coordinate $x_4(i)$ of monomer i reduced to being either 0 or 1. The partition function of the extended ensemble in a parallel tempering simulation therefore reads

$$Z = \prod_h \sum_{\{c\}} \exp \left\{ -\frac{H_0 + \sum_{i=1}^{N_{\text{part}}} h x_4(i)}{k_{\text{B}} T} \right\}, \quad (11)$$

where $\sum_{\{c\}}$ indicates a sum over all configurations at fixed h . The external field h determines the average number of ghost particles. For $h = 0$ half the particles are ghost particles and for $h \rightarrow \infty$ all particles are back in the original three-dimensional system. In this way excluded volume as well as connectivity constraints are partially removed. In the application to the equilibration of a collapsed polymer globule it became possible for the first time to identify a liquid to solid transition in the globular phase. In principle, this scheme is applicable to coarse-grained and chemically realistic continuum models as well, although its efficiency will depend on whether packing and connectivity or local energetics are the dominating factor for the slow-down of relaxation processes. The latter will be discussed in detail in the next section.

3 Local Relaxation in Polymer Melts

In the last section we dealt with the chain length dependence of the longest relaxation time in a polymer melt and with advanced MC algorithms to overcome this slowing down. We did not dwell on the prefactor $\tau_{\text{mes}}(T)$ and its temperature dependence which typically spans 14 orders of magnitude when we study a glass forming polymer melt above T_g . In order to understand the physical origin of this slowing down and the relative importance of energetic effects (intramolecular dihedral potential) versus packing effects (which are entropic in origin) one has to use chemically realistic modeling. We will present in this section MD simulations of a random copolymer of 50 % trans 1,4-PB, 40 % cis 1,4-PB and 10% vinyl groups. This polymer was synthesized with an average molecular weight corresponding to about 30 repeat units and characterized to possess the above chemical micro structure¹³. We have 40 chains of 30 repeat units in our simulation box and we will be using a united atom model for the CH, CH₂ and CH₃ groups. With a total of about 5000 united atoms, these systems can be simulated today over a time range of several 100 ns (about 2 ns of real time trajectory per day on a single processor).

The ability of the simulation to quantitatively reproduce experimental data on relaxation processes in polymer melts strongly rests on the implementation of the correct force fields for the dihedral angles. These typically possess barriers separating the isomeric states as was discussed in the section on the force field, and it is by correlated jumps over these barriers^{25,5,6} that the local relaxation processes come about. The frequency of these

jumps is exponentially sensitive to the height of the barriers. In order to be able to make a parameter free quantitative comparison with experiment one therefore needs to carefully determine this part of the force field¹³, which is done through high-level quantum chemistry calculations. The validation of this force field then proceeds through comparison with experiment.

3.1 NMR Experiments

An experimental technique which is very sensitive to the local dihedral barriers is C^{13} NMR spin lattice relaxation time measurements.

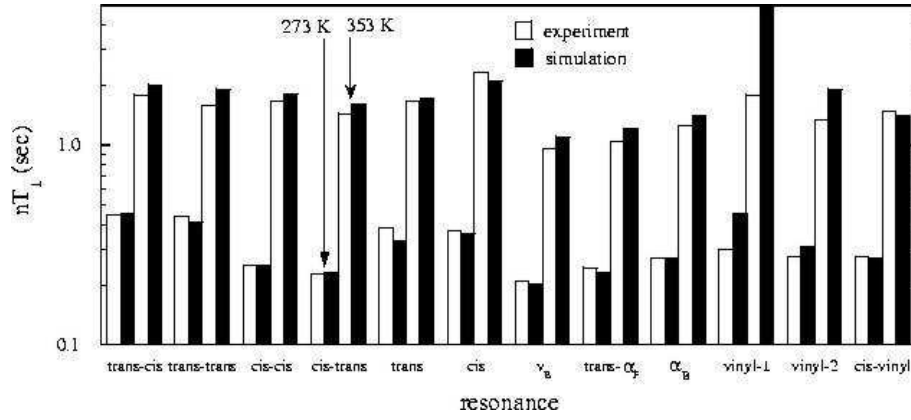


Figure 6. Comparison of the spin lattice relaxation times determined from MD simulations and experiment for two different temperatures. The figure shows data for 12 different resonances (cis-cis for example indicates an sp^2 carbon in a cis group next to another cis group, trans an sp^3 carbon of a trans group)

For polymers like PB this technique observes the reorientational motion of the CH bonds as quantified through the second Legendre polynomial of the CH bond orientational autocorrelation function. It is sensitive to the local chemical environment of the relaxing bond giving rise to 12 different resonances which can be identified²⁶. A CH bond at the sp^2 carbon of the double bond in a cis group next to a trans group (see Fig. 1) relaxes differently from one next to another cis group and differently from the ones in a trans group and differently from the CH bonds at a sp^3 carbon. The vinyl group alone gives rise to six different resonances.

In a computer simulation one measures the CH vector autocorrelation function and determines the second Legendre polynomial according to

$$P_2^{CH}(t) = \frac{1}{2} \left(3 \langle [e_{CH}(t) \cdot e_{CH}(0)]^2 \rangle - 1 \right) \quad (12)$$

From this function one obtains the spectral density by Fourier transformation

$$J(\omega) = \frac{1}{2} \int_{-\infty}^{\infty} P_2^{CH}(t) e^{i\omega t} dt. \quad (13)$$

The spin-lattice relaxation time finally is given by evaluating the spectral density at the Larmor frequencies of carbon and hydrogen atoms

$$\frac{1}{nT_1} = K [J(\omega_H - \omega_C) + 3J(\omega_C) + 6J(\omega_H + \omega_C)] \quad (14)$$

where ω_H, ω_C are Larmor frequencies, K is a constant depending on the hybridization of the nucleus and n is the number of hydrogen atoms attached to the carbon atom under study.

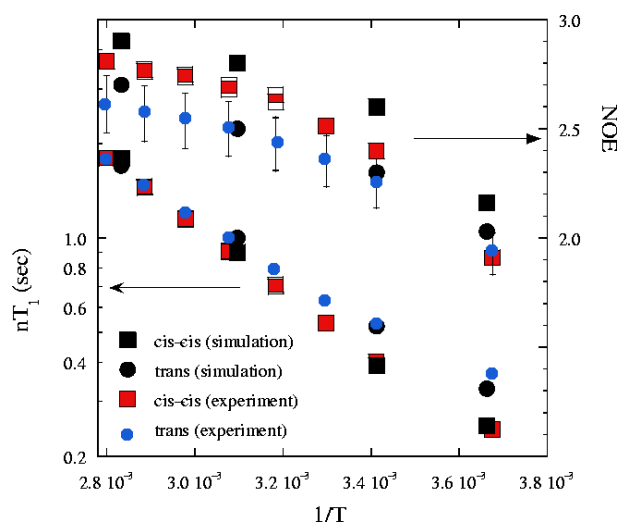


Figure 7. Comparison of the spin lattice relaxation times determined from MD simulations and experiment for the cis-cis (sp^2) and the trans (sp^3) resonance as a function of inverse temperature. Also shown are results for the nuclear Overhauser enhancement which is a measure of the non-exponentiality of the observed relaxation

To be able to evaluate this function from a stored trajectory of a simulation of a united atom model, we have to reinsert hydrogen atoms into the system. This is done using the information on equilibrium ($T = 0$) bond lengths and angles which allows the determination of the hydrogen positions from the positions of the carbon backbone atoms (united atoms) alone. In Fig 6 we show a comparison between experiment and simulation for the 12 resolvable resonances for two different temperatures. As one can see, there is a very good agreement (better than within 10%) between simulation and experiment for all but one resonance. The one that is not reproduced is a rigid body rotation of the side group which does not lead to a conformation relaxation and which we therefore did not try to match more accurately. The temperature dependence of the spin lattice relaxation time for two resonances can be seen in Fig. 7

The spin lattice relaxation time decreases when the autocorrelation time for the observed relaxation increases. For high temperatures, where the nuclear Overhauser enhancement is still close to three²⁷, the spin lattice relaxation time traces the temperature dependence of the autocorrelation time of the torsional transitions and the agreement between simulation and experiment is a quantitative validation of the torsional force field²⁶.

From comparing to the observed torsional autocorrelation function in the simulation one can learn that this identification of the physical motion seen in the NMR experiment breaks down below about 300 K. In Fig. 8 we compare different measures of the local orientational

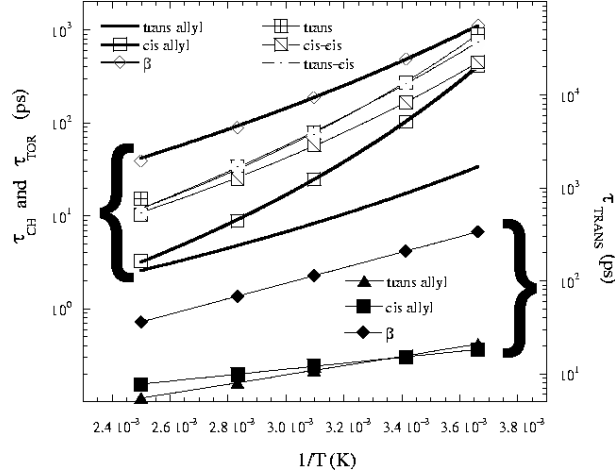


Figure 8. Temperature dependence of several measures of the local orientational mobility of the chain. The lower set of curves pertaining to the right abscissa shows the mean time between torsional transitions for the three relevant torsional potentials along the chain. The upper two sets of curves give the integrated autocorrelation time for the second Legendre polynomial of the CH vector orientation, τ_{CH} , and the integrated autocorrelation time for the torsion angle autocorrelation function, τ_{TOR} .

mobility of the chains. The most basic one is the mean time between torsional transitions which we determined with a time resolution of 1 ps along the simulated trajectory. This time scale shows an Arrhenius temperature dependence. It is different for the three relevant torsion potentials along the chain, the allyl bonds next to a cis or trans double bond and the β bond at the connection between two PB monomers. All these time scales, however, simply follow an Arrhenius law with their energetic dihedral barrier as activation energy in the observed temperature range, so that there are no packing effects observable on the dihedral dynamics in this range. Contrary to the mean time between torsional transitions, the autocorrelation time for the torsions (thick lines in Fig. 8) defined as the time integral of the following correlation function

$$f(t) = \frac{\langle \cos(\phi(t)) \cos(\phi(0)) \rangle - \langle \cos(\phi(0)) \rangle^2}{\langle \cos(\phi(0))^2 \rangle - \langle \cos(\phi(0)) \rangle^2}, \quad (15)$$

increases in a Vogel-Fulcher like manner. This shows the growing importance of back-jump correlations for the torsional transitions: a torsional transition which occurs with its Arrhenius rate is immediately reversed with an increasing probability upon reducing the temperature, thus leading to no decorrelation of the torsional angle. The final set of curves denoted by their resonance name in the figure legend gives the autocorrelation time of the second Legendre polynomial for the CH vectors, i.e., the time integral of $P_2(t)$ in equ. (12). These time scales pick up contributions from different torsions adjacent to the carbon

atom under study and lie intermediate to the autocorrelation functions of these torsions. Another technique which is taken as a measure of local relaxation processes in polymer melts is dielectric relaxation which we will discuss in the following.

3.2 Dielectric Relaxation Experiments

Dielectric relaxation measurements couple to the dynamics of the dipole moment of the sample. The dielectric permittivity is the Fourier-Laplace transform of the dipole moment autocorrelation function.

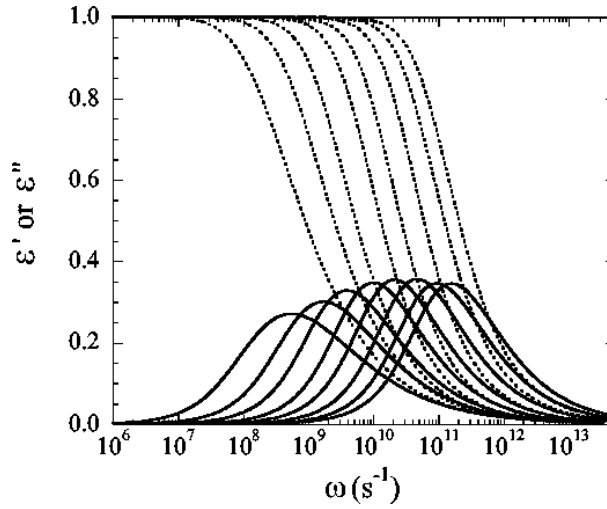


Figure 9. Real and imaginary part of the dielectric permittivity as a function of frequency.

$$\frac{\epsilon' + i\epsilon''}{\Delta\epsilon} = 1 - i\omega \int_0^\infty \Phi(t) e^{-i\omega t} dt \quad (16)$$

with $\Phi(t)$ defined as

$$\Phi(t) = \frac{\langle \mathbf{M}(t) \cdot \mathbf{M}(0) \rangle}{\langle \mathbf{M}(0) \cdot \mathbf{M}(0) \rangle}.$$

Furthermore we have used the following relation for the dielectric relaxation strength

$$\Delta\epsilon = \epsilon_r - \epsilon_u = \left(\frac{\epsilon_u + 2}{3} \right)^2 \frac{3\epsilon_r}{2\epsilon_r + \epsilon_u} \frac{\langle M(0)^2 \rangle}{3V\epsilon_0 k_B T}$$

where ϵ_u and ϵ_r are the unrelaxed and relaxed dielectric constants of the material, respectively. The dipole moment of the box is the sum of local dipole moments along the chains so that $\Phi(t)$ in principle contains correlations between all local dipole moments, intra-molecular and inter-molecular. A comparison with chemically realistic MD simulations can help to identify the independently relaxing dipoles observed in the experiment.

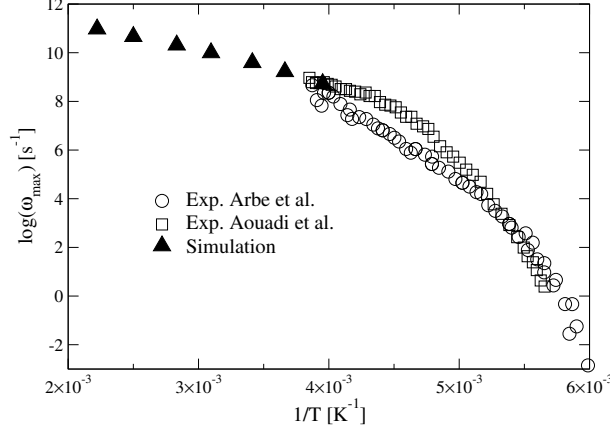


Figure 10. Position of the peak in the dielectric loss as a function of temperature for the simulation (filled triangles) and two sets of experimental data (open circles²⁸ and open squares²⁹). The experiments show some scatter, but the in the overlap temperature regime experiment and simulation agree.

For PB the dipole moments are located on the vinyl side group and the cis group. The partial charges, however, are small and that was the reason we did not include Coulomb forces into our MD simulation. Consequently, there can be no correlations due to the local dipole moment occurring in our simulation trajectory. We then have to see, how well our results for the dielectric permittivity agree with the experiments²⁷. To this end we reinsert partial charges into the simulated trajectory. We know that we have to find

$$\langle M^2(0) \rangle = N_{\text{chains}} \langle M_{\text{chain}}^2(0) \rangle$$

and we also find that the autocorrelation function of the box dipole moment can be calculated through the autocorrelation function of the chain dipole moment

$$\Phi(t) = \Phi_{\text{chain}}(t)$$

which has a 40 times better statistics. In Fig. 9 we show the real and imaginary part of the dielectric permittivity as obtained from the simulation. The frequencies of maximum loss can now be compared with experimental results. Experimentally the maximum loss frequency is typically measured for lower temperatures^{28,29} to study the temperature dependence of the structural glass transition or α process. Two sets of experiments in the literature show some discrepancies in an intermediate temperature range but agree with each other and with our simulation data at higher temperatures where we have an overlap temperature window between simulation and experiment. From this we can conclude, that also in the experiment one sees no correlations between the dipole moments of different chains. When we furthermore compare the time scale given by the maximum loss frequency with the time scales of the Rouse modes for our chains we can obtain from the simulation, we can say that the dielectric measurements on PB see the relaxation of a chain segment of about 6 backbone carbons, which is exactly the length of a statistical segment of the chains.

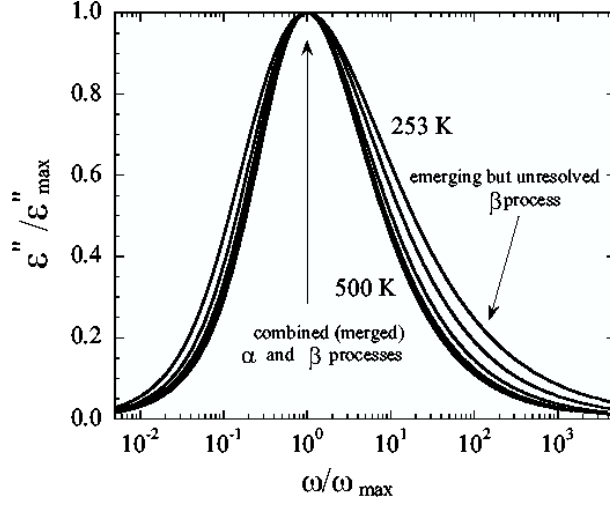


Figure 11. Time temperature superposition applied to the dielectric loss data obtained from the computer simulation in a temperature range from 500 K down to 253 K.

Many glass-forming materials show two distinct relaxation processes at low temperatures in their dielectric spectrum, called the α -process - which shows a Vogel-Fulcher-Tamman temperature dependence - and the β -process - which shows an Arrhenius temperature dependence. In the temperature regime accessible to our MD simulations so far we are actually observing the so-called combined $\alpha - \beta$ process, i.e., there has been no separation of time scales yet between the two processes which can be observed at lower temperatures. For a single type of relaxation mechanism underlying the dielectric spectrum we should be able to see time-temperature superposition if we scale the dielectric loss data in Fig. 9 by the peak position and peak height, respectively, which is shown in Fig. 11.

Our results in Fig. 11 show that the time temperature superposition is only borne out for temperatures above 300 K. Below this temperature the scaling breaks down, indicating a change in the molecular motion mechanisms underlying the dielectric data observed. We interpret this as the emergence of the β process which becomes clearly resolvable as an independent feature in the loss spectrum only at lower temperatures.

Let us now come to the question of increasing heterogeneity in the local mobilities upon decreasing the temperature. We have already identified a tendency for immediate back jumps after one torsional transition as the reason for the different temperature dependencies of the mean waiting time between torsional transitions $\langle t_{\text{wait}} \rangle$ and the torsional autocorrelation time. In a homogeneous system, where every chemically identical torsion shows identical dynamics on the time scales of observation, the probability distribution of waiting times should be Poissonian. As we can see in Fig. 12 which shows the waiting time distribution for the 10th jump to occur for temperatures $T = 500$ K, 450 K, 400 K, 353 K, 323 K, 293 K, 273 K and 253 K, at high temperatures the curves approach the Poisson distribution indicated by the fat line in the figure. At lower temperatures two features emerge. There is an increasing probability of very short waiting times which captures the increasing amount of direct back jumps for the dihedral transitions. Furthermore the distribution

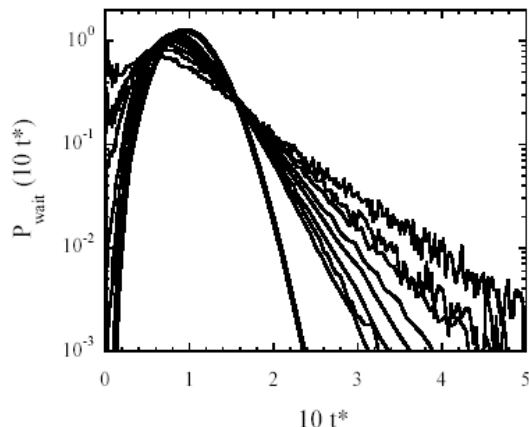


Figure 12. Distribution of waiting times for a total of 10 torsional transitions per dihedral degree of freedom to occur plotted versus 10 times $t/\langle t_{\text{wait}} \rangle$.

develops a fat tail toward very long waiting times indicating the increasing heterogeneity between the different torsion angles. For high temperatures this heterogeneity will also show up on the distribution for the average waiting time until the next jump occurs, but on the scale of 10 jumps the heterogeneity at high temperatures has been smeared out: slow angles have had time to become fast and vice versa. At low temperatures this homogenization time becomes much longer than the time scales observed for the jump distribution and the two extremes of the mobility distribution show up clearly. When we characterize the distribution of waiting times for several values of n , the number of transitions observed, by their normalized dispersion

$$D(n, T) = \sigma^2(n, T)/n$$

this means that at fixed temperature $D(n, T)$ becomes smaller with increasing n approaching the Poisson limit of $1/n$. For fixed n , $D(n, t)$ increases dramatically with decreasing temperature due to the fat tails developing in the waiting time distribution. We can define a measure for the homogenization time of the dihedral mobility by determining the number of jumps $n^*(T)$ (this will be a real value determined by interpolation) after which $D(n, T)$ reaches a fixed value, for example $D(n^*, T) = 1$. This corresponds to a time scale $t^*(T) = n^*(T)\langle t_{\text{wait}} \rangle(T)$. The temperature dependence of this time scale is compared in Fig. 13 with that of the average waiting time and the torsional autocorrelation time. It is clear from this comparison that the torsional autocorrelation time is sensitive to this homogenization process of the torsional mobilities and that the slowing down of this homogenization leads to the stronger than Arrhenius-like increase in relaxation time as compared to the average waiting time between transitions.

3.3 Changing the Model Hamiltonian

So far we have presented MD simulations of a chemically realistic united atom model for PB employing a carefully validated quantum chemistry based force-field. For the following

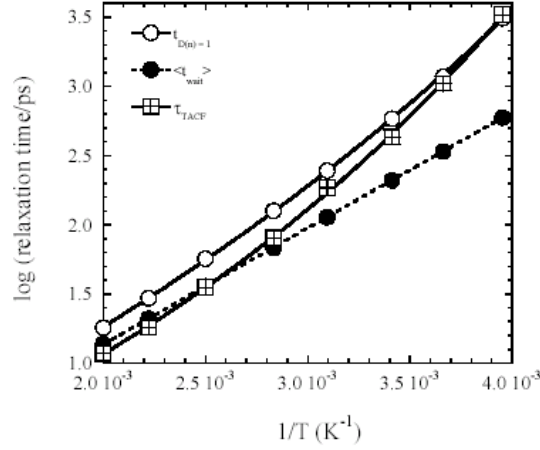


Figure 13. Homogenization time scale for the dihedral mobilities compared to the average waiting time and the dihedral angle autocorrelation time.

study we will slightly modify the chemically realistic model to one without vinyl side groups, i.e. it will be random copolymer of 55 % trans 1,4-PB and 45 % cis 1,4-PB. We will again have 40 chains of now 29 repeat units (115 backbone carbons) in our simulation box and will be using the same united atom force field. We have seen that we can understand the physics of local relaxation processes observed in NMR and dielectric experiments with an understanding of the dihedral potentials along the chain. But how important is this enthalpic effect compared to the entropic packing effects? To resolve this question is a necessary prerequisite to understanding the glass transition in real polymer melts as well as the dynamic response of polymer blends. To bring out the effects of the torsional barriers on the local dynamics of the atoms most dramatically, we use the ability of the simulation approach to modify parts of the force field selectively. We will compare the results from the chemically realistic chain (CRC) model simulations to those for a freely rotating chain (FRC) version of this CRC model, where we switch off all torsion potentials³⁰. Let us first show that in the case of PB this has no effect on the static structure of the melt.

When we discuss the static structure of an amorphous polymer melt we always have to take into account two different measures of that structure. One of these is the chain structure as described by the single chain structure factor

$$S(q) = \left\langle \frac{1}{N} \sum_{i=1}^N e^{iq \cdot r_i} \right\rangle \quad (17)$$

where the average is meant to include a spherical and a thermal average. This function is shown in Fig. 14 for $T = 240, 273$ and 353 K for the CRC model and for 273 K for the FRC model. The behavior for small momentum transfers agrees well with the Debye function $f_D(x) = \frac{2}{x^4}(e^{-x^2} + x^2 - 1)$ where $x = qR_g$, which describes the scattering of a Gaussian coil. The obtained value for the radius of gyration agrees well with the directly measured value. The single chain structure factor shows no temperature dependence in the

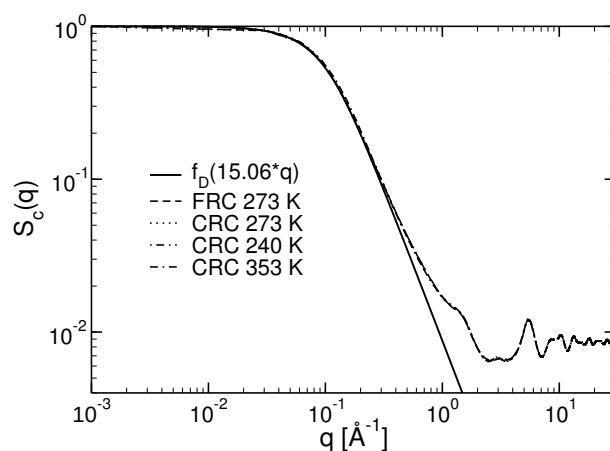


Figure 14. Single chain structure factor for the CRC model at different temperatures indicated in the legend and for the FRC model at 273 K. Also included is the Debye function which fits the behavior for small momentum transfers.

depicted temperature range and agrees perfectly between the FRC and CRC models. This is a peculiarity of PB since for this polymer all minima in the different dihedral potentials are iso-energetic, which also explains the lack of temperature dependence of the single chain structure factor of the CRC model.

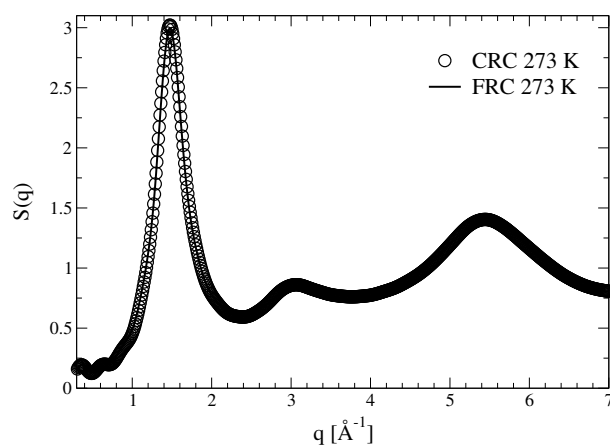


Figure 15. Melt structure factor for PB as obtained from simulations of the CRC model and the FRC model at 273 K.

The other quantity characterizing the melt structure is the liquid structure factor. This is shown over a wide momentum transfer range in Fig.15 at 273 K. For the calculation

we have used the united atoms as scattering centers of equal scattering strength, calculating in this way the structure of the actual simulated systems. A quantitative comparison to the structure factor of PB would be improved by reinserting the hydrogen atoms into their mechanical equilibrium positions⁶ and explicitly using the scattering lengths of the carbon and hydrogen atoms in the system. It is gratifying, that even so the position of the amorphous halo at $q = 1.47 \text{ \AA}^{-1}$ agrees nicely with the experimental results³¹ and also the behavior at higher momentum transfers is comparable. We have performed both sets of simulations at the equilibrium density of the CRC model and this result shows that under these conditions the liquid structure is the same in both models.

In the preceding paragraphs we have shown that we have two models at hand which show the same static structure (packing) on the level of the two-body correlation functions. Do they have the same dynamics? The high temperature behavior of the CRC model

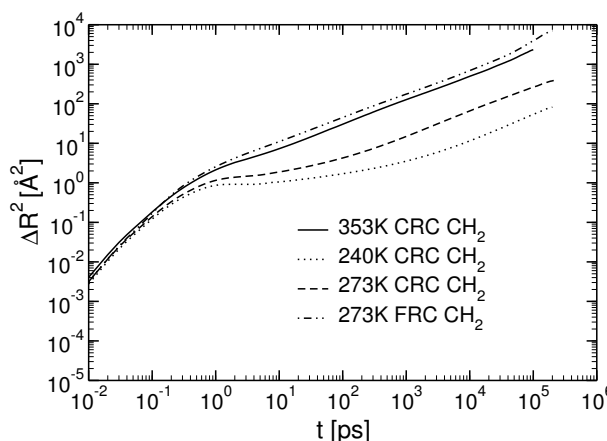


Figure 16. Mean square displacements of the sp^3 carbons along the chain back-bone as a function of time for several temperatures for the CRC model and for $T = 273 \text{ K}$ for the FRC model.

(curve at $T = 353 \text{ K}$ in Fig. 16) and the behavior of the FRC model agree. One observes a crossover from short time ballistic and vibrational motion to a subdiffusive Rouse-like regime determined by the connectivity of the chains. For the CRC model at 273 K , however, one observes a plateau intervening between the short time motion and the Rouse-like regime and this plateau becomes more pronounced upon lowering the temperature to 240 K . It starts at around $t = 1 \text{ ps}$ and extends almost to 100 ps for 240 K . This slowing down is not due to packing effects but obviously due to the presence of intramolecular barriers against dihedral rotation. On the time scale of 1 ps the fast vibrational dynamics of the bond angles and torsion angles is damped out and this time scale is not strongly dependent on temperature. The mean time between torsional transitions, however, as we have seen increases in an Arrhenius-like fashion with decreasing temperature. Consequently we are observing a separation of time scales between the vibrational dynamics and the relaxational dynamics governed by the torsional transitions. At 240 K the mean waiting time between torsional transitions has reached about 100 ps and this is exactly the time scale

of the break-up of the plateau. For shorter times the mean-squared displacement curves only pick up contributions from the fast-moving torsions in the waiting time distribution and upon lowering the temperature these become fewer and fewer (at a fixed time). This result extends to lower temperatures showing that the glass transition in PB is strongly influenced by enthalpic effects and not given by the intermolecular packing³² alone. In the next section we will now focus on the large scale dynamics in chemically realistic polymer melts.

4 Large Scale Relaxation as Observable in Neutron Scattering Experiments

The comparison between neutron scattering experiments and MD simulations can be done on several length and time scales. Comparing to incoherent scattering as for example measured in time of flight experiments^{33–35} one focuses on small length scales (momentum transfers of about one inverse Ångström) and time scales below 1 ns and the motion of individual atoms. This yields complementary information to the one obtained from the comparison with NMR experiments.

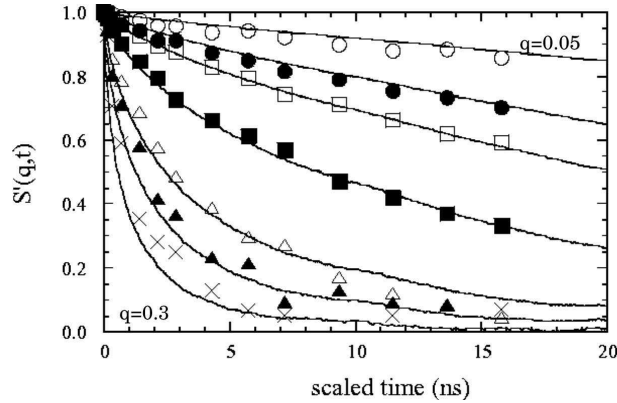


Figure 17. Single chain coherent intermediate scattering function for PB at 353 K compared between experiment and simulation. Simulation times are scaled by a factor of 0.8 to account for a difference in center of mass diffusion coefficient.

On larger length and time scales it is a challenge to reproduce in the simulations the dynamics of a single chain in the melt as observed in neutron spin echo (NSE) experiments³⁶. These experiments measure the configurational relaxation of a polymer chain on varying length scales in form of the intermediate coherent scattering function of the chain

$$S'(q, t) = S(q, t)/S(q, 0) = \frac{\langle \sum_{n,m} e^{i\mathbf{q} \cdot (\mathbf{r}_n(t) - \mathbf{r}_m(0))} \rangle}{\langle \sum_{n,m} e^{i\mathbf{q} \cdot (\mathbf{r}_n(0) - \mathbf{r}_m(0))} \rangle}. \quad (18)$$

Here the sums run over all atoms of the same chain, so one imagines an experiment where one has a few deuterated chains in a protonated matrix, giving rise to strong coherent scattering between the atoms of one single chain.

This quantity is easily calculated from the computer simulation and Fig. 17 shows a comparison¹³ of the results from the simulation (lines) to the experimental data (symbols) for a momentum transfer range of $q = 0.05 \text{ \AA}^{-1}$ to $q = 0.3 \text{ \AA}^{-1}$. It turned out that there is an overall difference in center of mass diffusivity of about 20 % between simulation and experiment similar to earlier experience³⁷, so for the figure the experimental time points are rescaled by a factor 0.8. This makes the whole set of scattering curves, measuring the configurational relaxation of the polymer chain on different length scales superimpose.

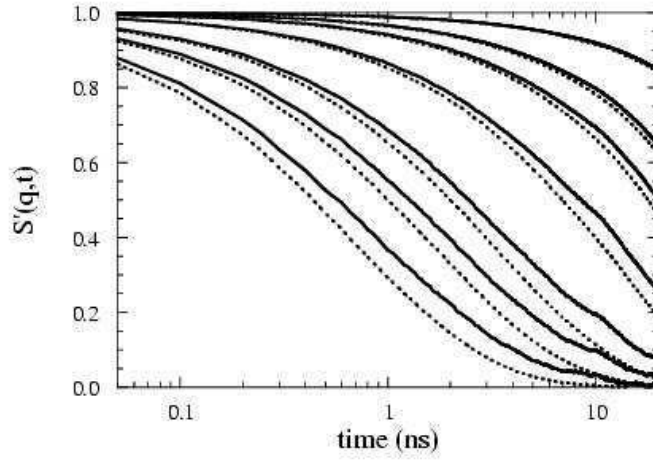


Figure 18. Single chain coherent intermediate scattering function for PB at 353 K (full lines) compared to the Rouse model prediction (dashed lines).

Polymer melt relaxation on these length and time scales and for chains below the entanglement molecular weight, as in our case, is typically analyzed within the Rouse model. The equation of motion for this chain of phantom beads

$$\begin{aligned} \zeta d\mathbf{r}_n &= \frac{3k_B T^2}{\sigma} (\mathbf{r}_{n+1} - 2\mathbf{r}_n + \mathbf{r}_{n-1}) dt + d\mathbf{W}_n(t) \\ \langle dW_{n\alpha}(t) \rangle &= 0 \\ \langle dW_{n\alpha}(t) dW_{m\beta}(t') \rangle &= \delta_{nm} \delta_{\alpha\beta} \delta(t - t') 2\zeta k_B T dt \end{aligned} \quad (19)$$

can be solved analytically and there is a closed expression for the single chain coherent intermediate scattering function

$$\begin{aligned} S(q, t) &= \frac{1}{N} \exp\{-q^2 D t\} \sum_{n,m=1}^N \exp\left\{-\frac{q^2 \sigma^2}{6} |n - m| - \frac{2Nq^2 \sigma^2}{3\pi^2}\right. \\ &\quad \left. \times \sum_{p=1}^N \cos\left(\frac{p\pi n}{N}\right) \cos\left(\frac{p\pi m}{N}\right) \left(1 - e^{-p^2 t/\tau_R}\right)\right\}. \end{aligned} \quad (20)$$

This function depends only on the length scale σ of the model and the segmental friction ζ entering the center of mass diffusion coefficient $D_N = k_B T / N \zeta$ and the longest relaxation time of the chain $\tau_R = \zeta N^2 \sigma^2 / 3 \pi^2 k_B T$. These quantities can be measured independently in the simulation so that we get a parameter free comparison to the prediction of the Rouse model (Fig. 18).

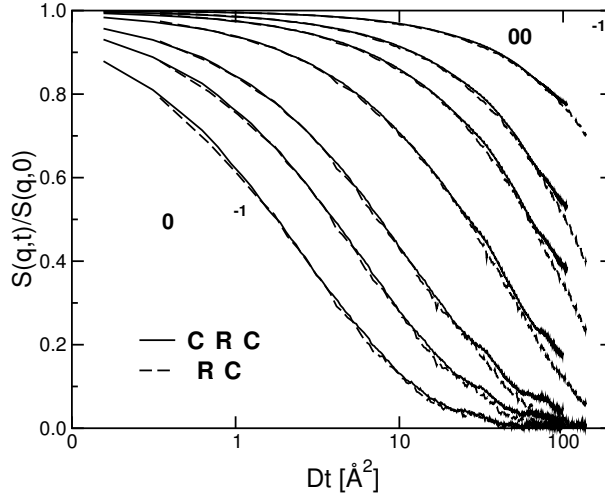


Figure 19. Comparison of the single chain coherent intermediate scattering function for the CRC (full lines) and FRC (dashed lines) models at 353 K. The difference in segmental friction is absorbed into a rescaling of the time axes by the chain center of mass diffusion coefficient.

The simulation data clearly show a stronger stretching compared to the Rouse model prediction, although there is an astonishingly good overall agreement considering the simple nature of the model. The differences are, however, significant and can be traced to the failure of the dynamic Gaussian assumption³⁸

$$\langle \exp\{i\mathbf{q} \cdot (\mathbf{r}_n(t) - \mathbf{r}_m(0))\} \rangle = \exp \left\{ -\frac{q^2}{6} \langle (\mathbf{r}_n(t) - \mathbf{r}_m(0))^2 \rangle \right\}$$

underlying the calculation of the scattering function. For the Rouse model this assumption is correct because it models the segmental motion as a Gaussian process, but in reality this assumption is not fulfilled due to interactions between a segment and its environment leading to non-Gaussian, heterogeneous mobility on the relevant time scales as we already discussed for the torsional dynamics.

As a final result on the large scale dynamics in polymer melts we can now come back to the question of universality. When we compare the results for the single chain intermediate scattering function of the CRC and FRC model³⁹ we expect them to be qualitatively the same. On the relevant time scales the difference in local dynamics observed in Fig. 16 gets absorbed into just one effective rate constant, the segmental friction ζ . One would therefore conclude that on these time and length scales the dynamics is the same after rescaling time scales for the difference in segmental friction and this is exactly what we find in Fig. 19.

5 Summary

In this contribution we have discussed on the one hand advanced MC methods for the equilibration of dense polymer systems, addressing slowing down due to entropic reasons, and chemically realistic MD simulations of 1,4-polybutadiene melts to exemplify the importance of local energetics for the understanding of relaxation processes in polymer melts. We have seen that modern connectivity altering MC algorithms are able to equilibrate the large-scale structure of polymer melts very efficiently, that they even increase in efficiency with increasing chain length. They can therefore provide starting configurations for MD simulations of chemically realistic polymer melts. These are targeted at understanding molecular structure and relaxations on small to intermediate length and time scales. They rely on carefully validated force fields which are able to quantitatively reproduce experimental data on well defined model systems. These force fields become available today thanks to a combination of experiments, high-level quantum chemistry calculations and simulations of simple model systems. Having established the ability of the simulation to reproduce available experimental data without any adjustable parameters one can then proceed to exploit the strong points of the simulation approach.

Where the range of thermodynamic parameters coverable in the simulation is often strongly limited, for those thermodynamic state points where a simulation in full equilibrium is possible one gets the complete information on the system under study down to every coordinate and momentum of every particle. This allows for measuring properties and correlations not available to experimental techniques, like for instance the distribution of waiting times between torsional transition, which are then instrumental in understanding and explaining effects of dynamic heterogeneity in polymer melts.

Another strong point of the simulation approach is its ability to selectively change parts of the model Hamiltonian. In this way we compared a chemically realistic model of PB with a freely rotating chain version of the same polymer and did not have to switch to a completely different polymer with some of the same properties like is unavoidable in experiments⁴⁰. With this approach we could establish that identical structure on the two-body correlation function level (single chain and liquid structure factors) does not imply identical dynamics. Understanding the energetics of the dihedral potential is instrumental to understanding relaxation processes in polymer melts. These results also raise questions on the applicability of mode-coupling theory^{41,42} to the glass transition in real polymer melts, because it relies on entropic packing effects as the cause for the glass transition.

Acknowledgments

It has been a pleasure to collaborate with G. D. Smith, S. Krushev, D. Bedrov and O. Borodin on the Molecular Dynamics simulations presented here. Obviously much of this work would not have been possible without the close collaboration with the following colleagues from the experiment side: M. D. Ediger, M. Monkenbusch, X. H. Qiu, D. Richter, L. Willner. The author acknowledges funding from the German Science Foundation under grant PA473/3-1,2 and BMBF under grant 03N6015.

References

1. K. Binder editor. *Monte Carlo and Molecular Dynamics Simulations in Polymer Science*. Oxford University Press, 1995.
2. M Doi and S. F. Edwards. *The Theory of Polymer Dynamics*. Oxford University Press, 1988.
3. S. C. Glotzer and W. Paul. *Annu. Rev. Mater. Res.* 32: 401, 2002.
4. R. Martoňák, W. Paul, and K. Binder. *Phys. Rev. E*, 57:2425, 1998.
5. G. D. Smith, Do Y. Yoon, W. Zhu, M. D. Ediger. *Macromolecules*, 27:5563, 1994.
6. W. Paul, Do Y. Yoon, and G. D. Smith, *J. Chem. Phys.*, 103:1702, 1995; W. Paul, G. D. Smith, and Do Y. Yoon, *Macromolecules*, 30:7772, 1997.
7. S. K. Nath, F. A. Escobedo, and J. J. de Pablo, *J. Chem. Phys.*, 108:9905, 1998; S. K. Nath, B. J Banaszak, and J. J. de Pablo, *J. Chem. Phys.*, 114:3612, 2001; J. Budzien, C. Raphael, M. D. Ediger, J. J. de Pablo, *J. Chem. Phys.*, 116:8209, 2002.
8. M. Mondello, G. S. Grest, A. R. Garcia, and B. G Silbernagel, *J. Chem. Phys.*, 105:5208, 1996; M. Mondello and G. S. Grest *J. Chem. Phys.*, 106:9327, 1997; M. Mondello, G. S. Grest, E. B. Webb, and P. Peczak, *J. Chem. Phys.*, 109:798, 1998.
9. V. G. Mavrantzas and D. N. Theodorou, *Macromolecules*, 31:6310, 1998; V. A. Harmandaris, V. G. Mavrantzas, and D. N. Theodorou, *Macromolecules*, 31:7934, 1998; *ibid.*, 33:8062, 2000; V. G. Mavrantzas and H. C. Öttinger, *Macromolecules*, 35:960, 2002.
10. M. Doxastakis, V. G. Mavrantzas, and D. N. Theodorou, *J. Chem. Phys.*, 115:11339, 2001; *ibid.*, 115:11352, 2001.
11. J. Colmenero, F. Alvarez, and A. Arbe, *Phys. Rev. E*, 65:041804, 2002.
12. K. Karatasos, F. Saija, and J.-P. Ryckaert, *Physica B*, 301:119, 2001; K. Karatasos and J.-P. Ryckaert, *Macromolecules*, 34:7232, 2001; K. Karatasos, J.-P. Ryckaert, R. Ricciardi, and F. Laupretre *Macromolecules*, 35:1451, 2002.
13. G. D. Smith and W. Paul. *J. Phys. Chem. A*, 102:1200, 1998; G. D. Smith, W. Paul, M. Monkenbusch, L. Willner, D. Richter, X. H. Qiu, and M. D. Ediger. *Macromolecules*, 32:8857, 1999.
14. R. Gee and R. H. Boyd, *J. Chem. Phys.*, 101:8028, 1994.
15. E. Kim, S. Misra, and W. L. Mattice, *Macromolecules*, 26:3424, 1993.
16. M. E. Tuckerman, B. J. Berne, and G. J. Martyna, *J. Chem. Phys.*, 97:1990, 1992; G. J. Martyna, M. E. Tuckerman, D. J. Tobias, and M. Klein, *Mol. Phys.*, 87:1117, 1996; A. Kopf, B. Dünweg, and W. Paul, *Comp. Phys. Commun.*, 101:1, 1997.
17. S. Geyler, T. Pakula, and J. Reiter, *J. Chem. Phys.*, 92:2676, 1990.
18. P. V. K. Pant and D. N. Theodorou, *Macromolecules*, 28:7224, 1995.
19. V. G. Mavrantzas, T. D. Boone, E. Zervopoulou, and D. N. Theodorou. *Macromolecules*, 32:5072, 1999; V. G. Mavrantzas and D. N. Theodorou, *Comput. Theor. Polym. Sci.*, 10:1 2000.
20. N. C. Karayiannis, V. G. Mavrantzas, and D. N. Theodorou, *Phys. Rev. Lett.*, 88:105503, 2002; N. C. Karayiannis, A. E. Giannousaki, V. G. Mavrantzas, and D. N. Theodorou, *J. Chem. Phys.*, 117:5465, 2002.
21. A. Uhlherr, V. G. Mavrantzas, M. Doxastakis, and D. N. Theodorou, *Macromolecules*, 34:8554, 2001; A. Uhlherr, M. Doxastakis, V. G. Mavrantzas, D. N. Theodorou, S. J. Leak, N. E. Adam, and P. E. Nyberg, *Europhys. Lett.*, 57:605, 2002.

22. M. G. Wu and M. W. Deem, *J. Chem. Phys.*, 111:6625, 1999.
23. D. N. Theodorou in *Bridging Time Scales: Molecular Simulations for the Next Decade*, P. Nielaba, M. Mareshal, and G. Ciccotti (Eds.), (Springer, Berlin, 2002), pp. 67-127.
24. W. Paul and M. Müller, *J. Chem. Phys.*, 115:630, 2001.
25. R. H. Boyd, R. H. Gee, J. Han, and Y. Jin. *J. Chem. Phys.*, 101:788, 1994.
26. G. D. Smith, O. Borodin, D. Bedrov, W. Paul, X. Qiu, and M. D. Ediger. *Macromolecules*, 34: 5192, 2001.
27. G. D. Smith, O. Borodin, and W. Paul. *J. Chem. Phys.*, 117:10350, 2002.
28. A. Arbe, D. Richter, J. Colmenero, and B. Farago. *Phys. Rev. E*, 54:3853, 1996.
29. A. Aouadi, M. J. Lebon, C. Dreyfus, B. Strube, W. Steffen, A. Patkowski, and M. R. Pick. *J. Phys. Condens. Matter*, 9:3803, 1997.
30. S. Krushev and W. Paul. *Phys. Rev. E*, 67:021806, 2003.
31. D. Richter, B. Frick, and B. Farago, *Phys. Rev. Lett.* 61:2465, 1988.
32. D. Bedrov, G. D. Smith, W. Paul, in preparation.
33. G. D. Smith, W. Paul, D. Y. Yoon, A. Zirkel, J. Hendricks, D. Richter, and H. Schober. *J. Chem. Phys.*, 107:4751, 1997.
34. K. Karatasos, F. Saija, J.-P. Ryckaert. *Physica B*, 301:119, 2001.
35. O. Ahumada, D. N. Theodorou, A. Triolo, V. Arrighi, C. Karatasos, and J.-P. Ryckaert. *Macromolecules*, 35:7110, 2002.
36. G. D. Smith, W. Paul, and D. Richter, *Chem. Phys.*, 261:61, 2000.
37. W. Paul, G. D. Smith, D. Y. Yoon, B. Farago, S. Rathgeber, A. Zirkel, L. Willner, and D. Richter. *Phys. Rev. Lett.*, 80:2346, 1998.
38. G. D. Smith, W. Paul, M. Monkenbusch, and D. Richer. *J. Chem. Phys.*, 114:4285, 2001.
39. S. Krushev, W. Paul, and G. D. Smith. *Macromolecules*, 35:4198, 2002.
40. A. Arbe, M. Monkenbusch, J. Stellbrink, D. Richter, B. Farago, K. Almdal, R. Faust. *Macromolecules*, 34:1281, 2001.
41. W. Götze and L. Sjögren. In *Transport Theory and Statistical Physics*. S. Yip and P. Nelson, eds., Marcel Dekker, 1995, pp. 801.
42. S. H. Chong SH and M. Fuchs, *Phys. Rev. Lett.* 88:185702, 2002.

Efficient Methods for Long Range Interactions in Periodic Geometries Plus One Application

Christian Holm

Max-Planck-Institut für Polymerforschung,
55021 Mainz, Germany

E-mail: holm@mpip-mainz.mpg.de

<http://www.espresso.mpg.de/>

We give an extensive introduction into the topic of how to compute efficiently long range interactions. We start out by reviewing the traditional Ewald sum for 3D Coulomb systems, discuss then in some detail the P³M method of Hockney and Eastwood, and briefly mention alternative ways of dealing with the Coulomb sum. We then discuss the best strategies to perform the sum under partially periodic boundary conditions. We also review the dipolar Ewald sum. For all our methods we provide error formulas which enable us to tune the algorithm at predefined accuracy. We then present the structure and design of our newly developed program package ESPResSo which contains already many of the presented methods. At the end we provide results of an extensive Molecular Dynamics simulations where we study the behavior of polyelectrolytes in poor solvents, taking explicitly care of the counterions. The conformations show pearl-necklace structures that are subject to strong conformational fluctuations, leading only to small signatures in the form factor. In addition we study how the necklace collapses as a function of Bjerrum length. At last we demonstrate that the position of the first peak in the inter-chain structure factor varies with the monomer density close to $\rho_m^{1/3}$ for all densities which is in strong contrast to polyelectrolyte solutions in good solvent.

1 Introduction

Soft condensed matter (or soft matter, as it is often called) is a term for materials in states of matter that are neither simple liquids nor hard solids of the type studied, for example, in solid state physics. Many such materials are familiar from everyday life - glues, paints, soaps, baby diapers - while others are important in industrial processes, such as polymer melts that are molded and extruded to form plastics¹. Biological materials are mainly made out of soft matter as well - membranes, actin filaments, DNA, RNA, and proteins belong to this class. Furthermore, most of the food we digest is soft matter. All these materials share the importance of length scales intermediate between atomic and macroscopic scales: The relevant range for soft matter lies between nanometers and micrometers. Examples are polymers, colloids, liquid crystals, glasses, and dipolar fluids. Typical energies between different structures are similar to thermal energies. Hence, Brownian motion or thermal fluctuations play a prominent role. Another key feature of soft matter systems is their propensity to self-assemble. Again the energy differences during this process are small such that many neighboring states are normally accessible through fluctuations. This often results in complex phase behaviors yielding a rich variety of accessible structures. Order does not necessarily arise on the single molecule level, but quite commonly exhibits a multitude of hierarchically ordered structures of sometimes tremendous intricacy and complexity. Most of the biological systems are usually not even in equilibrium but evolve among switchable steady states.

Given this wide field, research on soft material substances often acquires knowledge from different areas of research, such as physics, chemistry, and biology, such that a high level of interdisciplinary may be required for certain scientific questions.

In the past, our research has mainly focused on the study of charged polymers (polyelectrolytes) and charged colloids which serve as important substances for many technical applications. Charged systems also occur in biological environments (since most biological matter is charged)², and modeling explicit water molecules requires partial charges as well.

Computational simulations provide some unique ways to elucidate the properties of charged systems. However, they are time consuming, since the interaction between charged species is long ranged. Moreover, to compute the interactions of N particles fast and accurately, one needs smart algorithms to beat the unfavorable $\mathcal{O}(N^2)$ complexity which one obtains from simply counting all particle interactions.

In this contribution we give a short overview over recently developed methods to treat Coulomb interactions in fully or partially periodic geometries, and also show how to treat dipolar interactions. Next we present some features of our newly designed program package ESPResSo which implements most of these algorithms. We conclude with a short overview of an application to polyelectrolytes in poor solvent.

2 Methods for the Coulomb Sum in 3D

One of the biggest problems for the simulations of charged systems is the long range nature of the Coulomb interaction. In principle, each charge interacts with all others, leading to a computational effort of $\mathcal{O}(N^2)$ already within the central simulation box. For many physical investigations one wants to simulate bulk properties and therefore introduces periodic boundary conditions in all spatial directions to avoid boundary effects. The famous Ewald sum³⁻⁶ does a remarkable job in splitting the very slowly converging sum over the Coulomb potential into two exponentially converging sums. Still, this method suffers from two deficiencies. First, it is computationally demanding. This is due to the fact that one part of the problem is solved in reciprocal space implying the need for several Fourier transformations. Second, the algorithm scales like N^2 with N being the number of charged particles in the primary box, or at best like $N^{3/2}$, if one uses cutoffs which are optimized with respect to the splitting parameter⁷. The situation becomes worse if only partially periodic boundary conditions are applied. In this section we give an introduction to the Ewald summation, collecting the important equations for energy and forces. We discuss briefly some alternative methods. Some recent results of our research on how to deal with partially periodic boundary conditions will be deferred to the next two sections. The material has been mainly collected from the sources⁸⁻¹³. As an additional good textbook we can recommend the second edition of Frenkel and Smit¹⁴.

2.1 The Standard 3D Ewald Method

There are many examples of long-range interactions which can be treated by Ewald techniques¹⁵. Here, however, only the case of Coulomb point charges will be discussed, i.e., an interaction potential $1/r$. Consider therefore a system of N particles with charges q_i at positions $\mathbf{r}_i$ in an overall neutral and, for simplicity, cubic simulation box of length L . If

periodic boundary conditions are applied, the total electrostatic energy of the box is given by

$$E = \frac{1}{2} \sum_{\mathbf{n} \in \mathbb{Z}^3} \sum_{i,j=1}^N \frac{q_i q_j}{|\mathbf{r}_{ij} + \mathbf{n}L|}, \quad (1)$$

where $\mathbf{r}_{ij} = \mathbf{r}_i - \mathbf{r}_j$. Since this sum is only conditionally convergent, its value is not well-defined unless one specifies the precise way in which the cluster of simulation boxes should fill the $\mathbb{R}^3$, i.e., its *shape*, e.g. approximately spherical. A more detailed discussion of these points can be found in Refs.^{5,6,16,17}.

It is often stated that the long-range nature of the Coulomb potential complicates the treatment of electrostatic interactions. This is, however, only one part of the problem. In fact, the Coulomb potential bears *two* intrinsic difficulties. It is slowly decaying at large distances, and strongly varying at small distances. It is the combination of these two properties which leads to severe problems. If only one of them was present, everything would be comparatively easy, since a short-range potential could be treated by a simple cutoff, as it is done, e.g., for interactions of the Lennard-Jones type, and a long-range potential, which is periodic and slowly varying *everywhere*, can accurately be represented by the first few terms of its Fourier series.

Obviously, each of the two complications forbids the simple solution of the other, and the slowly decaying long-range part of the Coulomb potential renders a straightforward summation of Eqn. (1) impracticable. The trick is thus to split the problem into two parts by the trivial identity

$$\frac{1}{r} = \frac{f(r)}{r} + \frac{1-f(r)}{r}. \quad (2)$$

The underlying idea is to distribute the two complications between the two terms in Eqn. (2) by a suitable choice of the splitting function f . In particular:

- The first part $\frac{f(r)}{r}$ should be negligible, or even zero, beyond some cutoff $r_{\max}$, so that the summation up to the cutoff is a good approximation to (or the exact result of) this contribution to the total electrostatic potential.
- The second part $\frac{1-f(r)}{r}$ should be a slowly varying function for *all* r , so that its Fourier transform can be represented by only a few $\mathbf{k}$ -vectors with $|\mathbf{k}| \leq k_{\max}$. This permits an efficient calculation of this contribution to the total electrostatic potential in reciprocal space.

Since the field equations are linear, the sum of these two contributions gives the solution for the potential of the original problem.

The two requirements on the splitting function f mentioned above leave a large freedom of choice^{4,18,19}. The traditional selection is the complementary error function $\text{erfc}(r) := \frac{2}{\sqrt{\pi}} \int_r^\infty dt e^{-t^2}$. This results in the well known Ewald formula for the electrostatic energy of the primary box:

$$E = E^{(r)} + E^{(k)} + E^{(s)} + E^{(d)}, \quad (3)$$

where $E^{(r)}$ is the contribution from real space, $E^{(k)}$ the contribution from reciprocal space, $E^{(s)}$ the self energy and $E^{(d)}$ the dipole correction. They can be written as^{14,15}

$$E^{(r)} = \frac{1}{2} \sum_{\mathbf{m} \in \mathbb{Z}^3} \sum_{i,j} q_i q_j \frac{\text{erfc}(\alpha |\mathbf{r}_{ij} + \mathbf{m}L|)}{|\mathbf{r}_{ij} + \mathbf{m}L|} \quad (4)$$

$$E^{(k)} = \frac{1}{2} \frac{1}{V} \sum_{\mathbf{k} \neq \mathbf{0}} \frac{4\pi}{k^2} e^{-k^2/4\alpha^2} |\tilde{\rho}(\mathbf{k})|^2 \quad (5)$$

$$E^{(s)} = -\frac{\alpha}{\sqrt{\pi}} \sum_i q_i^2 \quad (6)$$

$$E^{(d)} = \frac{2\pi}{(1 + 2\epsilon')V} \left(\sum_i q_i \mathbf{r}_i \right)^2, \quad (7)$$

where the Fourier transformed charge density $\tilde{\rho}(\mathbf{k})$ is defined as

$$\tilde{\rho}(\mathbf{k}) = \int_V d^3r \rho(\mathbf{r}) e^{-i\mathbf{k} \cdot \mathbf{r}} = \sum_{j=1}^N q_j e^{-i\mathbf{k} \cdot \mathbf{r}_j} \quad \text{where} \quad \mathbf{k} \in \frac{2\pi}{L} \mathbb{Z}^3. \quad (8)$$

The advantage of rewriting Eqn. (1) this way is that the exponentially converging sums over $\mathbf{m}$ and $\mathbf{k}$ in (4,5) allow the introduction of comparatively small cutoffs without much loss in accuracy. Typically one chooses α large enough as to employ the minimum image convention in Eqn. (4). The inverse length α , which is often referred to as the Ewald (or splitting) parameter, tunes the relative weights of the real space and the reciprocal space contributions. However, the final result of the exact equation (3), not terminating the sums at some finite cutoff value, is independent of α . The form (7) given for the dipole correction assumes that the set of periodic replications of the simulation box tends spherically towards an infinite cluster and that the medium outside this sphere is a homogeneous dielectric with dielectric constant ϵ'^5 , whereas inside we set for simplicity $\epsilon = 1$. The derivation of this term is not straightforward and requires an accurate mathematical treatment of the conditional convergence of Eqn. (1), and of the image charges generated in the corresponding dielectric medium. Note that the case of a surrounding vacuum corresponds to $\epsilon' = 1$ and that the dipole correction vanishes for metallic (or “tin foil”) boundary conditions, since then $\epsilon' = \infty$. A detailed discussion of this term can be found in Ref.^{5,6,17}. The dipole correction in Eqn. (7) is independent of the Ewald parameter α . This again shows that the correction is not specific to the Ewald sum but more generally reflects the problems inherent to the conditional convergence of the $\mathbf{n}$ -sum in Eqn. (1). For the computation of $E^{(d)}$ the particle coordinates must *not* be folded back into the primary unit cell, for otherwise each boundary crossing produces an unphysical jump in the electrostatic energy, see Ref.¹⁶. The cutoffs $r_{\max}$ and $k_{\max}$ can be optimized with respect to α such that the required computer time scales like $N^{3/2}$, see Perram et al.⁷. This, however, may require that $r_{\max} > L/2$ prohibiting the simple minimum image convention in real space and rendering this procedure less tempting. For given finite real- and reciprocal space cutoffs there exists an optimal α such that the accuracy of the approximated Ewald sum is the highest possible. This optimal value can be determined easily with the help of the excellent estimates for the cutoff errors derived by Kolafa and Perram²⁰ – essentially by demanding that the real- and reciprocal space contribution to the error should be equal. If the system

under investigation is not electrostatically neutral, the infinite sum in Eqn. (1) diverges. It can be made convergent by adding a homogeneously distributed background charge which restores neutrality – a typical situation for one-component plasma simulations. This results in an additional electroneutrality-term $E^{(n)}$ to be included in Eqn. (3), which reads (see, e.g.,²¹)

$$E^{(n)} = -\frac{\pi}{2\alpha^2 V} \left(\sum_i q_i \right)^2. \quad (9)$$

Since the neutralizing background is homogeneous, the correction term in Eqn. (9) is independent of the particle positions.

The force $\mathbf{F}_i$ on particle i is obtained by differentiating the electrostatic potential energy E with respect to $\mathbf{r}_i$, i.e.,

$$\mathbf{F}_i = -\frac{\partial}{\partial \mathbf{r}_i} E. \quad (10)$$

Using Eqns. (3–8) one obtains the following Ewald formula for the forces:

$$\mathbf{F}_i = \mathbf{F}_i^{(r)} + \mathbf{F}_i^{(k)} + \mathbf{F}_i^{(d)}, \quad (11)$$

with the real space, Fourier space and dipole contributions given by:

$$\begin{aligned} \mathbf{F}_i^{(r)} = q_i \sum_j q_j \sum_{\mathbf{m} \in \mathbb{Z}^3} \left(\frac{2\alpha}{\sqrt{\pi}} \exp(-\alpha^2 |\mathbf{r}_{ij} + \mathbf{m}L|^2) \right. \\ \left. + \frac{\text{erfc}(\alpha |\mathbf{r}_{ij} + \mathbf{m}L|)}{|\mathbf{r}_{ij} + \mathbf{m}L|} \right) \frac{\mathbf{r}_{ij} + \mathbf{m}L}{|\mathbf{r}_{ij} + \mathbf{m}L|^2} \end{aligned} \quad (12)$$

$$\mathbf{F}_i^{(k)} = q_i \sum_j q_j \frac{1}{V} \sum_{\mathbf{k} \neq 0} \frac{4\pi \mathbf{k}}{k^2} \exp\left(-\frac{k^2}{4\alpha^2}\right) \sin(\mathbf{k} \cdot \mathbf{r}_{ij}) \quad (13)$$

$$\mathbf{F}_i^{(d)} = -\frac{4\pi q_i}{(1 + 2\epsilon')V} \sum_j q_j \mathbf{r}_j. \quad (14)$$

Since the self energy in Eqn. (6) and the neutralizing contribution in Eqn. (9) are independent of particle positions, they do not contribute to the force.

2.2 The Particle Mesh Ewald Idea

The Fourier transformations involved in Eqns. (5,13) are the most time consuming part of the Ewald sum. Several methods have been proposed to address this problem, e.g., tabulation of the complete Ewald potential²² or the use of polynomial approximations. A particularly successful approach for the latter is the expansion of the non-spherical contributions to the Ewald potential in cubic harmonics^{22,23}. Apart from the difficulty of computational overhead which may strongly increase with the desired accuracy, these methods do not solve the additional problem of unfavorable scaling with particle number. At best they scale as $N^{3/2}$, which is more costly than a plain cutoff or reaction field approach.

The essential idea is *not* to avoid the Fourier transformations but to modify the problem in such a way that it permits application of the Fast Fourier Transformation (FFT, see e.g.²⁴ and references therein). This reduces the complexity of the reciprocal part of the

Ewald sum to $N \log N$. For increasing number of particles the real space cutoff can be varied in such a way, that this scaling applies to the complete Ewald sum. Since the FFT is a grid transformation, there are discretization problems to be solved and corresponding discretization errors to be minimized. Performing the Fourier transformations in the reciprocal space part of the Ewald sum by FFT routines is by no means straightforward:

1. The point charges with continuous coordinates have to be replaced by a grid based charge density, since the FFT is a discrete and finite transformation.
2. It is neither obvious nor true that the best grid approximation to the continuum solution of the Poisson equation is achieved by using the continuum Green function²⁵.
3. There are at least three possibilities for implementing the differentiation needed in Eqn. (10). They differ in accuracy and speed.
4. The procedure of assigning the forces calculated on the mesh back to the actual particles can – under certain circumstances – lead to unwanted violations of Newton’s third law. They can be anything between harmless and disastrous.

The four steps involved in a particle mesh calculation are sources of various kinds of errors, originating, e.g., from discretization, interpolation or aliasing^a problems. Since these contributions are not independent of each other (reducing one might enhance another), the only reasonable demand is the minimization of the *total* error at given computational effort.

The original literature on particle mesh routines is usually not easy to digest for the non-specialist. It is obscured by the fact that various authors approach the problem from different directions and use different notations. There exist three mesh implementations of the Ewald sum – similar in spirit but different in detail. The oldest is the original particle-particle-particle-mesh (P³M) method of Hockney and Eastwood²⁵, and then there are two variants, namely, the Particle Mesh Ewald (PME) method of Darden *et al.*²⁶ and an extension of the latter by Essmann *et al.*²⁷, which is usually referred to as Smooth Particle Mesh Ewald (SPME).

In two papers by Deserno *et al.*^{8,9} it was shown how the three methods differ in detail, and it was demonstrated that the oldest method, namely the original P³M algorithm is actually the most accurate one, and since error estimates exist⁹, this method should be the preferred mesh method. The two new ideas of the more recent methods, namely the $i\mathbf{k}$ differentiation of PME, and the analytic differentiation of the charge assignment function in real space of SPME, can both also be used in P³M, making it the most versatile mesh algorithm on the market. Our experience shows that the break even point between a standard Ewald sum and the P³M algorithm is around 600-800 charges, depending on implementation details and desired accuracy. The algorithm can also be efficiently parallelized^{28,29}.

2.3 P³M in a Nutshell

The P³M method maps the system onto a mesh, such that the necessary Fourier transformations can be accomplished by Fast Fourier routines. At the same time the simple Coulomb

^a A finite grid cannot represent arbitrarily large $\mathbf{k}$ -vectors. Instead, they are folded back into the first “Brillouin zone” and distort there the true spectrum. This effect is usually referred to as “aliasing”. See, e.g., Sec. 12.1 in Ref.²⁴.

Green function $4\pi/k^2$ is adjusted by a functional minimization of an error functional as to make the result of the mesh calculation most closely resemble the continuum solution.

The first step, i.e., generating the mesh based charge density ρ_M (defined at the mesh points $\mathbf{r}_p$), is carried out with the help of a charge assignment function W :

$$\rho_M(\mathbf{r}_p) = \frac{1}{h^3} \sum_{i=1}^N q_i W(\mathbf{r}_p - \mathbf{r}_i) \quad (15)$$

Here h is the mesh spacing, and the number of mesh points $N_M = L/h$ along each direction should preferably be a power of two, since in this case the FFT is most efficient. The charge assignment function is classified according to its order P , i.e. between how many grid points – per coordinate direction – each charge is distributed. For W a cardinal B-spline³⁰ is chosen, which is a piecewise polynomial function of weight one. The order P gives the number of sections in the function. The first 7 cardinal-B-splines are sketched in Fig. 1. Its Fourier transform is

$$\tilde{W}(\mathbf{k}) = h^3 \left(\frac{\sin(\frac{1}{2}k_x h)}{\frac{1}{2}k_x h} \frac{\sin(\frac{1}{2}k_y h)}{\frac{1}{2}k_y h} \frac{\sin(\frac{1}{2}k_z h)}{\frac{1}{2}k_z h} \right)^P \quad (16)$$

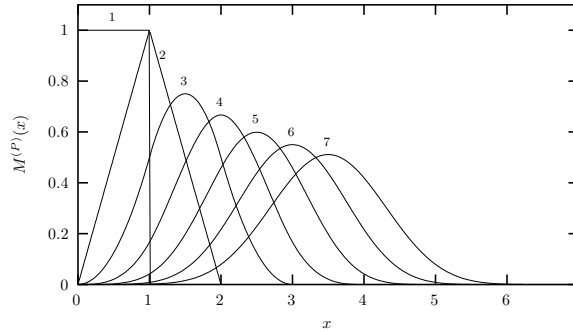


Figure 1. Sketch of the first 7 cardinal-B-splines $M^{(P)}(x)$, parameterized by P . Note that the charge assignment functions $W^{(P)}(x)$ for the P³M algorithm are just the “centered” B-splines.

In a second step the mesh based electric field $\mathbf{E}(\mathbf{r}_p)$ is calculated. Basically, the electric field is the derivative of the electrostatic potential, but there exist several alternatives for implementing the differentiation on a lattice⁸. Here we will restrict ourselves to the case of *i***k**-differentiation, which works by multiplying the Fourier transformed potential with *i***k**. In this case $\mathbf{E}(\mathbf{r}_p)$ can be written as

$$\mathbf{E}(\mathbf{r}_p) = \overleftarrow{\text{FFT}} \left[-i\mathbf{k} \times \overrightarrow{\text{FFT}} [\rho_M] \times \hat{G}_{\text{opt}} \right] (\mathbf{r}_p) \quad (17)$$

In words, $\mathbf{E}(\mathbf{r}_p)$ is the *backward* finite Fourier transform of the product of $-i\mathbf{k}$, the *forward* finite Fourier transform of the mesh based charge density ρ_M and the so-called optimal

influence function $\hat{G}_{\text{opt}}$, given by

$$\hat{G}_{\text{opt}}(\mathbf{k}) = \frac{\tilde{D}(\mathbf{k}) \cdot \sum_{\mathbf{m} \in \mathbb{Z}^3} \tilde{U}^2(\mathbf{k} + \frac{2\pi}{h}\mathbf{m}) \tilde{R}(\mathbf{k} + \frac{2\pi}{h}\mathbf{m})}{|\tilde{D}(\mathbf{k})|^2 \left[\sum_{\mathbf{m} \in \mathbb{Z}^3} \tilde{U}^2(\mathbf{k} + \frac{2\pi}{h}\mathbf{m}) \right]^2} \quad (18)$$

$$\text{with } \tilde{R}(\mathbf{k}) := -i\mathbf{k} \frac{4\pi}{k^2} e^{-k^2/4\alpha^2} \quad (19)$$

$$\text{and } \tilde{U}(\mathbf{k}) := \tilde{W}(\mathbf{k})/h^3 \quad (20)$$

Here $\tilde{D}(\mathbf{k})$ is the Fourier transform of the employed differentiation operator, which is simply $i\mathbf{k}$ in our case. Finally, one arrives at the force on particle i , i.e. the replacement of Eqn. (13):

$$\mathbf{F}_i = q_i \sum_{\mathbf{r}_p \in \mathbb{M}} \mathbf{E}(\mathbf{r}_p) W(\mathbf{r}_i - \mathbf{r}_p) \quad (21)$$

Hereby the sum extends over the complete mesh $\mathbb{M}$.

Although the presented formulas (15–21) look somewhat complicated, it is rather easy to implement them step by step. If the real space cutoff r_{max} is chosen small enough (so that the real space contribution (12) can be calculated in order N), the complete algorithm is essentially of order $N \log N$, see the article by Petersen³¹ for a nice exposition of this point.

2.4 How and Why to Control Errors

An investigation of the errors connected with particle mesh Ewald methods is important for three reasons:

1. The procedure of discretization introduces new sources of errors in addition to the ones originating from real and reciprocal space cutoffs.
2. Comparing the efficiency of different mesh methods is only fair if it is done at the same level of accuracy.
3. The tuning parameters should be chosen in such a way as to run the algorithm at its optimal operation point.

The two last points of course apply to any numerical method to compute energy or forces for the Coulomb (or dipolar) sum. Errors can tell us if we might see artifacts in simulations due to a poorly converged sum. They can tell us how the algorithm scales at its optimal point, and they can help us save a lot of expensive computer time. Unfortunately some of the free or commercially available computer programs choose parameter combinations automatically according to some more or less unknown rules, hidden in the program, which is a very dangerous route since after all the user needs to interpret the data. Therefore we stress here the point that for all our implemented routines we have error estimates, and we *use them*.

However, there is no unique or optimal measure of accuracy. If molecular dynamics (MD) simulations are performed, the main interest lies in errors connected with the force,

while in Monte-Carlo (MC) simulations one is concerned with accurate energies. In the simulation of ensemble averages it is the global accuracy, measured e.g. by root mean square quantities, which is important. In the simulation of rare events local accuracy and maximal errors are relevant as well. Errors in the force can be due to their magnitude or due to their direction. And finally, one might be interested in absolute or relative errors.

Whatever quantity one decides to look at, it can be investigated as a function of system parameters like particle separation or distribution, *tuning* parameters like α , mesh size or interpolation order and *components* of the algorithm, e.g., interpolation or differentiation scheme or splitting function $f(r)$. This obviously gives rise to a very large number of combinations. In other words: The corresponding parameter space is large and nontrivial, i.e., general statements concerning the performance of one method can usually not be extracted from low-dimensional cuts through this space, because different methods scale differently with respect to their parameters.

We are concerned with one measure of accuracy, namely the root mean square (rms) error in the force given by

$$\Delta F := \sqrt{\frac{1}{N} \sum_{i=1}^N (\mathbf{F}_i - \mathbf{F}_i^{\text{exa}})^2}, \quad (22)$$

where $\mathbf{F}_i$ is the force on particle number i calculated via some mesh method and $\mathbf{F}_i^{\text{exa}}$ is the *exact* physical force on that particle, calculable e.g. by a well converged standard Ewald sum. There exist error estimates for the real space and Fourier space contribution to this error for the standard Ewald sum²⁰, for the PME method³¹, and for the P³M method⁹.

Here we show the dependence of the root mean square error in the force in Eqn. (22) on the number of charged particles and their valence.^b Since the assumptions and arguments involved are of a rather general nature, the result is not specific to a certain type of Ewald method.

It is reasonable to assume that the error in the force on particle i can be written as

$$\Delta \mathbf{F}_i := \mathbf{F}_i - \mathbf{F}_i^{\text{exa}} = q_i \sum_{j \neq i} q_j \chi_{ij}. \quad (23)$$

The idea behind this Ansatz is that – just as it is true for $\mathbf{F}_i$ – the *error* in $\mathbf{F}_i$ originates from the $N - 1$ interactions of particle i with the other charged particles, and each contribution should be proportional to the product of the two charges involved. The vector χ_{ij} gives the direction and magnitude of this error for two unit charges and depends on their separation and orientation as well as on the specific algorithm used for calculating the electrostatic forces. For this term it is further assumed that

$$\langle \chi_{ij} \cdot \chi_{ik} \rangle = \delta_{jk} \langle \chi_{ij}^2 \rangle =: \delta_{jk} \chi^2, \quad (24)$$

where averaging over the particle configurations is denoted by the angular brackets $\langle \dots \rangle$. The underlying assumption that contributions from different particles are uncorrelated is certainly not always true, think, e.g., of highly ordered or strongly inhomogeneous particle distributions. However, it is a sensible one for *random* systems. Obviously, the term $\langle \chi_{ij}^2 \rangle$

^bNote that this is by no means the only interesting measure of accuracy

– the mean square force error for two unit charges – can no longer depend on i and j and is thus written as χ^2 . Using Eqns. (23,24), it follows

$$\langle (\Delta \mathbf{F}_i)^2 \rangle = q_i^2 \sum_{j \neq i} \sum_{k \neq i} q_j q_k \langle \chi_{ij} \cdot \chi_{ik} \rangle \approx q_i^2 \chi^2 \mathcal{Q}^2, \quad (25)$$

where the important quantity $\mathcal{Q}^2$ is defined as

$$\mathcal{Q}^2 := \sum_{j=1}^N q_j^2. \quad (26)$$

We further assume that

$$\left\langle \sqrt{\frac{1}{N} \sum_{i=1}^N (\Delta \mathbf{F}_i)^2} \right\rangle \approx \sqrt{\frac{1}{N} \sum_{i=1}^N \langle (\Delta \mathbf{F}_i)^2 \rangle}, \quad (27)$$

which can be shown true for reasonably large systems by the law of large numbers along the line of reasoning of Ref.⁹ to end up with the final relation

$$\Delta F \approx \chi \frac{\mathcal{Q}^2}{\sqrt{N}}. \quad (28)$$

Thus, the scaling of the rms error in the force with particle number and valence is given by the factor $\mathcal{Q}^2 N^{-1/2}$, whereas the prefactor χ – which cannot be obtained by such simple arguments – contains the details of the method and is independent of the simulated system. Note that any information on the valence distribution enters only through the value of $\mathcal{Q}^2$.

The most interesting ingredient of the P³M method is the optimal influence function from Eqn. (18). It is constructed such that the result of the mesh calculation is as close as possible to the solution of the original continuum problem. More precisely, the P³M method is derived from the requirement that the resulting Fourier space contribution to the force minimizes the the following error measure Q :

$$Q := \frac{1}{h^3} \int_{h^3} d^3 r_1 \int_{L^3} d^3 r [\mathbf{F}(\mathbf{r}; \mathbf{r}_1) - \mathbf{R}(\mathbf{r})]^2 \quad (29)$$

$\mathbf{F}(\mathbf{r}; \mathbf{r}_1)$ is the Fourier space contribution of the force between two unit charges at positions $\mathbf{r}_1$ and $\mathbf{r}_1 + \mathbf{r}$ as calculated by the P³M method (note that due to broken rotational and translational symmetry this does in fact depend on the coordinates of *both* particles), and $\mathbf{R}(\mathbf{r})$ is the corresponding exact reference force (whose Fourier transform is just Eqn. (19)). The inner integral over $\mathbf{r}$ scans all particle separations, whereas the outer integral over $\mathbf{r}_1$ averages over all possible locations of the first particle within a mesh cell. Obviously, up to a factor L^{-3} this expression is just the mean square error in the force for two unit charges, in other words, the quantity χ^2 from Eqn. (24). This provides a link between the rms error of an N particle system and the error Q from Hockney and Eastwood: Using Eqn. (28) one obtains

$$\Delta F \approx \mathcal{Q}^2 \sqrt{\frac{Q}{NL^3}}. \quad (30)$$

It is important to realize that Hockney and Eastwood not only provide a closed expression for the optimal influence function $\hat{G}_{\text{opt}}$, but also a closed expression for the corresponding “optimal error” $Q_{\text{opt}} = Q[\hat{G}_{\text{opt}}]$:

$$Q_{\text{opt}} = \frac{1}{L^3} \sum_{\mathbf{k} \in \hat{\mathbb{M}}} \left\{ \sum_{\mathbf{m} \in \mathbb{Z}^3} \left| \tilde{\mathbf{R}}(\mathbf{k} + \frac{2\pi}{h}\mathbf{m}) \right|^2 - \frac{\left| \tilde{\mathbf{D}}(\mathbf{k}) \cdot \sum_{\mathbf{m} \in \mathbb{Z}^3} \tilde{U}^2(\mathbf{k} + \frac{2\pi}{h}\mathbf{m}) \tilde{\mathbf{R}}^*(\mathbf{k} + \frac{2\pi}{h}\mathbf{m}) \right|^2}{|\tilde{\mathbf{D}}(\mathbf{k})|^2 \left[\sum_{\mathbf{m} \in \mathbb{Z}^3} \tilde{U}^2(\mathbf{k} + \frac{2\pi}{h}\mathbf{m}) \right]^2} \right\}. \quad (31)$$

The outer sum extends over all $\mathbf{k}$ -vectors of the Fourier transformed mesh $\hat{\mathbb{M}}$, and the asterisk denotes the complex conjugate. Once again, in the special case of ik-differentiation one has $\tilde{\mathbf{D}}(\mathbf{k}) = i\mathbf{k}$.

Admittedly, Eqn. (31) looks rather complicated. Still, in combination with Eqn. (30) it gives the rms force error of the P³M method or – more precisely – of its Fourier space contribution. After all, the computation of Q_{opt} and that of $\hat{G}_{\text{opt}}$ are quite similar. It should be emphasized that the formula (31) for the optimal Q -value, just like the one for the optimal influence function in Eqn. (18), is of a very general nature. It does also work for different charge assignment functions, reference forces or any differentiation scheme which can be expressed by an operator $\tilde{\mathbf{D}}(\mathbf{k})$.

The corresponding rms error in the force from the real space contribution in Eqn (12) has been derived by Kolafa and Perram²⁰ and is provided here for reference purpose:

$$\Delta F^{(r)} \approx \frac{2Q^2}{\sqrt{N}r_{\text{max}}L^3} \exp(-\alpha^2 r_{\text{max}}^2). \quad (32)$$

With these two estimates at hand it is easy to determine the optimal value of the splitting parameter α via a stand-alone program, which takes the relevant system parameters (N , Q^2 , L) and specifications of the algorithm (r_{max} , N_{M} , P) as its input, where N_{M} denotes the number of mesh points. If real and reciprocal space contribution to the error, $\Delta F^{(r)}$ and $\Delta F^{(k)}$ respectively, are assumed to be statistically independent, the total error is given by

$$\Delta F = \sqrt{(\Delta F^{(r)})^2 + (\Delta F^{(k)})^2}. \quad (33)$$

This quantity has to be minimized with respect to α . In most cases it is, however, accurate enough to use the following approximation: Determine the value of α at which the real and reciprocal space contribution to the rms force error are equal.

2.5 Alternatives to Ewald

Another way of computing Eq.(1) is via a convergence factor

$$E = \lim_{\beta \rightarrow 0} \frac{1}{2} \sum_{\mathbf{n}} \sum_{ij}^I \frac{q_i q_j \exp(-\beta |\mathbf{r}_{ij} + \mathbf{n}L|)}{|\mathbf{r}_{ij} + \mathbf{n}L|}. \quad (34)$$

This approach is used in the Lekner³² and Sperb^{33,34} methods to efficiently sum up the 3D Coulomb sum. Although the method in its original versions has $\mathcal{O}(N^2)$ complexity,

Strebel and Sperb have developed a factorization approach which yields an $\mathcal{O}(N \log N)$ algorithm³⁵.

Other advanced methods of $\mathcal{O}(N \log N)$ are tree algorithms³⁶, which are the first order approximation of even better, so-called fast multipole methods³⁷ (FMM). These can reach a linear complexity, but at the expense of a heavy computational overhead which makes these methods advantageous only for a very large number of charges ($N \approx 100\,000$)³⁸, see also the contribution of Grubmüller in this volume. Another recent approach is to use multigrid methods which solve the Poisson equation in real space^{39,40}. These have the advantage that they should be more useful on massively parallel architectures since an efficient implementation of 3D FFTs is difficult on those machines. Another advantage could be the fact that for MD simulations the previous Poisson solution can still be used as a starting point for the multigrid iteration thereby improving drastically the convergence speed.

All these alternative methods (except for the Strebel algorithm) share the property that in principle they can be applied also to geometrical situations where only one or two dimensions are periodically replicated. The Ewald like algorithms can be formulated for those geometries as well, however they loose tremendously in speed. However, several tricks can be applied to accelerate them.

3 Two Periodic Dimensions

For thin polyelectrolytes films, interactions of charged species with membranes, or in general geometries, where only two dimensions are periodically replicated, and the third one has a finite extend h ($2D + h$ geometry), the original Ewald sum has to be modified. For such systems Ewald based formulas are only slowly convergent, have mostly $\mathcal{O}(N^2)$ scalings and no “a priori” error estimates exist⁴¹. A fast Ewald based method has been put forward in⁴², where the Fourier sum is specially treated with an integral approximation. However, in our opinion, this introduces non-controllable errors, which makes the method useless⁴³. Also the Lekner method, which is based on a resummation of the force sum, can be used, however it has an $\mathcal{O}(N^2)$ scaling, and special care has to be taken for particles which are nearby in the non-periodic dimension, see also the discussion of Mazars⁴⁴ on accuracy problems. Recently we proposed a new method called **MMM2D**^{11,45} which has an $\mathcal{O}(N^{5/3})$ complexity and full error control. It is based on a convergence factor approach similar to **MMM**³³. In two dimensions the convergence factor based methods and the Ewald sum methods yield exactly the same results. However, this will still only allow simulations including up to a few thousand charges. There have been early attempts to use the full 3D Ewald sum also for these slab problems. The main idea is to fill only parts of the simulation box with charges and to leave some space empty, in an attempt to decouple the interactions in the third dimension^{46–48}. Since each image layer is globally neutral, one hopes that their interactions decay as they become more and more distant, i.e. as the size of the gap is increased. In this way one could make use of any advanced 3D Ewald implementation, see also Ref.⁴⁹ for a variant of this idea.

3.1 The ELC Method

Up to recently the 3D methods using a gap had to be checked on a trial and error basis. We recently greatly improved this situation by deriving a term, called electrostatic layer correction (**ELC**), which subtracts the interactions due to the unwanted layers^{12,13}. In this way the necessary gap length can be largely reduced. The combination of that term with any three dimensional summation method with slab-wise summation order will yield the exact electrostatic energy. Since the change in the summation order is done by adding a very simple term, any three dimensional summation method with the standard spherical summation order can be used. The **ELC**-term can be evaluated easily in a time linear in the number of charges, hence the whole method scales like the underlying standard summation method. We also developed an error formula for the maximal pairwise error in the energy and forces of the layer correction term, hence the precision of this method can be tuned to any desired value, when used in conjunction with other error estimates for the standard summation method^{20,9}. Using, for example, the P³M method one obtains an $N \log N$ scaling with well controlled errors also for the $2D + h$ geometry, which up to now seems to be the optimal choice.

We just try to give here some insight into the basic idea of changing the summation order. Consider this time a system of N particles with charges q_i and positions $\mathbf{r}_i = (x_i, y_i, z_i)$ that reside in a box of edges $L \times L \times h$, where $h = \max_{i,j} |z_i - z_j|$ is the maximal z -distance of two particles. The basic idea is to expand this slab system in the non-periodic z -coordinate to a system with periodicity in all three dimensions. More precisely, the original box of size $L \times L \times h$ is placed inside a box of size $L \times L \times L_z$ where $L_z \gg h$ sufficiently large. Then this box is replicated periodically in all three dimensions. The result is a three-dimensional periodic system with empty space regions (“gaps”) of height $\delta := L_z - h$ (see Fig. 2). δ will be called gap size in the following.

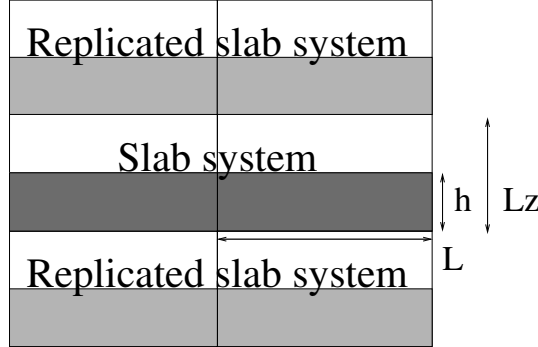


Figure 2. Schematic representation of a fully periodically replicated slab system

Since the electrostatic potential is only finite if the total system is charge neutral, the additional image layers (those layers above or below the original slab system) are charge neutral, too. Now let us consider the n^{th} image layer which has an offset of nL_z to the original layer. If nL_z is large enough, each particle of charge q_j at position $(x_j, y_j, z_j + nL_z)$ and its replicas in the x, y -plane can be viewed as constituting a homogeneous charged sheet

of charge density $\sigma_j = \frac{q_j}{L_z}$. The potential of such a charged sheet at distance z is $2\pi\sigma_j|z|$. Now we consider the contribution from a pair of image layers located at $\pm nL_z$, $n > 0$ to the energy of a charge q_i at position (x_i, y_i, z_i) in the central layer. Since $|z_j - z_i| < nL_z$, we have $|z_j - z_i + nL_z| = nL_z + z_j - z_i$ and $|z_j - z_i - nL_z| = nL_z - z_j + z_i$, and hence the interaction energy from those two image layers with the charge q_i vanishes by charge neutrality:

$$2\pi q_i \sum_{j=1}^N \sigma_j (|z_j - z_i + nL_z| + |z_j - z_i - nL_z|) = 4\pi q_i nL_z \sum_{j=1}^N \sigma_j = 0. \quad (35)$$

The only errors occurring are those coming from the approximation of assuming homogeneously charged, infinite sheets instead of discrete charges. This assumption should become better when increasing the distance nL_z from the central layer.

However, in a naive implementation, even large gap sizes will result in large errors⁴⁸. This is due to the order of summation for the three dimensional Coulomb sum, which is spherical by convention. This order implies that with increasing shell cutoff S the number of image shells grows faster than the number of shells of the primary layer, namely $\mathcal{O}(S^3)$ versus $\mathcal{O}(S^2)$ (see Fig. 3(a)). In other words, we include the unwanted terms faster than the actually wanted terms. Also the image layers are not really infinite charged sheets but are truncated due to the cutoff. Yeh and Berkowitz⁴⁸ already suggested that this problem can be solved by changing the order of summation. Smith has shown that by adding to the Coulomb energy the term

$$E_c = 2\pi M_z^2 - \frac{2\pi \mathbf{M}^2}{3}, \quad (36)$$

where $\mathbf{M} = \sum q_i \mathbf{r}_i$ is the total dipole moment, one obtains the result of a slab-wise summation instead of the spherical limit⁵⁰. Slab-wise summation refers to the sum $\sum_{|n| \geq 0} E_l(n)$, where $E_l(n)$ denotes the energy, calculated in circular summation order, resulting from the image layer with shift nL_z in the z -coordinate. Technically this is the order where we first treat the original layer and then add the image layers grouped in symmetrical pairs (see Fig. 3(b)). Obviously this summation order fits much better to the charged sheet argumentation given above. Although this is a major change in the summation order, the difference given by Eq. (36) is a very simple term. In fact, Smith shows that changes of the summation order always result in a difference that depends only on the total dipole moment.

Applying this slab-wise summation order, Yeh and Berkowitz showed that a gap size of at least h is normally sufficient to obtain an moderately accurate result. Therefore the result of a standard three dimensional summation method plus the shape-dependent term given by Eq. (36), which we refer to as a *slab-wise method*, can be used to obtain a good approximation to the result for the slab geometry with the same computational effort as for the underlying three dimensional summation method (no matter if a simple or sophisticated method is used). One drawback is that no theoretical estimates exist for the error introduced by the image layers. Therefore one might be forced to use even larger gaps to assure that no artifacts are produced by the image layers. One deducible artifact is that the pairwise error will be position dependant. Particles in the middle of the slab will see no effect of the image layers due to symmetry, and particles near the surface will encounter for the same reason the largest errors, which is definitely an unwanted feature for studying surface

	$S^2=3$	$S=2$	$S^2=3$	
$S^2=3$	$S^2=2$	$S=1$	$S^2=2$	$S^2=3$
$S=2$	$S=1$	$S=0$	$S=1$	$S=2$
$S^2=3$	$S^2=2$	$S=1$	$S^2=2$	$S^2=3$
	$S^2=3$	$S=2$	$S^2=3$	

(a) Schematic view of the spherical summation order. S is the length of the box offset.

$n=2$	$n=2$	$n=2$
$n=1$	$n=1$	$n=1$
$n=0$	$n=0$	$n=0$
$n=1$	$n=1$	$n=1$
$n=2$	$n=2$	$n=2$

(b) Schematic view of the slab-wise summation order. n is the z offset of the box, the spherical summation order in the x, y -plane is not shown.

Figure 3. Schematic views of different summation orders.

effects. Therefore averaging error measures like the commonly used RMS force error should not be applied without additional checks for the particles near the surfaces.

The other drawback is that normally the box now will have a significantly larger L_z/L . But at least for Ewald type methods the computation time is proportional to this fraction. This is easy to see as the number of k -space vectors in the z direction must be proportional to L_z to maintain a fixed resolution and therefore error. It is verified experimentally that a gap of at least h is needed. For a cubic system $h = L$ therefore the computation time at least doubles.

Nevertheless because of the bad scaling of the known methods for slab geometries like the one by Parry^{51,52} ($\mathcal{O}(N^2)$) or **MMM2D**^{11,45} ($\mathcal{O}(N^{5/3})$), for particle numbers above $N \approx 1000$ using slab-wise methods is a great improvement. Fortunately one can do even better!

We briefly show how the **ELC** term can be computed; consult the original article for more details^{12,13}. The method relies on the far formula of **MMM2D**^{11,45}. For the following analysis there is no special restriction on h except for $h < L_z$, which is true even if the

$L \times L \times L_z$ -box is completely filled.

We start with a formal definition of the Coulomb energy of the slab system

$$E = \frac{1}{2} \sum_{S=0}^{\infty} \sum_{\substack{\mathbf{n} \in \mathbb{Z}^2 \times \{0\} \\ n_x^2 + n_y^2 = S}} \sum'_{i,j=1}^N \frac{q_i q_j}{|\mathbf{r}_i - \mathbf{r}_j + \Lambda \mathbf{n}|}. \quad (37)$$

$\Lambda = \text{diag}(L, L, L_z)$ is a diagonal matrix describing the shape of the box. The image boxes are denoted with the vector $\mathbf{n} = (n_x, n_y, n_z)$, where $n_z = 0$ for now. The prime on the inner summation indicates the omission of the self-interaction $i = j$ in the primary box $\mathbf{n} = (0, 0, 0)$ (i. e. the singular case). For the surrounding dielectric medium we assume vacuum boundary conditions.

We now expand the system to a fully three-dimensional periodic system, where L_z determines the period in the z -coordinate as in the previous section. We can rewrite the energy as

$$E = E_s + E_c + E_{lc}, \quad (38)$$

where

$$E_s = \frac{1}{2} \sum_{S=0}^{\infty} \sum_{\substack{\mathbf{n} \in \mathbb{Z}^3 \\ \mathbf{n}^2 = S}} \sum'_{i,j=1}^N \frac{q_i q_j}{|\mathbf{r}_i - \mathbf{r}_j + \Lambda \mathbf{n}|}. \quad (39)$$

denotes the standard three-dimensional Coulomb-sum with spherical limit. To evaluate this expression one can use any of the efficient algorithms, starting with the classical Ewald summation up to modern methods like fast multipole methods³⁷ or mesh based algorithms⁸. E_c again denotes the shape-dependent term given by Eq. (36) and finally

$$E_{lc} = -\frac{1}{2} \sum_{\substack{T \in \mathbb{Z} \\ T > 0}} \sum_{n_z = \pm T} \sum_{S=0}^{\infty} \sum_{\substack{\mathbf{n} \in \mathbb{Z}^2 \times \{n_z\} \\ \mathbf{n}^2 = S}} \sum'_{i,j=1}^N \frac{q_i q_j}{|\mathbf{r}_i - \mathbf{r}_j + \Lambda \mathbf{n}|}. \quad (40)$$

denotes the contribution of the image layers, for which we are going to derive a new expression in the following.

We start with the expression for the energy induced by an image layer at z -offset $n_z \neq 0$:

$$E_l(n_z) = -\frac{1}{2} \sum_{S=0}^{\infty} \sum_{\substack{\mathbf{n} \in \mathbb{Z}^2 \times \{n_z\} \\ \mathbf{n}^2 = S}} \sum'_{i,j=1}^N \frac{q_i q_j}{|\mathbf{r}_i - \mathbf{r}_j + \Lambda \mathbf{n}|}. \quad (41)$$

It can be shown rigorously, although this is non-trivial, that

$$E_l(n_z) = -\frac{1}{2} \lim_{\beta \rightarrow 0} \sum_{\mathbf{n} \in \mathbb{Z}^2 \times \{n_z\}} \sum'_{i,j=1}^N \frac{q_i q_j e^{-\beta |\mathbf{r}_i - \mathbf{r}_j + \Lambda \mathbf{n}|}}{|\mathbf{r}_i - \mathbf{r}_j + \Lambda \mathbf{n}|}. \quad (42)$$

This is a convergence factor approach with a convergence factor of $e^{-\beta |\mathbf{r}_i - \mathbf{r}_j + \Lambda \mathbf{n}|}$. Note that this approach is exact only for two-dimensional systems, for three-dimensional systems Eqs. (41) and (42) differ by a multiple of the dipole moment^{5,6}.

In^{53, 11} one can find a proof for this equation and an efficient way of calculating E_l for charge neutral systems. We do not want to go through the full derivation again; it consists of the application of Poisson's summation formula along both periodic coordinates and performing the limit $\beta \rightarrow 0$ analytically. One obtains

$$E_{lc}(n_z) = -\frac{1}{2} \sum_{i,j=1}^N q_i q_j \phi(\mathbf{r}_i - \mathbf{r}_j + \Lambda \mathbf{n}), \quad (43)$$

where ϕ is an artificial pairwise potential that yields the total Coulomb energy and its derivative produces the pairwise forces for the periodic system, and $k_{\parallel} = (k_x, k_y)$ is a Fourier variable with integer values.

For now we only have a formula for the contribution of one image layer, so we still have to sum over all n_z . This task can be performed analytically. The terms $2\pi|z|/L^2$ can be omitted since they are exactly the homogeneous sheet potential and we have seen before that this cancels out for charge neutral systems (see Eq. (35)).

The summation over n_z of the remaining sums over p and q is fairly easy to perform using the geometric series (as these sums are absolutely convergent, exchanging the summation over n_z and the summations over (k_x, k_y) is possible). Combining the terms for $\pm n_z$ again we obtain

$$E_{lc} = \sum_{i,j=1}^N q_i q_j \psi(\mathbf{r}_i - \mathbf{r}_j + \Lambda \mathbf{n}), \quad (44)$$

where

$$\begin{aligned} \psi(x, y, z) = & \frac{4}{L} \sum_{k_x, k_y > 0} \frac{\cosh(2\pi k_{\parallel} z_{ij}/L)}{k_{\parallel} (e^{2\pi k_{\parallel} L_z/L} - 1)} \cos(2\pi k_x x_{ij}/L) \cos(2\pi k_y y_{ij}/L) + \\ & \frac{2}{L} \sum_{k_x > 0} \frac{\cosh(2\pi k_x z_{ij}/L) \cos(2\pi k_x x_{ij}/L)}{k_x (e^{2\pi k_x L_z/L} - 1)} + \\ & \frac{2}{L} \sum_{k_y > 0} \frac{\cosh(2\pi k_y z_{ij}/L) \cos(2\pi k_y y_{ij}/L)}{k_y (e^{2\pi k_y L_z/L} - 1)}. \end{aligned} \quad (45)$$

The forces can be obtained from that by simple differentiation since the sums are absolutely convergent. Although the form in Eq.(45) has a much better convergence than the original form in Eq.(40), its main advantage is a linear computation time with respect to the number of particles N . To see this, the equation has to be rewritten using the addition theorems for the cosine and the hyperbolic cosine. For each $k_{\parallel}$ one first calculates the

sixteen terms

$$\begin{aligned}
\chi(c/s, c/s, c/s) &= \sum_{i=1}^N q_i \cosh / \sinh(2\pi k_{\parallel} z_i / L) \cos / \sin(2\pi k_x x_i / L) \cos / \sin(2\pi k_y y_i / L), \\
\chi(x, c/s, c/s) &= \sum_{i=1}^N q_i \cosh / \sinh(2\pi k_x z_i / L) \cos / \sin(2\pi k_x x_i / L), \\
\chi(y, c/s, c/s) &= \sum_{i=1}^N q_i \cosh / \sinh(2\pi k_y z_i / L) \cos / \sin(2\pi k_y y_i / L),
\end{aligned} \tag{46}$$

where the indices in the obvious way determine which of the functions cosine (hyperbolicus) or sinus (hyperbolicus) are used. Then we evaluate

$$\begin{aligned}
E_{lc} &= \frac{4}{L} \sum_{k_x, k_y > 0} \frac{1}{(e^{2\pi k_{\parallel} L_z / L} - 1) k_{\parallel}} \left(\chi_{(ccc)}^2 + \chi_{(csc)}^2 + \chi_{(ccs)}^2 + \chi_{(css)}^2 - \right. \\
&\quad \left. \chi_{(scc)}^2 - \chi_{(ssc)}^2 - \chi_{(scs)}^2 - \chi_{(sss)}^2 \right) + \\
&\quad \frac{2}{L} \sum_{k_x > 0} \frac{1}{(e^{2\pi k_x L_z / L} - 1) k_x} \left(\chi_{(xcc)}^2 + \chi_{(xcs)}^2 - \chi_{(xsc)}^2 - \chi_{(xss)}^2 \right) + \\
&\quad \frac{2}{L} \sum_{k_y > 0} \frac{1}{(e^{2\pi k_y L_z / L} - 1) k_y} \left(\chi_{(ycc)}^2 + \chi_{(ycs)}^2 - \chi_{(ysc)}^2 - \chi_{(yss)}^2 \right)
\end{aligned} \tag{47}$$

Similar expansions using the same sixteen terms can also be found for the forces. Obviously this has linear computation time with respect to the number of particles, as the only summations over all the particles occur in the χ_* . The infinite summation over $k_{\parallel}$ can be cutoff at some value to achieve the desired accuracy, using the error estimates presented by Arnold et al¹².

To summarize, the main profits of the **ELC** term are, that it scales as the number N of particles and that it has a rigorous error bound. Moreover this error bound can be used to estimate the size of the image layer contribution and therefore gives a bound on the error introduced by slab-wise methods as proposed by Yeh and Berkowitz⁴⁸. We found that the error for these methods decays exponentially in L_z/L . However, the errors are not uniformly distributed over the slab, namely they are worst at the surfaces of the slabs, hence the maximal pairwise error should be used instead of the usually RMS-errors of (22). In the second paper on **ELC**¹³ we considered in detail the implementation of the layer correction for the standard Ewald and the P³M method. There we also derived anisotropic Ewald error formulas, gave some fundamental guidelines for optimization, demonstrated the accuracy, and gave error formulas and computation times for typical systems.

4 The Coulomb Sum in One Periodic Dimension

Boundary conditions where only one dimension is periodic, and the other two are open or finite appear, for example, in physical situations such as electronic structures of supported

structures on metal surfaces, such as steps and atomic chains in one dimensions, or in systems which can model a charged stiff polymer or DNA piece, where for avoiding end effects the rod is made infinitely long^{54–56}. It was only very recently that a 1D Ewald method (EW1D) for these systems has been developed^{57,58}, although also the Lekner method³² developed earlier can be used, see however⁴⁴. Again, both are $\mathcal{O}(N^2)$ methods which converge rather slow, and no error estimates have been provided.

We recently developed a method which we called **MMM1D**⁵⁹, which is basically the **MMM2D** method adapted to 1D. Again, rigorous error bounds can be derived, and although the method is $\mathcal{O}(N^2)$, it is much faster than EW1D, and easier to use. We applied this method already to charged rod systems with good results^{55,56}.

As mentioned in (2.5) in principle the FMM, multigrid methods, or tree codes can handle this situation, but they are quite involved for the normally only small number of charges involved, and error estimates are not easy to obtain.

Also a modified Ewald method in which the summation of the reciprocal-scape vectors was modified⁶⁰, similar to the one used by Kawata and Mikami⁶¹ exists, but again the approximations made seem hard to control which render the method rather useless.

And finally, one can in principle use any 3D Ewald method and separate the rod images sufficiently to ensure that the interactions are negligible, which can be ensured for systems where mobile counterions are present which can screen the interactions. This has been applied often by Deserno et al.^{10,62,63}. As a side remark, similar to the **ELC** method, by changing the summation order and using the prescription of Smith⁵⁰, a corrected 3DEW summation should also be possible, but it might again not be easy to solve the problem of how to control the error.

5 The Dipolar Ewald Summation in 3D

In the previous section we treated methods to deal with the Coulomb, or more generally, interactions which vary with $1/r$. However, there are also systems of interest, which can be modeled by interacting point dipoles. Substances of that kind are ferrofluids, which are basically dispersed magnetic particles⁶⁴, or magneto-rheological (MR) or electro-rheological (ER) fluids, or solvents which can be modeled approximately by dipolar interactions like water. For periodic geometries the Ewald method is again applicable. In the simplest implementations the involved computation time grows like N^2 , or at best like $N^{3/2}$, if the cutoff is optimally varied with the splitting parameter⁷. For MC simulations, knowing the energy formulas is sufficient, whereas for MD simulations, we need to know forces and torques. In this part we give a reliable error estimate for the energy, for the forces, and for the torques, when computed via the standard Ewald sum. Moreover we will give a detailed discussion on the optimization of the parameters, which will lead to the most efficient parameters for a predefined error in each observable quantity. This can all be done prior to the actual simulation, ensuring thus optimal performance at optimally controlled errors. The exposition is heavily based on the paper by Wang and Holm⁶⁵.

Consider a system of N particles with a point-dipole μ_i at their center position $\mathbf{r}_i$ in a cubic simulation box of length L . If periodic boundary conditions are applied, the total

electrostatic energy of the box is given by

$$U = \frac{1}{2} \sum_{\mathbf{n} \in \mathbb{Z}^3} \sum_{i,j=1}^N \left\{ \frac{\boldsymbol{\mu}_i \cdot \boldsymbol{\mu}_j}{|\mathbf{r}_{ij} + \mathbf{n}|^3} - \frac{3[\boldsymbol{\mu}_i \cdot (\mathbf{r}_{ij} + \mathbf{n})][\boldsymbol{\mu}_j \cdot (\mathbf{r}_{ij} + \mathbf{n})]}{|\mathbf{r}_{ij} + \mathbf{n}|^5} \right\}. \quad (48)$$

where $\mathbf{r}_{ij} = \mathbf{r}_i - \mathbf{r}_j$. The sum over $\mathbf{n}$ is over all simple cubic lattice points, $\mathbf{n} = (n_x L, n_y L, n_z L)$ with n_x, n_y, n_z integers. The slowly decaying long range part of the dipolar potential makes the straightforward summation of Eqn.(48) too time consuming. The Ewald trick splits the problem again into two rapidly convergent parts, one in real space and one in reciprocal space. The details of the method are discussed in Refs.^{3,5,7,15}, here we only give the final expressions

$$U = U^{(r)} + U^{(k)} + U^{(s)} + U^{(d)}, \quad (49)$$

where the real-space $U^{(r)}$, the k-space (reciprocal-space) $U^{(k)}$, the self $U^{(s)}$ and the dipole (surface) $U^{(d)}$ contributions are respectively given by:

$$U^{(r)} = \frac{1}{2} \sum_{\mathbf{n} \in \mathbb{Z}^3} \sum_{i,j=1}^N \left\{ (\boldsymbol{\mu}_i \cdot \boldsymbol{\mu}_j) B(|\mathbf{r}_{ij} + \mathbf{n}|) - [\boldsymbol{\mu}_i \cdot (\mathbf{r}_{ij} + \mathbf{n})][\boldsymbol{\mu}_j \cdot (\mathbf{r}_{ij} + \mathbf{n})] C(|\mathbf{r}_{ij} + \mathbf{n}|) \right\}, \quad (50)$$

$$U^{(k)} = \frac{1}{2L^3} \sum_{\mathbf{k} \in \mathbb{Z}^3, \mathbf{k} \neq 0} \frac{4\pi}{k^2} e^{-(\pi k/\alpha L)^2} \sum_{i,j=1}^N (\boldsymbol{\mu}_i \cdot \mathbf{k})(\boldsymbol{\mu}_j \cdot \mathbf{k}) e^{2\pi i \mathbf{k} \cdot \mathbf{r}_{ij}/L}, \quad (51)$$

$$U^{(s)} = -\frac{2\alpha^3}{3\sqrt{\pi}} \sum_{i=1}^N \mu_i^2, \quad (52)$$

$$U^{(d)} = \frac{2\pi}{(2\epsilon' + 1)L^3} \sum_{i,j=1}^N \boldsymbol{\mu}_i \cdot \boldsymbol{\mu}_j. \quad (53)$$

The sums over i and j are for the particles in the central box and

$$B(r) = [\text{erfc}(\alpha r) + (2\alpha r/\sqrt{\pi}) \exp(-\alpha^2 r^2)]/r^3, \quad (54)$$

$$C(r) = [3\text{erfc}(\alpha r) + (2\alpha r/\sqrt{\pi})(3 + 2\alpha^2 r^2) \exp(-\alpha^2 r^2)]/r^5. \quad (55)$$

The inverse length α is the splitting parameter of the Ewald summation which should be chosen so as to optimize the performance. The form Eqn.(53) given for the surface correction assumes that the set of the periodic replications of the simulation box tends in a spherical way towards an infinite cluster and that the medium outside this sphere is an uniform dielectric with dielectric constant $\epsilon'^{15,5}$. The case of a surrounding vacuum corresponds to $\epsilon' = 1$ and the surface term vanishes for the metallic boundary conditions ($\epsilon' = \infty$).

In practical calculations, the infinite sums in Eqns.(50, 51) are truncated by only taking into account distances which are smaller than some real space cutoff r_c and wave vectors with a modulus smaller than some reciprocal space cutoff k_c . If $r_c \leq L/2$, the sum in real space (Eqn.(50)) reduces to the normal minimum image convention. The double sum over particles in $U^{(k)}$ can be replaced by a product of two single sums which is more suitable for numerical calculations.

The force $\mathbf{F}_i$ acting on particle i is obtained by differentiating the potential energy U with respect to $\mathbf{r}_i$, i.e.,

$$\mathbf{F}_i = -\frac{\partial U}{\partial \mathbf{r}_i} = \mathbf{F}_i^{(r)} + \mathbf{F}_i^{(k)}, \quad (56)$$

with the real-space and k-space contributions given by:

$$\begin{aligned} \mathbf{F}_i^{(r)} = \sum_{\mathbf{n} \in \mathbb{Z}^3} \sum_{j=1}^N \left\{ \left\{ (\boldsymbol{\mu}_i \cdot \boldsymbol{\mu}_j)(\mathbf{r}_{ij} + \mathbf{n}) + \boldsymbol{\mu}_i[\boldsymbol{\mu}_j \cdot (\mathbf{r}_{ij} + \mathbf{n})] + [\boldsymbol{\mu}_i \cdot (\mathbf{r}_{ij} + \mathbf{n})]\boldsymbol{\mu}_j \right\} C(|\mathbf{r}_{ij} + \mathbf{n}|) \right. \\ \left. - [\boldsymbol{\mu}_i \cdot (\mathbf{r}_{ij} + \mathbf{n})][\boldsymbol{\mu}_j \cdot (\mathbf{r}_{ij} + \mathbf{n})]D(|\mathbf{r}_{ij} + \mathbf{n}|)(\mathbf{r}_{ij} + \mathbf{n}) \right\}, \end{aligned} \quad (57)$$

$$\mathbf{F}_i^{(k)} = \frac{2\pi}{L^4} \sum_{j=1}^N \sum_{\mathbf{k} \in \mathbb{Z}^3, \mathbf{k} \neq 0} \frac{4\pi\mathbf{k}}{k^2} (\boldsymbol{\mu}_i \cdot \mathbf{k})(\boldsymbol{\mu}_j \cdot \mathbf{k}) \exp[-(\pi k/\alpha L)^2] \sin(2\pi\mathbf{k} \cdot \mathbf{r}_{ij}/L), \quad (58)$$

where

$$D(r) = [15\text{erfc}(\alpha r) + (2\alpha r/\sqrt{\pi})(15 + 10\alpha^2 r^2 + 4\alpha^4 r^4) \exp(-\alpha^2 r^2)]/r^7. \quad (59)$$

Since the self and surface energy terms [Eqns.(52, 53)] are independent of the particle positions, they have no contributions to the force, unlike the Ewald summation for the Coulomb systems where the surface term contributes.

The torque $\boldsymbol{\tau}_i$ acting on particle i is related to the electrostatic field $\mathbf{E}_i$ at the location of this particle,

$$\boldsymbol{\tau}_i = \boldsymbol{\mu}_i \times \mathbf{E}_i = \boldsymbol{\tau}_i^{(r)} + \boldsymbol{\tau}_i^{(k)} + \boldsymbol{\tau}_i^{(d)}, \quad (60)$$

with

$$\mathbf{E}_i = -\frac{\partial U}{\partial \boldsymbol{\mu}_i}, \quad (61)$$

and thus

$$\begin{aligned} \boldsymbol{\tau}_i^{(r)} = - \sum_{\mathbf{n} \in \mathbb{Z}^3} \sum_{j=1}^N \left\{ (\boldsymbol{\mu}_i \times \boldsymbol{\mu}_j) B(|\mathbf{r}_{ij} + \mathbf{n}|) - \right. \\ \left. [\boldsymbol{\mu}_i \times (\mathbf{r}_{ij} + \mathbf{n})][\boldsymbol{\mu}_j \cdot (\mathbf{r}_{ij} + \mathbf{n})] C(|\mathbf{r}_{ij} + \mathbf{n}|) \right\}, \end{aligned} \quad (62)$$

$$\boldsymbol{\tau}_i^{(k)} = -\frac{1}{L^3} \sum_{j=1}^N \sum_{\mathbf{k} \in \mathbb{Z}^3, \mathbf{k} \neq 0} \frac{4\pi}{k^2} (\boldsymbol{\mu}_i \times \mathbf{k})(\boldsymbol{\mu}_j \cdot \mathbf{k}) e^{-(\pi k/\alpha L)^2} e^{2\pi i \mathbf{k} \cdot \mathbf{r}_{ij}/L}, \quad (63)$$

$$\boldsymbol{\tau}_i^{(d)} = -\frac{4\pi}{(2\epsilon' + 1)L^3} \sum_{j=1}^N \boldsymbol{\mu}_i \times \boldsymbol{\mu}_j. \quad (64)$$

5.1 Error Formulas

We now give estimates for the rms error caused by cutting off the Ewald summation in real-space and k-space for the forces, the total energy and the torques. There are no errors involved in the self and surface contributions (Eqns(52, 53, 64)), because no cutoff operations are applied to them.

As can be shown similarly to the way as it was presented in Sec.2.4 the rms error for the force can be cast in the following form

$$\Delta F^{(r)} \approx \chi^{(r)} \frac{\mathcal{M}^2}{\sqrt{N}}, \quad (65)$$

where the quantity $\mathcal{M}$ is defined as

$$\mathcal{M}^2 := \sum_{j=1}^N \mu_j^2. \quad (66)$$

After some lengthy calculation one obtains

$$\chi^{(r)2} \approx L^{-3} r_c^{-9} \alpha^{-4} \left(\frac{13}{6} C_c^2 + \frac{2}{15} D_c^2 - \frac{13}{15} C_c D_c \right) \exp(-2\alpha^2 r_c^2), \quad (67)$$

where the terms C_c and D_c are given by

$$C_c = 4\alpha^4 r_c^4 + 6\alpha^2 r_c^2 + 3, \quad (68)$$

$$D_c = 8\alpha^6 r_c^6 + 20\alpha^4 r_c^4 + 30\alpha^2 r_c^2 + 15. \quad (69)$$

The resulting rms expectation of the real-space cutoff error in the forces is thus

$$\Delta F^{(r)} \approx \mathcal{M}^2 (L^3 \alpha^4 r_c^9 N)^{-1/2} \left(\frac{13}{6} C_c^2 + \frac{2}{15} D_c^2 - \frac{13}{15} C_c D_c \right)^{1/2} \exp(-\alpha^2 r_c^2). \quad (70)$$

The derivation of the expected real-space cutoff errors in the total potential energy and torques follows the same way. For calculating the fluctuation of the error in total energy, it is noted that the interaction energy between two dipoles is evenly shared between them. That means the sum of $\langle (\Delta U^{(r)})^2 \rangle$ over all particles contains each pair contribution twice and thus the fluctuation of the real-space cutoff error is one half of the sum²⁰. Then the rms value of the real-space cutoff error of the total potential energy is estimated as

$$\Delta U^{(r)} \approx \mathcal{M}^2 (L^3 \alpha^4 r_c^7 N)^{-1/2} \left[\frac{1}{4} B_c^2 + \frac{1}{15} C_c^2 - \frac{1}{6} B_c C_c \right]^{1/2} \exp(-\alpha^2 r_c^2) \quad (71)$$

with

$$B_c = 2\alpha^2 r_c^2 + 1. \quad (72)$$

The rms error on the torques is estimated similarly to the force as

$$\Delta \tau^{(r)} \approx \mathcal{M}^2 (L^3 \alpha^4 r_c^7 N)^{-1/2} \left[\frac{1}{2} B_c^2 + \frac{1}{5} C_c^2 \right]^{1/2} \exp(-\alpha^2 r_c^2). \quad (73)$$

Eqn.(70, 71, 73) all contain the exponential $\exp(-\alpha^2 r_c^2)$. For sufficiently low errors, αr_c has to be larger than one, for example $\alpha r_c \approx \pi$ for an error of $\exp(-\pi^2) \approx 5 \times 10^{-5}$. If

only the highest powers of αr_c are retained, the estimates of the real-space cutoff errors in the total energy, forces and torques can be reduced to

$$\Delta U^{(r)} \approx 4\mathcal{M}^2\alpha^2(r_c/15NL^3)^{1/2} \exp(-\alpha^2 r_c^2), \quad (74)$$

$$\Delta F^{(r)} \approx 8\mathcal{M}^2\alpha^4(2r_c^3/15NL^3)^{1/2} \exp(-\alpha^2 r_c^2), \quad (75)$$

$$\Delta \tau^{(r)} \approx 4\mathcal{M}^2\alpha^2(r_c/5NL^3)^{1/2} \exp(-\alpha^2 r_c^2), \quad (76)$$

where Eqn.(74) is a factor of $\sqrt{6/5}$ slightly larger than that given in Eqn.(35) of Ref.²⁰. The advantage of these simplified formulas is that they reflect the dependence of the rms errors on α and r_c more directly and thus could be used more easily in determining the optimal values of these parameters.

In deriving the estimates of the reciprocal-space (k-space) cutoff errors, we further assume that the radial distribution function of the particles is approximately unity at all distances. Following Eqn.(58), the k-space cutoff error in the force acting on particle i is given by

$$\Delta \mathbf{F}_i^{(k)} = \sum_{j=1}^N \sum_{\mathbf{k}, k > k_c} \frac{8\pi^2 \mathbf{k}}{L^4 k^2} (\boldsymbol{\mu}_i \cdot \mathbf{k})(\boldsymbol{\mu}_j \cdot \mathbf{k}) \exp[-(\pi k/\alpha L)^2] \sin(2\pi \mathbf{k} \cdot \mathbf{r}_{ij}/L). \quad (77)$$

Note that the diagonal term ($j = i$) in the sum does not depend on the positions of the particles. It will provide a systematic contribution to the cutoff error in k-space²⁰. In Eqn.(77) this contribution equals to zero, thus there is no systematic part of the error in the forces. The same thing happens to the error in the torques. But for the total energy the diagonal terms are positive and the systematic contribution plays a dominant role in the cutoff error.

The off-diagonal terms in Eqn.(77) do depend on the positions of the particles and have alternating signs. Similarly as before, the off-diagonal contribution to the cutoff error in $\Delta \mathbf{F}_i^{(k)}$ is given by

$$\Delta \mathbf{F}_{i,off}^{(k)} = |\boldsymbol{\mu}_i| \sum_{j \neq i} |\boldsymbol{\mu}_j| \chi_{ij}^{(k)} \quad (78)$$

with

$$\chi_{ij}^{(k)} = \sum_{\mathbf{k}, k > k_c} \frac{8\pi^2 \mathbf{k}}{L^4} \cos \vartheta(\hat{\boldsymbol{\mu}}_i, \hat{\mathbf{k}}) \cos \vartheta(\hat{\boldsymbol{\mu}}_j, \hat{\mathbf{k}}) \exp[-(\pi k/\alpha L)^2] i \exp(2\pi i \mathbf{k} \cdot \mathbf{r}/L), \quad (79)$$

where $\mathbf{r}$ stands for $\mathbf{r}_{ij}$, $\sin(2\pi \mathbf{k} \cdot \mathbf{r}/L)$ is re-written as $i \exp(2\pi i \mathbf{k} \cdot \mathbf{r}/L)$ according to the symmetrical character of the summation over $\mathbf{k}$, and $\hat{\mathbf{k}}$ is the unit vector along $\mathbf{k}$. Since the particles are assumed to be randomly distributed over the simulation box, the fluctuation of $\Delta \mathbf{F}_{i,off}^{(k)}$ can also be written as

$$\langle (\Delta \mathbf{F}_{i,off}^{(k)})^2 \rangle = \mu_i^2 \sum_{j \neq i} \sum_{m \neq i} |\boldsymbol{\mu}_j| |\boldsymbol{\mu}_m| \langle \chi_{ij}^{(k)} \cdot \chi_{im}^{(k)*} \rangle \approx \mu_i^2 \mathcal{M}^2 \chi^{(k)2}, \quad (80)$$

where again $\chi^{(k)2}$ is independent of i and j .

Choosing the z-axis of the spherical coordinates (k, θ, ϕ) along the $\hat{\boldsymbol{\mu}}$ orientation, we find

$$\chi^{(k)2} \approx 128\pi^3 L^{-6} \alpha^2 k_c^3 \exp[-2(\pi k_c/\alpha L)^2]/15. \quad (81)$$

The rms expectation of the k-space cutoff error in the forces is thus

$$\Delta F^{(k)} \approx 8\pi\mathcal{M}^2L^{-3}\alpha(2\pi k_c^3/15N)^{1/2}\exp[-(\pi k_c/\alpha L)^2]. \quad (82)$$

Here the notation $\Delta F_{off}^{(k)}$ is replaced directly with $\Delta F^{(k)}$ due to the fact of no diagonal contribution.

The derivation of the off-diagonal parts of the cutoff errors in the total energy and torques proceeds in the same way. That the sum over $\langle(\Delta U_{i,off}^{(k)})^2\rangle$ contains each pair contribution twice has also been considered in the error estimate of the total energy. The results are given by

$$\Delta U_{off}^{(k)} \approx 4\mathcal{M}^2L^{-2}\alpha(\pi k_c/15N)^{1/2}\exp[-(\pi k_c/\alpha L)^2], \quad (83)$$

$$\Delta \tau^{(k)} \approx 4\mathcal{M}^2L^{-2}\alpha(\pi k_c/5N)^{1/2}\exp[-(\pi k_c/\alpha L)^2]. \quad (84)$$

$\Delta \tau^{(k)}$ is also used directly instead of $\Delta \tau_{off}^{(k)}$.

The diagonal (systematic) part of the cutoff error in the total energy can be written as

$$\begin{aligned} \Delta U_{diag}^{(k)} &= \frac{1}{2\sqrt{N}} \sum_{i=1}^N \sum_{\mathbf{k}, k > k_c} \frac{4\pi}{L^3} \mu_i^2 \cos^2 \vartheta(\hat{\boldsymbol{\mu}}_i, \hat{\mathbf{k}}) \exp[-(\pi k/\alpha L)^2] \\ &\approx \frac{2\pi}{L^3\sqrt{N}} \mathcal{M}^2 \int_{k_c}^{\infty} \exp[-(\pi k/\alpha L)^2] k^2 dk \int_0^{\pi} \sin \theta d\theta \int_0^{2\pi} \cos^2 \vartheta(\hat{\boldsymbol{\mu}}, \hat{\mathbf{k}}) d\phi \\ &\approx \frac{4}{3} \mathcal{M}^2 L^{-1} \alpha^2 k_c N^{-1/2} \exp[-(\pi k_c/\alpha L)^2], \end{aligned} \quad (85)$$

where the sum is again approximated by an integral and an integral expansion formula²⁰ is used to get the final estimate. The total k-space cutoff error in the total energy is thus

$$\Delta U^{(k)} = \Delta U_{diag}^{(k)} + \Delta U_{off}^{(k)}. \quad (86)$$

Comparing Eqn.(85) with (83), it can be seen that the systematic part of the error is a factor of $\sim L\alpha k_c^{1/2} (\gg 1)$ larger than the statistical part. Hence the systematic contribution is dominant in the k-space cutoff error of the total energy.

Assuming that the real-space and reciprocal-space contributions to the error are independent, the total cutoff error in Ewald summation can be written as

$$\Delta \Theta = \sqrt{\Delta \Theta^{(r)2} + \Delta \Theta^{(k)2}}, \quad (87)$$

where Θ stands for U, F and τ .

5.2 Optimization of Parameters

Now we show how to use the analytical formulas derived in Sec. 5.1 to determine the optimal values of α , r_c and k_c by which the required accuracy could be satisfied and the computation time is minimized. The detailed discussions on this subject can also be found in Refs.^{7,31}.

The overall computation time for computing the forces with the Ewald method is approximately given by³¹

$$\mathcal{T} = a_r N^2 (r_c/L)^3 + a_k N k_c^3, \quad (88)$$

where the primitive overheads a_r and a_k highly depend on the implementation of the code and need to be determined by numerical experiments. As an example, we have carried out timing experiments on a DEC personal workstation (CPU 433MHz) using a standard Fortran 77 compiler. In the implementation the complementary error function and its derivative are calculated with table lookup and the reciprocal-space summation is optimized as in Refs.^{7,15}. The linked-cell method is used to deal with the short-range forces (when doing simulations). The primitive overheads are then found roughly to be $a_r = 2.5\mu s$ and $a_k = 0.7\mu s$.

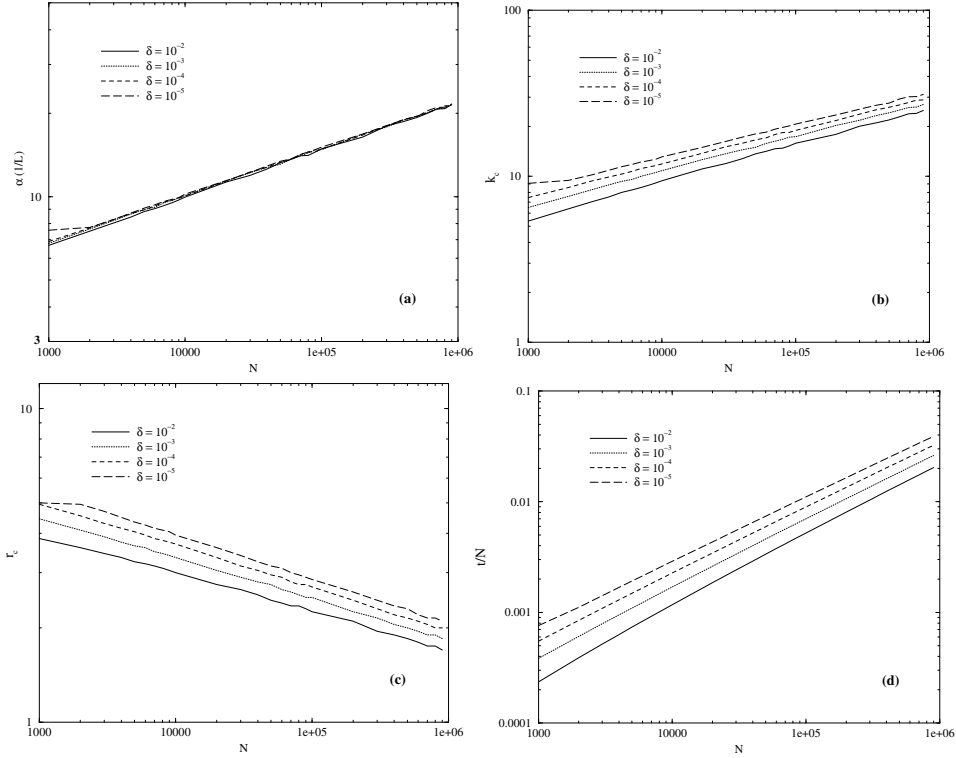


Figure 4. Optimal values of the parameters α (a), k_c (b) and r_c (c) as well as the corresponding minimized computation time T/N (d) as a function of the number of particles.

For a required accuracy δ , the parameters α , r_c and k_c should be chosen to minimize $\mathcal{T}$ with respect to the two constraints of the error bounds [Eqns.(70) and (82)], which are restated as

$$\frac{\delta}{\sqrt{2}} = \mathcal{M}^2 (L^3 \alpha^4 r_c^9 N)^{-1/2} \left(\frac{13}{6} C_c^2 + \frac{2}{15} D_c^2 - \frac{13}{15} C_c D_c \right)^{1/2} \exp(-\alpha^2 r_c^2), \quad (89)$$

$$\frac{\delta}{\sqrt{2}} = 8\pi \mathcal{M}^2 L^{-3} \alpha (2\pi k_c^3 / 15N)^{1/2} \exp[-(\pi k_c / \alpha L)^2]. \quad (90)$$

In case of $\delta \leq 5 \times 10^{-5}$, Eqn.(75) could be used instead of (70) so as to show the de-

pendence of the accuracy on the parameters more clearly. Eqn.(89) and (90) provide the qualitative function relations of r_c and k_c with α as: $r_c(\alpha) \approx -A\sqrt{\ln \delta}/\alpha$ and $k_c(\alpha) \approx -B\sqrt{\ln \delta}\alpha$. Inserting them into Eqn.(88) and differentiating it with respect to α yields $\alpha \propto N^{1/6}$ and thus $r_c \propto N^{-1/6}$ and $k_c \propto N^{1/6}$. The minimized computation time is then proportional to $N^{3/2}$ with the proportionality constant depending on the accuracy. The same results can be found for the Coulomb Ewald method in Refs.^{7,20,31}. This can be easily understood by comparing Eqns.(70) and (82) in Sec. 5.1 of this paper with Eqns.(18) and (32) in Ref.²⁰, where one finds the same exponential dependences of the cutoff errors on α , r_c and k_c for the dipolar and Coulomb Ewald summations.

The numerical investigation of the functional dependence of the optimal parameters on N and δ are performed by using the primitive overheads obtained above. For each given N and δ , we at first choose different values for r_c within the inequality $r_c \leq L/2$. For each r_c the parameters α and k_c are calculated by solving Eqns.(89) and (90). These values are then introduced into Eqn.(88) to figure out the optimal value of r_c which gives the minimum computation time. In calculations the size of the simulation cell is fixed to a dimensionless length of $L = 10$. The range of accuracy requirement and number of particles are chosen to be $\delta = 10^{-2}$ to 10^{-5} measuring in $\mathcal{P}^2/\mathcal{L}^4$ and $N = 10^3$ to $N = 10^6$ which should cover most of the applications. The particles are supposed to have an uniform dipole moment of $\mathcal{P}$. The results for the optimal values of the parameters and the corresponding computation time per particle are shown in Fig. 4 (a-d), respectively. It can be clearly seen that the functional dependence of the parameters and the overall computation time on N are just as discussed above. Fig.4(c) shows that when a high accuracy is required for a system with a small number of particles, the predicted real-space cutoff is larger than half of the box length and $r_c = L/2$ must be used. The optimal α values hardly depend on the accuracies. These results are very similar to that obtained for the Coulomb Ewald summation in Ref.³¹, except for $r_c \propto N^{-1/6}$ here and $r_c \propto N^{1/6}$ there. This is because they considered a system of constant density, while we choose the volume of the simulation cell to be constant.

6 The ESPResSo Project

As shown before, the algorithms for various fully or partially periodic geometries can be quite complex, and thus the production of single data points can take weeks or even more for complex biomolecular problems like protein folding, see for example other contributions in this volume. Not only data production takes time, also newcomers to the field experience often a painfully long learning period to get familiar with the intricacies of the methods and the code. This in turn implies, that modifications of the algorithms are difficult and time consuming. Considerable amounts of valuable research time is spent on coding issues with the final result of a highly specialized research tool. Often this results in human resources being kept away either from performing algorithmic improvements or scientific investigations. Since our group has been developing fast algorithms for quite some time now, mostly in specialized programs, we wanted to bundle them into one package. Moreover, other algorithms for soft matter research, coming from either colleagues at the theory group at the MPI-P, or interested research groups around the world, should possibly be included in the future. Our requirements therefore were, that the program should be easy to use, but scientifically sound; it should grant experts access to state-of-the-art techniques, but enable beginners to become experts as well, therefore it should be well-documented

providing exhaustive informations. And finally, new algorithmic additions should be fairly easy to be included, and certainly the program should not be a “black box”.

Looking at other available simulation packages, e.g. BALL⁶⁶, GISMOS⁶⁷, GRO-MOS⁶⁸, LAMMPS⁶⁹, NAMD⁷⁰, polyMD⁷¹, and OCTA⁷², we did not find a single package which met all our needs.

This led us to design a newly structured program for research on soft matter, which we called ESPResSo : an **E**xtensible **S**imulation **P**ackage for **R**esearch on **S**oft Matter Systems⁷³. The program is intended to enable us to study soft matter model systems via MD and MC algorithms, with particular emphasis on extensibility for new, highly complex force/energy algorithms. Since the problems under investigation are located along scientific frontiers, meaning they are complex and computationally time-consuming, the program is parallelizable, fast, accurate, and easily modifiable. Here, we present only a short overview of the ESPResSo-package. Updates and more documentation can be found on the web page <http://www.espresso.mpg.de/>. The distribution of the source code adheres to the open source standards and can be obtained upon request. We hope that other researchers can test, improve or even enlarge our code into a valuable research tool for the soft matter community.

6.1 Design

The ESPResSo design was developed to specifically serve the demands of our computational research group. While most simulation programs focus only on single aspects of such a project, ESPResSo is suited to help researchers in the *whole* process between the specification of a scientific problem and the interpretation of the results. Since ESPResSo offers a variety of methods and combines the knowledge of tens of man-years of research expertise on soft matter it helps newcomers to get into simulation techniques and to choose the right method for a certain problem. On the other hand experts can easily implement their own special routines into the framework of ESPResSo, enabling them to explore paths outside the scope of their own programs. Often new problems require new algorithmic solutions. ESPResSo helps the user implement new features due to its hierarchical structure, its modularity, its general data structures and its well-defined interfaces. A test-suite which is part of ESPResSo helps in checking if new features reproduce well-known physical properties of model systems. The hierarchical structure is well-suited for running simulations without required knowledge of the whole program package. The whole system setup is contained within a Tcl⁷⁴ script. A large number of sample scripts for various simulation problems help in developing new applications. Inside a simulation script, one can handle the entire simulation process from the specification of a system, the actual simulation, its analysis and the graphical output of the results. We want to emphasize our goal to design a practical research tool which is easy to learn, use and extend. This is especially supported by ESPResSo being a team project, since this ensures that every part of the code has to pass through a discussion process provoking a simple, effective and understandable implementation.

A difficult task in the design process arises from conflicts between different requirements regarding the simultaneous optimization of several aspects. In order to ensure that new researchers do not need too much time to learn how ESPResSo works, the code has to be kept simple, which is sometimes in contradiction to code optimization for computational speed. Aiming at being able to handle a wide variety of topics instead of solving only

specific problems leads to the same challenge of countering the code's tendency towards more complexity, hence less understandability.

Hierarchical program structure: ESPResSo is built up using three hierarchical program levels. The steering of the program is done on a script language level. All tasks are implemented as extensions to the script language dealing with input and output of data, setting of particle properties, interactions and parameters, and performing the integration and analysis of a given system. The basic simulation level is implemented in C. It contains the integrator as well as the calculation of fundamental observables like forces, torques, energies, pressure and temperature. These first two levels build up the part that is common to all investigated problems. Consequently, it is the part which should be known by a researcher using ESPResSo. Therefore, special emphasis was placed on simplicity and readability. The third level, also implemented in C, ensures the speed, efficiency and great generality of ESPResSo. This includes algorithms to accelerate the force and energy calculations used by the integrator as well as special algorithms to treat long-range interactions. Parallelization of all time-critical parts of the program enable efficient large scale simulations (see the benchmarks in Section 6.2). On this level, one also finds all implemented potentials for the particle interactions, and interfaces to other programs like VMD⁷⁵ for on-the-fly visualization of simulations.

Modularity: The hierarchical structure is accomplished by splitting the levels into different modules, which subdivides the otherwise large program package into manageable pieces. This is particularly important for the aspect of extensibility because it ensures that an extension does not affect the entire package but rather one or few modules. It also allows the user to concentrate on understanding those modules and their scientific background that are actually used for the particular problem under investigation.

Generality: To serve as a general research tool, ESPResSo needs to be able to handle a wide variety of problems. This includes different topologies, short- and long-range interactions, external fields, constraints, different boundary conditions, and various methods like MD or MC. In order to treat large scale simulations, it is also necessary to have an efficient parallelized code which runs on multiple CPU architectures. The abilities of the parallelization is demonstrated in 6.2.

The connectivity between the particles, which in other programs is often stored as a global topology, is incorporated locally at every particle. In this way the user is free specifying the needed topologies on the script level, and the connectivity information can easily be parallelized. Since this allows for any kind of topology, additional concepts such as molecules, polymer chains or proteins are unnecessary. There are however a number of predefined sample scripts and auxiliary routines provided for the user's convenience which set up polymer chains, simple model networks, or more complex structures such as fullerenes, representing tutorial-like shortcuts that facilitate writing new task scripts. Since both single molecule experiments as well as investigations of confined systems have gained increasing importance over the last years, ESPResSo also contains features to handle corresponding simulations. The program is able to deal with periodic boundary conditions in any combination of up to three spatial directions using any of the previously described methods. When simulating bulk systems, e.g. a small representative portion of a solution, normally periodic boundary conditions are applied to avoid boundary effects. For simulations of thin films or surface effects, periodic boundary conditions in all three spatial dimensions do not make sense. The proper boundary conditions are periodic only in

two out of the three spatial dimensions, while the remaining coordinate has a non periodic boundary condition. If rods are the object of interest, only one coordinate is left to have periodic boundary conditions. As we have elucidated in the previous sections, the complexity of an electrostatic simulation dramatically changes with different boundary conditions.

In the case of a thin film, the particles have to be confined to a fixed layer. For this ESPResSo supports constraints like walls, cylinders or spheres. It is also possible to simulate particles subject to external forces or fields. In order to cover a wide range of thermodynamic environments one can switch between simulating different thermodynamic ensembles like the NVE-, NVT- or NPT-ensemble.

Documentation: ESPResSo is not intended to be a black-box-like package. Users are encouraged to try to understand its algorithms and routines, developers are strongly advised to do so before extending it. In order to preserve the knowledge about algorithms and their physical/chemical background, it is important to provide and maintain a well structured documentation. This is mainly done inside the code itself and then extracted and processed by `doc++`⁷⁶ into a user-friendly html-manual mainly addressing specific code-/function-/procedure-related issues. It is supplemented by a stand-alone documentation on general topics such as the usage of the script commands, the general organization of the data structures, communication schemes, and analysis options. The code development itself is done in a CVS-environment (concurrent version system⁷⁷). This helps to keep track of all changes, and provides information on what, when and by whom something has changed in the program.

Programming environment: We decided to use C as the only programming language in order to keep the code as simple to read as possible. Compared to C++ we think that this is still the language of choice in a research environment since it is easier to learn for people having a natural science rather than a computer science background. At the same time it provides all necessary features to create a modular and concise program package.

We use Tcl⁷⁴ as the script language since it contains a simple and effective interface to build C programs as extensions to the script language itself. Syntax and programming style are similar to C which makes it easy to learn. Another advantage is that there exist a large variety of extensions for Tcl. For example, with the Tk-extension it is straight forward to build a graphical user interface. This has already been done for presentation purposes. Tcl also gives us the possibility to easily create interfaces to other programs. Examples are gnuplot or xmgr for graphical processing of analysis results.

Another important choice was the type of communication for the parallelization. We decided on using MPI, as it is available for virtually all architectures; unlike e.g. OpenMP which requires shared memory, MPI also works on distributed memory computers such as Linux clusters. ESPResSo relies on the fact that MPI-implementations are normally well optimized for the underlying architecture.

Ease of use: For the ESPResSo-package to live up to its full potential, a straightforward and simple access is mandatory. The layered hierarchical program structure allows the user to focus on any aspect of his scientific simulation. This can be for example modeling complex physical systems with particle insertion/deletion, pressure-dependent volume changes and/or varying constraints can be done by simply creating a corresponding script file which specifies the basic rules of such a setup. Or it can be tweaking computational routines to utmost performance can be achieved by simply adding/modifying/replacing one single module of ESPResSo, immediately granting other users access to that im-

provement.

6.2 Benchmarks

Even though the primary goals of the ESPResSo-package are accessibility, modularity, flexibility, and extensibility, its secondary -and equally important- objective accounts for the scientific realities of tight schedules and time constraints: Being as optimized for speed as possible without sacrificing its primary benefits. Considering the timings we measured, the state-of-the-art algorithms adapted meet both demands extremely well. The direct comparison of benchmark timings of some test scenarios to those of the corresponding highly specialized codes show that despite the unique nature of ESPResSo representing a very general multi-purpose tool, it still performs similarly (e.g. compared to LAMMPS, but around 1.5 times slower than polyMD) with a firm robustness among different architectures (e.g. AMD, Intel, IBM, Alpha), wrapper (e.g. mpicci, mpich, mpicc), and operating systems (e.g. AIX, OSF1, Linux). To our knowledge there is no other available program package similar in scope and design of ESPResSo. Any potential difference in performance to highly specialized codes is negligible compared to either the time to be spend for producing even faster code (particularly when recalling ESPResSo's goal to remain understandable, prohibiting most low-level trickery), or to the advancements in hardware technology, or simply overcompensated for by ESPResSo's inherent design advantages since e.g. the modularity allows for an easy incorporation of any algorithmical improvements which might arise in the future.

Besides absolute speed, another fundamental feature of our scientific simulation system is its intrinsic parallelizability, which also distinguishes it from other projects: In ESPResSo the choices of data structures and algorithm implementations were optimized in this respect, so that the program is now able to use any reasonable number of processors on any computer system supporting one of the available MPI-environments. To demonstrate the parallel performance, Table 1 presents benchmarking results for three standard test scenarios. Note that the charged systems used the P³M routines, which is why for

# of Processors N	1	2	4	8	16	32
LJ-system (fixed)	1.0000	0.9888	0.9743	0.9587	0.9171	0.8437
LJ-system (scaled)	1.0000	0.9873	0.9726	0.9166	0.8913	0.8211
dense ES (fixed)	1.0000	0.9738	0.9475	0.8817	0.8115	0.6946
dense ES (scaled)	1.0000	0.9637	0.9955	0.7609	0.7446	0.5033
dilute ES (fixed)	1.0000	0.8722	0.8307	0.7343	0.5992	0.4594
dilute ES (scaled)	1.0000	0.8574	0.6823	0.6597	0.6080	0.3549

Table 1. Efficiency of the ESPResSo-code on an IBM Regatta H Server (eServer 690 Model 681 with 32 Power4 Processors at 1.3 GHz each) for three different systems: A neutral LJ fluid composed of either $32000 \cdot N$ (scaled) or 32000 particles (fixed) at a density of 0.8442 in a NVE-ensemble; a dense electrolyte system with either $2000 \cdot N$ (scaled) or 8000 particles (fixed) at a density of 0.07, friction and temperature of 1.0, Bjerrum length of 20.0 in a NVT-ensemble; a dilute electrolyte system which differs only in density ($1 \cdot 10^{-4}$), Bjerrum length (2.0), and fixed size (16000 particles). After warm-up and equilibration period the execution time T_N for integrating 1000 time steps (integrate 1000) was measured and compared between using one and N nodes: For a fixed-size system the efficiency reads $\frac{T_1/N}{T_N}$, for scaled-size $\frac{T_1}{T_N}$.

$N = 4$ the scaled dense ES-system seems more efficient compared to the scaled case at $N = 2$ or to the fixed system because the performance critically depends on the choice of P³M-parameters whose optimizability in turn depends on the number of particles and processors.

6.3 Further Developments

As of this writing the ESPResSo-package continues to undergo significant enlargement. We are currently implementing a standard dipolar Ewald sum⁶⁵, which will enhance the capabilities to simulate ferrofluids or dipolar fluids like simple water models, and add an enhanced leap-frog algorithm for the rotational degrees of freedom. Also work has started to include an anisotropic short range potential, namely the Gay-Berne potential, which will allow the study of liquid crystals.

For the dynamics of soft matter systems it is often necessary to include hydrodynamic interactions. Since in practice one cannot include all molecular details of the systems this can be achieved on a coarse grained level by coupling the solvent degrees of freedom to the simulated particles. This will be implemented via an advanced lattice Boltzmann algorithm from the group of B. Dünweg that has already proven its usefulness in polymer dynamics simulations⁷⁸. An alternative way of coarse graining hydrodynamics, called dissipative particle dynamics (DPD)⁷⁹, is based on a momentum conserving thermostat. We will implement a version according to Soddemann et al.⁸⁰. And finally, a non-equilibrium molecular dynamics algorithm⁸¹, especially useful for driven systems, will also be added.

On top of the present strengths of ESPResSo concerning the efficient treatment of electrostatics, we plan to implement two very new ideas which promise to be a significant improvement in investigating media with varying local dielectric constants: While the first is a purely local algorithm by T. Maggs⁸² which seems to be very well suited for MC simulations and can also be very useful for dense systems by using a constrained MD algorithm^c, the second is a finite difference multigrid scheme for electro- and magnetostatics^d which appears to be better for parallel applications, promising to be fast for MD algorithms as well due to its recursiveness.

There are many more improvements planned for the next year, and hopefully the capabilities of ESPResSo will grow even further once other researchers take up our idea and contribute to ESPResSo by using, customizing and extending it. More information, and the latest version can be found at <http://www.espresso.mpg.de/>

7 One Application: Polyelectrolytes in Poor Solvent

In the last section we want to present one application of a soft matter system, namely a polyelectrolyte system, where charges play a dominant role. Polyelectrolytes (PEs) are polymers which have the ability to dissociate charges in polar solvents resulting in charged polymer chains (macroion) and mobile counterions. They represent a broad and interesting class of soft matter^{83,2} that enjoy an increasing attention in the scientific community. In technical applications PEs are used as viscosity modifiers, as precipitating agents, and as

^cB. Dünweg, private communication

^dI. Tsukerman – private communication

superabsorbers. A thorough understanding of charged soft matter has become of great interest also in biochemistry and molecular biology. This is due to the fact that virtually all proteins, as well as other biopolymers such as DNA, actin, or microtubules are PEs.

Many PEs have a hydrocarbon based backbone for which water is a very poor solvent. Therefore, in aqueous solution, there is a competition between solvent quality, Coulombic interaction, and entropic degrees of freedom. The conformation of individual chains can under certain conditions assume pearl-necklace like structures^{84–86}. In earlier simulations^{87,88} we found that the polymer density can be used as a very simple parameter to separate different conformational regimes. Here we analyze in more detail the single chain behavior and the scaling of the peak in the inter-chain structure function.

7.1 Simulation Model

Our PE model and MD approach has been described in Refs.^{87,88,29,89,90} and consists of a bead spring chain of Lennard-Jones (LJ) particles. Chain monomers interact via a LJ interaction up to a distance $R_c = 2.5\sigma$ and experience an attraction with $\epsilon = 1.75 k_B T$. The Θ -point for this model is at $\epsilon = 0.34 k_B T$ ⁸⁷. The counterions interact via a purely repulsive LJ interaction. For bonded monomers we add a FENE bond potential. Charged particles at separation r interact via the Coulomb energy $k_B T \ell_b q_i q_j / r$, with $q_i = 1, (-1)$ for the charged chain monomers (counterions) and the Bjerrum length $\ell_b = e^2 / (4\pi\epsilon_S \epsilon_0 k_B T)$ (e : unit charge, ϵ_0 and ϵ_S : permittivity of the vacuum and of the solvent). We simulated various systems with several chains in the central simulation box at various monomer densities ρ_m and different values of ℓ_b . Each chain consists of $N_m = 48 \dots 478$ monomers, with a charge fraction $f = 1 \dots 1/3$. The pressure p was found to be always positive, with the pV diagram being convex at all densities, thus our simulations are stable, reach true thermal equilibrium, and reside in a one phase region.

7.2 Single Chain Properties

With the help of a specially developed cluster algorithm^{29,90} that automatically recognizes the number of pearls in a conformation, we have analyzed all equilibrium conformations in our systems. We found large coexistence regimes between structures consisting of conformations with different pearl numbers. Even a single chain shows over the course of time many transitions between different pearl structures, hence the different pearl states are not frozen or metastable. Also the position and size of the pearls and strings is constantly changing^{29,89}, compare Fig.5. We furthermore found, that the lower the pearl number, the stronger the counterions are attracted to the pearls. This is easy to rationalize, since smaller number of pearls mean larger pearls due to mass conservation, and thus to a higher local charge density. The integrated ion distribution versus the chain distance displays an inflection point, which is a signal of counterion condensation⁵⁴. In contrast to analytical theories⁹¹, the pearl structures are stable, even though there are counterions localized near and/or inside the pearls.

When one starts in a necklace conformation and increases ℓ_b the counterions will get attracted more and more towards the chain. Scaling theories have predicted that with the onset of condensation the necklace state should collapse in a first order transition into the globular state^{92,86,91,93}. However, the 'onset' of condensation is not a sharp border, rather

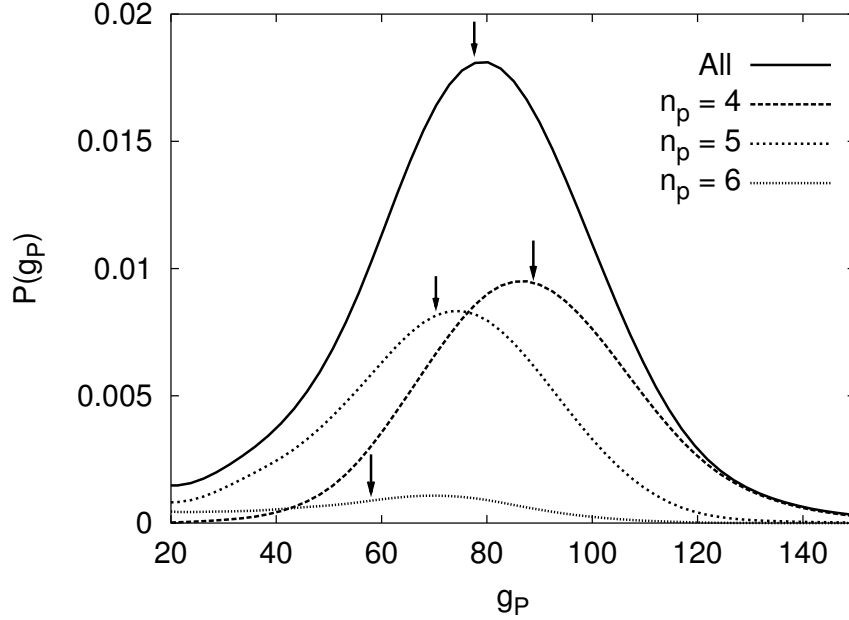


Figure 5. Probability distributions P for the pearl size g_P for a system with chain length $N_m = 382$, $\ell_b = 1.5\sigma$, and $f = 1/3$. Shown is the distribution for all chains as well as the distributions for the different structure types. The arrows mark the mean value of the corresponding probability distribution.

like within Poisson-Boltzmann theory one finds a smooth distribution of counterions which gets weighted closer to the macroion as the coupling is increased⁵⁴. Accordingly, in the simulation we do not observe a collapse transition. The picture is qualitatively the same as in the good solvent case⁹⁴. At $\ell_b = 0$ (no electrostatic interaction) the chain is in a collapsed conformation. By increasing ℓ_b the chain first extends up to a maximum, and then slowly shrinks back to a collapsed state. The non-monotonic behavior of the extension is qualitative the same as in the good solvent case⁹⁴ however the decrease is faster and more pronounced here^{87,95}. There is also a subtle dependence on f . The scaling variable which determines R_E of the necklace is $f^2\ell_b$ at fixed N_m and ϵ ⁸⁶. In Fig.6 we show snapshots of chains with the same value of $f^2\ell_b$, but different f . The chain extension is drastically different, and depends on the local interactions mediated by the counterions. This effect is obviously not captured by the scaling Ansatz! Note also, that the conformations on the way to the collapsed globule are very much reminiscent of a cylindrical shape⁹², since the strings become very short, and the pearls coalesce slowly on a sausage like string until they reach the globular state. The collapsed state is reached roughly at the same value of ℓ_b , which is reminiscent of the critical point of a Coulomb fluid.

Next we computed the spherically averaged form factor $S_1(q)$ of a single chain, shown in Fig.7, since this is an observable that is accessible in experiments, and for which also theoretical expressions have been developed^{86,96,97}.

In the range $1 < q\sigma < 2$ we denote a sharp decrease in S_1 , which reflects the intra pearl scattering, because it shows the typical Porod scattering of $S_1(q) \simeq q^{-4}$. The kink

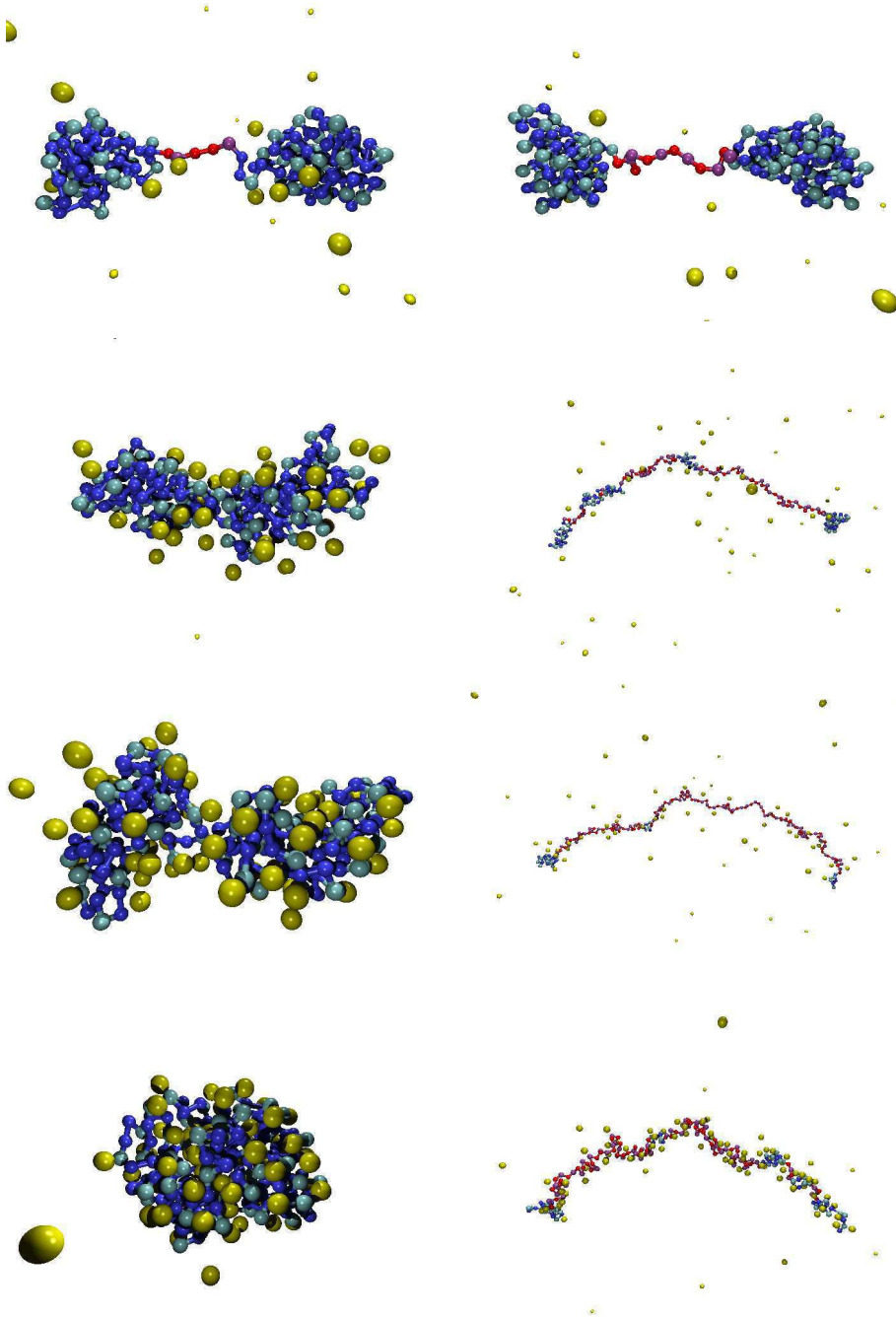


Figure 6. Snapshots for different values of the scaling variable $f^2 \ell_b$. Left row with $f = 1/3$, right row with $f = 1/2$. From top to bottom $f^2 \ell_b$ has the values: 0.08σ , 0.25σ , 0.5σ , 1.0σ . System: 8 chains with $N_m = 199$ monomers at $\rho_m = 5.0 \times 10^{-5} \sigma^{-3}$.

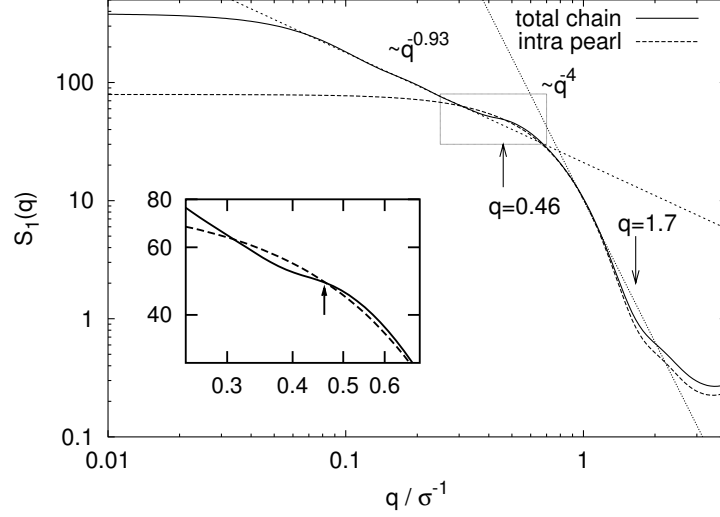


Figure 7. Spherically averaged form-factor $S_1(q)$. Shown are the single chain form-factor (solid line), together with the part of the form factor coming from the intra-pearl scattering (dashed line). The dotted and short dashed fits show the elongated chain part, and the Porod scattering part (globular conformation). System: 7 chains with 382 monomers, $f = 1/3$, $\ell_b = 1.5\sigma$, $\rho_m = 1.0 \times 10^{-5}\sigma^{-3}$.

at $q\sigma \approx 1.66$ appears at the position expected from the pearl size, but is broadly smeared out due to large size fluctuations. The shoulder which can be seen at $q\sigma \approx 0.5$ does not come from the intra-pearl scattering but is due to the scattering of neighboring pearls along the chain (inter-pearl contribution), which have a mean distance of $\langle r_{PP} \rangle = 13.3\sigma$. It is also smeared out due to the large distribution of inter-pearl distances. We conclude that the signatures of the pearl-necklaces are weak already for monodisperse samples. A possible improvement could be achieved for chains of very large molecular weights and low pearl numbers, which could lead to stable and large signatures.

7.3 Solution Properties: Scaling of the Correlation Length ξ

The overall scattering function $S(q)$ of the solution contains additional experimental information. For good solvent PEs, experiments⁹⁸, theory⁹⁹, and simulations⁹⁴ find a pronounced first peak of $S_{IC} = S/S_1$ at $q^* = (2\pi)/\xi$, where ξ is the correlation length. The position varies as $q^* \propto \rho_m^{1/3}$ in the very dilute regime and crosses over to a $\rho_m^{1/2}$ regime at higher concentrations. In Fig. 8 we have plotted the density dependence of q^* in poor solvent for different chain lengths. Within the error bars we find that for poor solvent chains q^* scales proportional to $\rho_m^{0.35 \pm 0.04}$ for *all* concentrations and chain lengths. This is in accord with very recent experiments¹⁰⁰, but theoretically not well understood. The response of the PE conformation to density changes is much larger in the poor solvent case^{87,29} than in the good solvent case⁹⁴, and the chain extension behaves non-monotonic as a function of density^{87,29}. Furthermore, in the density regime between $\rho_m\sigma^3 = 10^{-2} \dots 10^{-4}$ the chain extension and the pearl number varies most strongly, and almost all monomers

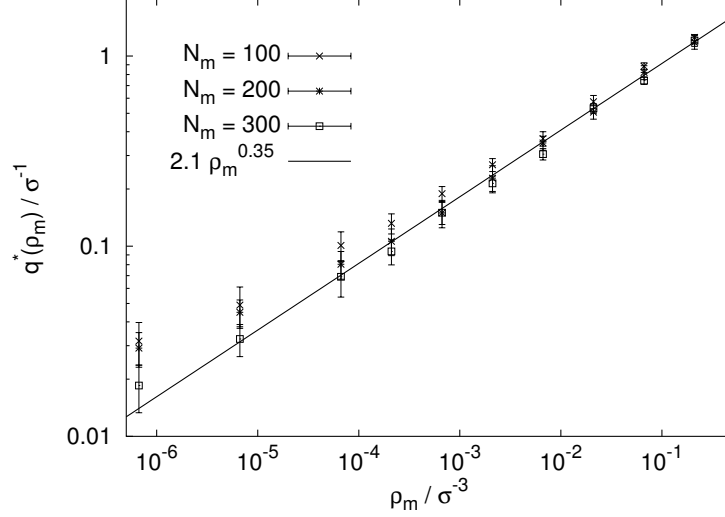


Figure 8. Density dependence of the peak q^* in the structure factor for three different chain length $N_m = 100, 200, 300$ with $f = 0.5$ and $\ell_b = 1.5\sigma$. The black line is a fit to the data with $N_m = 200$.

are located within the pearls. Upon approaching the dense regime, the string length tends to zero and we find a chain of touching pearls, indicating that the conventional necklace picture breaks down. Our result is compatible to scaling exponents found in scattering experiments^{96, 101, 102}.

Scaling theories^{103, 104} have predicted a $\rho_m^{1/2}$ regime to start at ρ_o^* , which is defined at the density where $R_E \approx \xi$, and to extend until $\xi \approx r_{pp}$ where a bead-controlled $\rho_m^{1/3}$ regime starts. We find $\rho_o^* \sigma^3 \simeq 5 \times 10^{-2}, 10^{-3}, 10^{-4}$ for $N_m = 100, 200, 300$. A pearl-pearl separation of the order of the correlation length, $r_{pp} \approx \xi$, is reached between $\rho_m \sigma^3 = 10^{-2}$ and 10^{-1} , which is roughly independent of N . Especially for the longer chains ($N_m = 200, 300$) a clear signature of a different power law, i.e. $\rho_m^{1/2}$, should be visible. One possible reason for our different findings is that the strong inter-chain coupling and the influence of the counterions on the conformations are not sufficiently taken into account in the (mean field) scaling approach. It is not clear at this stage if the $\rho_m^{1/2}$ regime can be recovered for much longer chain length. In addition we observe that the chains form a transient physical network at $\rho_m \sigma^3 = 0.2$ for $N_m \geq 200$ which has neither been seen in previous simulations nor predicted by theoretical approaches but is in accord with experimental studies^{96, 101, 102}. During the simulation time these networks reconstruct several times, e.g. chains are not trapped!

A more detailed account of the presented material can be found in the work of Limbach et al.^{89, 105, 90, 106, 107}.

8 Concluding Remarks

We have given a review of methods to compute long range interactions in fully or, for the case of the Coulomb sum, also partially periodic boundary conditions. All methods presented in more detail, have in common, that analytic error estimates exists, which allow easy tuning of the algorithms for speed and accuracy. For our research purposes, these methods are more or less the optimal choices at the moment. They are, or will be shortly, implemented in a newly designed program package, called ESPResSo which will be our main simulation program for the future. This program is open source, and informations regarding the latest version, or informations on how to participate in the further development of this package can be found on <http://www.espresso.mpg.de/>. Finally we have given a short summary of our recent findings about the conformational properties of polyelectrolytes in poor solvent.

Acknowledgments

This article would have been impossible without the contributions of A. Arnold, J. Dejoannis, M. Deserno, H.J. Limbach, B.A. Mann, and Z. Wang. Helpful discussions with K. Kremer and other members of our pep group are gratefully acknowledged. This work was supported by the “Zentrum für Multifunktionelle Werkstoffe und Miniaturisierte Funktionseinheiten”, grant BMBF 03N 6500, and the DFG within the SFB 625, TR6, and grant HO-1108/11-1.

References

1. Richard A. L. Jones. *Soft Condensed Matter*. Oxford University Press, Great Clarendon Street, Oxford OX2 6DP, 2002.
2. Christian Holm, Patrick Kékicheff, and Rudi Podgornik, editors. *Electrostatic Effects in Soft Matter and Biophysics*, volume 46 of *NATO Science Series II - Mathematics, Physics and Chemistry*. Kluwer Academic Publishers, Dordrecht, NL, December 2001.
3. P. P. Ewald. Die Berechnung optischer und elektrostatischer Gitterpotentiale. *Ann. Phys.*, 64:253–287, 1921.
4. D. M. Heyes. Electrostatic potentials and fields in infinite point charge lattices. *J. Chem. Phys.*, 74(3):1924–1929, February 1981.
5. Simon W. de Leeuw, John W. Perram, and E. R. Smith. Simulation of electrostatic systems in periodic boundary conditions. i. lattice sums and dielectric constants. *Proc. R. Soc. Lond. A*, 373:27–56, 1980.
6. Simon W. de Leeuw, John W. Perram, and E. R. Smith. Simulation of electrostatic systems in periodic boundary conditions. ii. equivalence of boundary conditions. *Proc. R. Soc. Lond. A*, 373:57–66, 1980.
7. J. Perram, H. G. Petersen, and Simon de Leeuw. An algorithm for the simulation of condensed matter which grows as the $3/2$ power of the number of particles. *Mol. Phys.*, 65:875, 1988.

8. Markus Deserno and Christian Holm. How to mesh up Ewald sums. i. a theoretical and numerical comparison of various particle mesh routines. *J. Chem. Phys.*, 109:7678, 1998.
9. Markus Deserno and Christian Holm. How to mesh up Ewald sums. ii. an accurate error estimate for the p3m algorithm. *J. Chem. Phys.*, 109:7694, 1998.
10. M. Deserno. *Counterion condensation for rigid linear polyelectrolytes*. PhD thesis, Universität Mainz, 2000.
11. Axel Arnold and Christian Holm. Mmm2d: A fast and accurate summation method for electrostatic interactions in 2d slab geometries. *Comp. Phys. Comm.*, 148(3):327–348, 1 November 2002.
12. Axel Arnold, Jason de Joannis, and Christian Holm. Electrostatics in periodic slab geometries i. *J. Chem. Phys.*, 117:2496–2502, 2002.
13. Jason de Joannis, Axel Arnold, and Christian Holm. Electrostatics in periodic slab geometries ii. *J. Chem. Phys.*, 117:2503–2512, 2002.
14. Daan Frenkel and Berend Smit. *Understanding Molecular Simulation*. Academic Press, San Diego, second edition, 2002.
15. Mike P. Allen and Dominik J. Tildesley. *Computer Simulation of Liquids*. Oxford Science Publications. Clarendon Press, Oxford, 1 edition, 1987.
16. Jean-Michel Caillol. Comments on the numerical simulation of electrolytes in periodic boundary conditions. *J. Chem. Phys.*, 101(7):6080–6090, 1994.
17. Stefan Boresch and Othmar Steinhauser. Presumed versus real artifacts of the ewald summation technique: The importance of dielectric boundary conditions. *Ber. Bunsenges. Phys. Chem.*, 101(7):1019–29, 1997.
18. H. J. C. Berendsen. In Wilfred F. van Gunsteren, P. K. Weiner, and A. J. Wilkinson, editors, *Computer Simulation of Biomolecular Systems*, volume 2, pages 161–81, The Netherlands, 1993. ESCOM.
19. Philippe H. Hünenberger. Optimal charge-shaping functions for the particle-particle-particle-mesh (p3m) method for computing electrostatic interactions in molecular simulations. *J. Chem. Phys.*, 113(23):10464–10476, 2000.
20. Jiri Kolafa and John W. Perram. Cutoff errors in the ewald summation formulae for point charge systems. *Molecular Simulation*, 9(5):351–68, 1992.
21. Gerhard Hummer, Lawrence R. Pratt, and Angle E. García. Molecular theories and simulation of ions and polar molecules in water. *J. Phys. Chem. A*, 102(41):7885–97, 1998.
22. M. J. L. Sangster and M. Dixon. Interionic potentials in alkali-halides and their use in simulations of molten-salts. *Adv. Phys.*, 25:247, 1976.
23. D. J. Adams and G. S. Dubey. Taming the ewald sum. *J. Comp. Phys.*, 72:156, 1987.
24. W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. *Numerical Recipes in C*. Cambridge University Press, Cambridge, 2 edition, 1992.
25. R. W. Hockney and J. W. Eastwood. *Computer Simulation Using Particles*. IOP, London, 1988.
26. T. Darden, D. York, and L. Pedersen. Particle mesh ewald: An $n \log(n)$ method for ewald sums in large systems. *J. Chem. Phys.*, 98:10089, 1993.
27. U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, and L. Pedersen. A smooth particle mesh ewald method. *J. Chem. Phys.*, 103:8577, 1995.
28. E. L. Pollock and Jim Glosli. Comments on p^3m , fmm, and the ewald method for

- large periodic coulombic systems. *Comp. Phys. Comm.*, 95:93–110, 1996.
29. Hans Jörg Limbach. *Struktur und Eigenschaften von Polyelektrolyten im schlechten Lösungsmittel*. PhD thesis, Johannes Gutenberg Universität, Mainz, Germany, 2001.
 30. I. J. Schoenberg. *Cardinal Spline Interpolation*. Society for Industrial and Applied Mathematics, Philadelphia, 1973.
 31. Henrik G. Petersen. Accuracy and efficiency of the particle mesh ewald method. *J. Chem. Phys.*, 103(9):3668–79, 1995.
 32. J. Lekner. Summation of coulomb fields in computer simulated disordered systems. *Physica A*, 176:485–498, 1991.
 33. R. Sperb. An alternative to ewald sums - part i: Identities for sums. *Molecular Simulation*, 20(3):179–200, 1998.
 34. R. Sperb. An alternative to ewald sums, part 2: The coulomb potential in a periodic system. *Molecular Simulation*, 22(3):199–212, 1999.
 35. R. Strebel and R. Sperb. An alternative to ewald sums, part 3: Implementation and results. *Molecular Simulation*, 27(1):61–74, 2001.
 36. J. E. Barnes and P. Hut. A hierarchical $O(n \log n)$ force/calculation algorithm. *Nature*, 324:446, 1986.
 37. L. Greengard and V. Rhoklin. A fast algorithm for particle simulations. *J. Comp. Phys.*, 73:325, 1987.
 38. K. Esselink. A comparison of algorithms for long-range interactions. *Comp. Phys. Comm.*, 87:375–395, 1995.
 39. C. Sagui and T. Darden. Multigrid methods for classical molecular dynamics simulations of biomolecules. *J. Chem. Phys.*, 2001.
 40. I. Tsukermann. Flexible local approximation method for electro- and magnetostatics. In *IEEE Trans. Magn.*, to appear in 2004.
 41. Albert H. Widmann and David B. Adolf. A comparison of ewald summation techniques for planar surfaces. *Comp. Phys. Comm.*, 107:167–186, 1997.
 42. M. Kawata and U. Nagashima. Particle mesh ewald method for three-dimensional systems with two-dimensional periodicity. *Chem. Phys. Lett.*, 340:165–172, 2001.
 43. Mazars M. Comment on “rapid calculation of the coulomb component of the stress tensor for three-dimensional systems with two-dimensional periodicity”. *J. Chem. Phys.*, 117(7):3524, Aug. 2002.
 44. Mazars M. Lekner summations. *J. Chem. Phys.*, 115(7):2955, Aug 2001.
 45. Axel Arnold and Christian Holm. A novel method for calculating electrostatic interactions in 2d periodic slab geometries. *Chemical Physical Letters*, 354:324–330, 2002.
 46. J. C. Shelley and G. N. Patey. Boundary condition effects in simulations of water confined between planar walls. *Mol. Phys.*, 88:385, 1996.
 47. E. Spohr. Effect of electrostatic boundary conditions and system size on the interfacial properties of water and aqueous solutions. *J. Chem. Phys.*, 107:6342, 1997.
 48. In-Chul Yeh and Max L. Berkowitz. Ewald summation for systems with slab geometry. *J. Chem. Phys.*, 111(7):3155–3162, 1999.
 49. P. Mináry, M. E. Tuckerman, K. A. Pihakari, and G. J. Martyna. A new reciprocal space based treatment of long range interactions on surfaces. *J. Chem. Phys.*, 116:5351, 2002.
 50. E. R. Smith. Electrostatic energy in ionic crystals. *Proc. R. Soc. Lond. A*, 375:475–

- 505, 1981.
51. D. E. Parry. Electrostatic potential in surface region of an ionic-crystal. *Surf. Sci.*, 49:433, 1975.
 52. D. E. Parry. Correction. *Surf. Sci.*, 54:195, 1976.
 53. Axel Arnold. Berechnung der elektrostatischen wechselwirkung in 2d + h periodischen systemen. Diploma thesis, Johannes Gutenberg-Universität, may 2001.
 54. Markus Deserno, Christian Holm, and Sylvio May. The fraction of condensed counterions around a charged rod: Comparison of Poisson-Boltzmann theory and computer simulations. *Macromolecules*, 33:199–206, 2000.
 55. Markus Deserno, Axel Arnold, and Christian Holm. Attraction and ionic correlations between charged stiff polyelectrolytes. *Macromolecules*, 36(1):249–259, 2003.
 56. A. Naji, A. Arnold, C. Holm, and R. R. Netz. Attraction of similarly-charged rods: Simulations and strong coupling theory. submitted to *Europhys. Lett.*
 57. Markus Porto. Ewald summation of electrostatic interactions of systems with finite extent in two of three dimensions. *J. Phys. A*, 33:6211, 2000.
 58. D. J. Langridge, J. F. Hart, and S. Crampin. Ewald summation technique for one-dimensional charge distributions. *Comp. Phys. Comm.*, 134:78, 2001.
 59. Axel Arnold and Christian Holm. Mmm1d. in preparation.
 60. A. Bródka. Ewald type summation method for electrostatic interactions in computer simulations of a three-dimensional system periodic in one direction. *Chem. Phys. Lett.*, 2002.
 61. M. Kawata and M. Mikami. Rapid calculation of two-dimensional ewald summation. *Chem. Phys. Lett.*, 340:157–164, 2001.
 62. M. Deserno, C. Holm, and K. Kremer. *Molecular dynamics simulations of the cylindrical cell model*, volume 99 of *Surfactant science series*, chapter 2, pages 59–110. Marcel Dekker, New York, 2001.
 63. M. Deserno and C. Holm. Theory and simulations of rigid polyelectrolytes. *Mol. Phys.*, 100(18):2941–2956, September 2002.
 64. R. E. Rosensweig. *Ferrohydrodynamics*. Cambridge Univ. Press, Cambridge, 1985.
 65. Z. W. Wang and C. Holm. Estimate of the cutoff errors in the ewald summation for dipolar systems. *J. Chem. Phys.*, 115:6277–6798, 2001.
 66. N. Boghossian, O. Kohlbacher, and H.-P. Lenhof. Ball: Biochemical algorithm library. Research report, Max-Planck-Institut für Informatik, Saarbrücken, 1999.
 67. C. J. Lejdfors. Gismos home page, 1998. <http://www.teokem.lu.se/gismos/>.
 68. W. F. van Gunsteren. Gromos96 homepage, 1996. <http://www.igc.ethz.ch/gromos/>.
 69. S. J. Plimpton. Fast parallel algorithms for short-range molecular dynamics. *J. Comp. Phys.*, 117:1–19, 1995.
 70. M. Nelson, W. Humphrey, A. Gursoy, A. Dalke, L. Kale, R. D. Skeel, and K. Schulten. Namd— a parallel, object-oriented molecular dynamics program. *Int. J. Supercomput. Ap.*, 10(4):251–268, 1996.
 71. M. Pütz and A. Kolb. Optimization techniques for parallel molecular dynamics using domain decomposition. *Comp. Phys. Comm.*, 113:145–167, 1998.
 72. M. Doi. Octa homepage, 2003. <http://octa.jp/OCTA/whatsOCTA.html>.
 73. A. Arnold, B. A. Mann, H. J. Limbach, and C. Holm. Espresso - an extensible

- simulation package for research on soft matter systems. *to be published*, 2004.
74. Tcl/Tk. Homepage, 2003. <http://tcl.activestate.com/>.
 75. VMD. Homepage, 2003. <http://www.ks.uiuc.edu/Research/vmd/>.
 76. R. Wunderling and M. Zöckler. DOC++: A documentation system for C/C++ and Java, 2003. <http://www.zib.de/Visual/software/doc++/>.
 77. CVS. Concurrent versions system - homepage, 2003. <http://www.cvshome.org/>.
 78. P. Ahlrichs and B. Dünweg. Simulation of a single polymer chain in solution by combining lattice boltzmann and molecular dynamics. *J. Chem. Phys.*, 111(17):8225–8239, 1999.
 79. Pep Español and Patrick Warren. Statistical mechanics of dissipative particle dynamics. *Europhys. Lett.*, 30:191, 1995.
 80. T. Soddemann, B. Dünweg, and K. Kremer. Dissipative particle dynamics: A useful thermostat for equilibrium and nonequilibrium molecular dynamics simulations. *Phys. Rev. E*, 68:046702, 2003.
 81. Florian Müller-Plathe. Reversing the perturbation in non-equilibrium molecular dynamics: An easy way to calculate the shear viscosity of fluids. *Phys. Rev. E*, 59:4894–4899, 1999.
 82. A. C. Maggs and V. Rosseto. Local simulation algorithms for coulombic interactions. *Phys. Rev. Lett.*, 88:196402, 2002.
 83. M. Hara, editor. *Polyelectrolytes: Science and Technology*. Marcel Dekker, New York, 1993.
 84. Y. Kantor and M. Kardar. Excess charge in polyampholytes. *Europhys. Lett.*, 27:643, 1994.
 85. Y. Kantor and M. Kardar. Instabilities of charged polyampholytes. *Phys. Rev. E*, 51:1299, 1995.
 86. A. V. Dobrynin, M. Rubinstein, and S. P. Obukhov. Cascade of transitions of polyelectrolytes in poor solvents. *Macromolecules*, 29(8):2974, 1996.
 87. U. Micka, C. Holm, and K. Kremer. Strongly charged, flexible polyelectrolytes in poor solvents - a molecular dynamics study. *Langmuir*, 15:4033, 1999.
 88. H. J. Limbach and C. Holm. End-effects of strongly charged polyelectrolytes - a molecular dynamics study. *J. Chem. Phys.*, 114(21):9674–9682, 2001.
 89. H. J. Limbach and C. Holm. Conformational properties of poor solvent polyelectrolytes. *Comp. Phys. Comp.*, 147:321–324, 2002.
 90. H. J. Limbach and C. Holm. Single-chain properties of polyelectrolytes in poor solvent. *J. Phys. Chem. B*, 107(32):8041–8055, 2003.
 91. H. Schiessel and P. Pincus. Counterion-condensation-induced collapse of highly charged polyelectrolytes. *Macromolecules*, 31:7953–7959, 1998.
 92. A. R. Khokhlov. On the collapse of weakly charged polyelectrolytes. *J. Phys. A*, 13:979, 1980.
 93. T. A. Vilgis, A. Johner, and J.-F. Joanny. Stretching necklaces. *Eur. Phys. J. E*, 2(3):289–300, 2000.
 94. M. J. Stevens and Kurt Kremer. The nature of flexible linear polyelectrolytes in salt free solution: A molecular dynamics study. *J. Chem. Phys.*, 103(4):1669–1690, 1995.
 95. M. O. Khan, S. M. Melnikov, and B. Jönsson. *Macromol.*, 32:8836, 1999.
 96. C. Heitz, M. Rawiso, and J. François. X-ray scattering study of a poly (methacrylic

- acid) sample as a function of its neutralization degree. *Polymer*, 40(7):1637–1650, 1999.
97. R. Schweins and K. Huber. preprint. *to be published*, 2004.
 98. M. Nierlich, C. Williams, F. Boue, J. Cotton, M. Daoud, B. Garnoux, G. Jannink, C. Picot, M. Moan, C. Wolff, M. Rinaudo, and P. G. de Gennes. Small-angle neutron-scattering by semi-dilute solutions of polyelectrolyte. *J. Physique*, 40:701, 1979.
 99. J. F. Joanny. Scaling description of charged systems. In C. Holm, P. Kékicheff, and R. Podgornik, editors, *Electrostatic Effects in Soft Matter and Biophysics*, volume 46 of *NATO Science Series II - Mathematics, Physics and Chemistry*, pages 149–170. Kluwer Academic Publishers, Dordrecht, NL, December 2001.
 100. D. Baigl, R. Ober, D. Qu, A. Fery, and C. E. Williams. Correlation length of hydrophobic polyelectrolyte solutions. *Europhys. Lett.*, 63(4):588–594, 2003.
 101. W. Essafi, F. Lafuma, and C. E. Williams. Effect of solvent quality on the behaviour of highly charged polyelectrolytes. *J. Phys. II*, 5:1269–1275, 1995.
 102. T. A. Waigh, R. Ober, C. E. Williams, and J.-C. Galin. Semidilute and concentrated solutions of a solvophobic polyelectrolyte in nonaqueous solvents. *Macromolecules*, 34(6):1973–1980, 2001.
 103. A. V. Dobrynin and M. Rubinstein. Hydrophobic polyelectrolytes. *Macromolecules*, 32(3):915–922, 1999.
 104. A. V. Dobrynin and M. Rubinstein. Counterion condensation and phase separation in solutions of hydrophobic polyelectrolytes. *Macromolecules*, 34(6):1964–1972, 2001.
 105. H. J. Limbach, C. Holm, and K. Kremer. Structure of polyelectrolytes in poor solvent. *Europhys. Lett.*, 60(4):566–572, 2002.
 106. C. Holm, H. J. Limbach, and K. Kremer. Poor solvent polyelectrolytes. *J. Phys.: Condens. Matter*, 15:S205–S211, 2003.
 107. H. J. Limbach and C. Holm. *in preparation*, 2004.

Phase Behavior and Chain Conformations in Polymer Blends: Monte Carlo Simulation vs Mean Field Theory

Marcus Müller

Institut für Physik, Johannes Gutenberg-Universität,
55099 Mainz, Germany
E-mail: marcus.mueller@uni-mainz.de

Using Monte Carlo simulations we investigate the phase behavior and interface properties in incompressible binary polymer blends. Special attention is focused on the relation between the Flory-Huggins parameter χ , that describes the incompatibility of the monomeric units in the mean field theory, and the particle-based model used in the simulations. The role of chain conformations on the thermodynamics and structure of the blend is investigated. We present a quantitative comparison between Monte Carlo simulations and mean field theory and the role of fluctuations in the bulk and at interfaces is explored.

1 Introduction

Melt blending of polymers is a promising route for tailoring materials to specific application properties: Polymeric materials in daily life are generally multicomponent systems. Chemically different polymers are “alloyed” as to design a material which combines the favorable characteristics of the individual components.¹ Clearly the miscibility behavior of the blend is crucial for understanding and tailoring properties relevant for practical applications. Miscibility on a microscopic length scale is desirable for a high tensile strength of the material. Unlike metallic alloys, however, chemically different polymers often do not mix on microscopic length scales. Rather a complicated morphology of droplets of one component dispersed into the other component forms on a mesoscopic length scale, and the blend can be conceived as an assembly of interfaces. While the detailed structure on this mesoscopic length scale depends strongly on the way the material is processed, the local properties of interfaces are certainly crucial for understanding the material properties. For instance, the interfacial width sets the length scale on which entanglement between polymers of the different components form. Experiments² suggest that the mechanical strength increases if the interfacial width exceeds the entanglement length. Alternatively, the interfacial tension is important for the breaking-up of droplets under shear.^{3,4} The lower the interfacial tension is the finer are dispersed the two components.

In the following we consider a dense incompressible mixture of two components A and B . Upon increasing the incompatibility between the two species the mixture will phase separate into an A -rich and a B -rich liquid. Of course, in compressible mixtures more complex phase behavior can occur. In fact, six qualitatively different types of phase diagrams can be distinguished for compressible binary mixtures according to a scheme by Konynenburg and Scott.⁵ In the following we restrict our attention to liquid-liquid immiscibility, which is characterized by a single order parameter, ϕ , being the composition of the mixture.

Much of the qualitative behavior of dense multicomponent polymer systems can be understood from a coarse-grained description: The large size of the chain molecules imparts a

rather universal behavior onto dense polymer mixtures, which can be characterized by only a small number of parameters: the end-to-end distance R_e of the molecules, and the incompatibility per chain, χN . The Flory-Huggins parameter χ describes the repulsion between unlike segments and N is the number of segments per molecule. R_e sets the characteristic length scale of spatial inhomogeneities, e.g., the width of interfaces between coexisting phases. Since this length scale is much larger than the size of the repeat units along the backbone of the polymer chain, one can use the Gaussian chain model, which captures the long-wavelength behavior of the polymer conformations in a melt. The parameters, R_e and χN , encode the chemical structure of the polymer on microscopic length scales. The properties on the mesoscopic length scale do depend on the microscopic structure only via the parameters R_e and χN .

Even in simple cases, the Flory-Huggins parameter χ results from small differences of dispersion forces between the different chemical segments and is typically only of the order $10^{-2} - 10^{-4}$ per segment. The extension of a polymer in a homogeneous melt results from a delicate screening of excluded volume interactions along the chain by surrounding molecules, and R_e depends on the density and temperature. Deriving those coarse-grained parameters, R_e and χN , from a microscopic model (e.g., atomistic force fields) is a formidable task. When these coarse-grained parameters are determined independently (e.g., by experiments) and used as an input, however, coarse-grained models are successful in making quantitative predictions.

Coarse-grained, particle-based models represent chains as a string of bonded segments on a lattice interacting via potentials of finite range. Field theoretical models represent chains as Gaussian walks in continuum space with end-to-end distance R_e and zero-ranged segmental repulsions of strength χ . By virtue of universality both are expected to yield identical descriptions for properties on mesoscopic length scales.

Computer simulations of coarse-grained, particle-based models numerically investigate the structure and thermodynamics without any approximation. Studies of the field theoretical model often invoke a mean field approximation: The self-consistent field (SCF) theory⁶⁻¹⁰ neglects fluctuations of the composition in the multicomponent system. Due to the extended shape of the molecules, one molecule interacts with many neighbors, and the mean field approximation is quite accurate under many circumstances. Let us introduce the quantity $\sqrt{\mathcal{N}} = \Phi R_e^3$, where Φ denote the polymer number density. This quantity characterizes the number of molecules inside the volume of the reference chain. It measures the degree of interdigitation of the molecules. In a dense melt, $\mathcal{N}$ is proportional to the number of segments N per chain. For systems which differ in $\mathcal{N}$ but are characterized by the same coarse-grained parameters, χN and R_e , the SCF theory will make identical predictions. The quantity $\mathcal{N}$ plays an important role as it controls the strength of fluctuations.

Of course, one can describe the properties of the blend also by the characteristics on the scale of monomeric units: The number of segments N , the statistical segment length $b = R_e^2/N$, the Flory-Huggins parameter χ , and the monomer density $\rho = \Phi/N$. However, the results must not change upon representing the chains by a different number of effective segments N' . Such a reparameterization leaves the combinations χN , R_e and $\mathcal{N}$ invariant.

We can address two types of questions by quantitatively comparing computer simulations of coarse-grained, particle-based polymer models with SCF theory: (i) How to identify the coarse-grained parameter, χN , for a specific, particle-based model? For which properties and parameters is the coarse-grained description of the polymer conformations

by the Gaussian chain model is valid? For which aspects are microscopic properties of the underlying model important? (ii) What is the validity of the mean field approximation (or additional approximations invoked in less accurate analytical treatments of the field theoretical model)?

In the following we shall describe computer simulations in the framework of a coarse-grained lattice model targeted at investigating the phase behavior and the local structure of the interfaces in polymer blends and systems containing diblock copolymers. The results are compared to SCF calculations. In the next section we provide some background on the qualitative behavior of these systems. Then, we describe our computational model and the Monte Carlo (MC) technique. Subsequently, we detail our results on symmetric binary blends, binary blends of polymers with different stiffness, a ternary blend of two homopolymers and a symmetric diblock copolymer.

2 Background

2.1 Bulk Phase Behavior of a Binary Blend: Flory-Huggins Theory

We consider a dense mixture of polymers that comprises two species - A and B . Each molecule comprises N segments, and unlike segments repel each other with strength χ . The free energy of mixing per molecule takes the form:

$$\frac{F_{\text{FH}}(\phi)}{T\Phi V} = \phi \ln \phi + (1 - \phi) \ln(1 - \phi) + \chi N \phi(1 - \phi), \quad (1)$$

where Φ denotes the number density of polymers and ϕ the composition of the incompressible mixture. T is the temperature and we set Boltzmann's constant $k_B = 1$. The first two terms describe the entropy of mixing; it stems entirely from the translational entropy. The last term describes the energy of mixing per molecule. Note that this Flory-Huggins free energy¹¹ does not include any contribution from the conformational entropy of the extended macromolecules. Implicitly, one assumes that the conformations of a single chain in a homogeneous system are independent from the environment (i.e., local composition and temperature).

When χ is used as an adjustable parameter or extracted from experiments, this mean field theory¹¹ of Flory and Huggins is quite successful in describing many experimental observations. Most notably the theory rationalizes the fact that long macromolecules tend to demix at very high temperatures. The critical point of the blend is located at:

$$\frac{N}{T_c} \sim \chi_c N = 2 \quad \text{and} \quad \phi_c = \frac{1}{2}, \quad (2)$$

The Flory-Huggins theory provides simple analytical expressions for the free energy of mixing. This simple form of the bulk free energy lays at the basis of SCF calculations⁶⁻¹⁰ for spatially inhomogeneous polymer systems in the framework of the Gaussian chain model. Nevertheless, the Flory-Huggins theory cannot rationalize the following observations:

(i) The temperature dependence of the measured values on the χ parameter often takes the form $\chi = A + B/T$. Following a common convention, B is denoted as the enthalpic

contribution, whereas A is referred to as an additional entropic contribution. In the framework of the Flory-Huggins theory,¹¹ the only contribution to the entropy of mixing stems from the translation of the molecules as a whole. Hence, the entropy of mixing per polymer is of the order unity. Additional entropic contributions e.g., due to the dependence of the packing arrangement of the segments on the local environment, may become important. In experiments, however, the temperature dependence of χ is only accessible over a rather small temperature range and there is a strong interplay between entropic and enthalpic contributions, which makes the standard decomposition of χ into an entropic and an enthalpic part difficult: e.g., the temperature dependence of the chain conformations or equation of state effects might yield a nonlinear relation between χ and the inverse temperature. Moreover, the χ parameter often depends on composition and chain length as well.

(ii) According to the original mean field theory, the demixing temperature is independent of the molecular architecture. Hence, blends of two homopolymers and two ring or star polymers with the same number of monomeric units are predicted to have the same miscibility behavior. Similarly, the theory does not capture the dependence of the miscibility behavior of the chain stiffness or the degree of branching.^{12,13}

(iii) Being a mean field theory, the Flory-Huggins theory neglects fluctuations of the local composition and invokes a random mixing approximation. Thus, the behavior in the vicinity of the critical point is described by mean field exponents and the binodals have a parabolic shape. In the ultimate vicinity of the critical point, the correlation length grows very large and the polymeric properties become irrelevant. In this critical region the behavior is characterized by the 3D Ising universality class¹⁴ which applies to all binary mixtures with short-ranged interactions. The latter behavior manifests itself in a much flatter binodal at the critical point and a stronger divergence of composition fluctuations. This has been observed in neutron scattering experiments extremely close to the critical point.^{15,16}

(iv) The same intermolecular forces which determine the miscibility behavior alter the conformation of the extended flexible macromolecules. MC simulations^{14,17–22} for rather short chain lengths reveal a contraction of the polymer coils in the minority phase. Experiments in highly incompatible poly(methyl methacrylate) (PMMA) and poly(vinyl acetate) (PVAc) blends of rather low molecular weight indicate a relative contraction of isolated PMMA chain extensions by 13–15%.²³ These observations indicate a possible coupling between the single chain conformations and the thermodynamic state (i.e., temperature and composition of the mixture).

2.2 Interfaces Between Coexisting Phases: SCF Theory

In spatially inhomogeneous systems, the structure of the polymer molecules becomes more apparent. The SCF theory considers a mixture of Gaussian chains with end-to-end distance R_e . The statistical mechanics of the interacting multi-chain system is analytically intractable and therefore one invokes a mean field approximation: The interaction of an A -chain with its neighbors is approximated by an effective external field $W_A[\phi]$, that depends, in turn, on the local composition $\phi(\mathbf{r})$. Then, the local density of A -chains is calculated as the Boltzmann average of the single chain density in the external field, and likewise for B -polymers. This relation and $W_A[\phi]$ constitute a closed self-consistent set of equations, which allows to describe inhomogeneous systems.^{6–10}

The single chain problem in the spatially varying external field can be solved only numerically. Using the Gaussian chain model, one has to deal with a modified diffusion equation. Alternatively, one can solve the single chain problem by brute force, partially enumerating the polymer configurations. To this end, one extracts a large number of A -chain conformations (typically $10^7 - 10^8$) from the simulations and calculates the Boltzmann average of the conformations in the external field W_A , and does likewise for B -chains.^{24,25} This method has the advantage of describing the chain conformations on all length scales (including the rod-like behavior and/or self-avoiding walk like behavior on small length scales), but it is considerably more involved computationally.

In some limiting cases simple analytic expressions can be obtained:

(i) In the vicinity of the critical point, the variation of the composition can be treated as a small parameter. Expanding the spatial variation of the composition in a Fourier-basis, $\phi(\mathbf{q}) = \phi_A(\mathbf{q}) = -\phi_B(\mathbf{q})$ for $q \neq 0$, and only considering terms to quadratic order, one derives the random phase (RPA) approximation for the free energy per molecule:²⁶

$$\frac{F_{\text{RPA}}[\phi(\mathbf{q})]}{T\Phi V} = \frac{1}{2} \sum_{\mathbf{q}} \phi(\mathbf{q}) \left\{ \frac{1}{\phi g(\mathbf{q})} + \frac{1}{(1-\phi)g(\mathbf{q})} \right\} \phi(-\mathbf{q}) - \frac{1}{2} \sum_{\mathbf{q}} \phi(\mathbf{q}) 2\chi N \phi(-\mathbf{q}) , \quad (3)$$

where the Debye function g characterizes the single chain structure factor of a Gaussian chain

$$g(\mathbf{q}) = \frac{12}{q^2 R_e^2} (\exp[-q^2 R_e^2/6] - 1 + q^2 R_e^2/6) . \quad (4)$$

The first two terms in Eq.(3) represent the conformational entropies of the two polymer species, the last term is the energy of mixing.

Calculating fluctuations of the composition around the homogeneous phase above the unmixing transition temperature, we obtain for the correlation length ξ

$$\frac{\xi}{R_e} = \frac{1}{\sqrt{18[1 - 2\chi N \phi(1 - \phi)]}} . \quad (5)$$

For larger incompatibilities $\chi N > \chi_c N = 2$, the blend phase separates into an A -rich and a B -rich bulk. The two phases are separated by an interface, which is described by a tanh-profile

$$\phi(z) = \frac{1}{2} (1 + \tanh[z/w_{\text{WSL}}]) \quad \text{with} \quad w_{\text{WSL}} = 2\xi \quad (6)$$

In the weak segregation limit (WSL) a polymer mixture behaves similar to a mixture of small molecules, because the characteristic size of composition variations is on the order of the molecular extension R_e , and therefore the Gaussian conformational statistics is not important, only the overall size of the molecule matters.

(ii) Far below the critical temperature, $\chi N \gg 2$, the SCF calculations take a particularly simple form (ground state dominance). In this case, the entropy loss the chain suffers at a spatial inhomogeneity can be described by the Lifshitz formula:²⁷

$$\frac{S[\phi]}{T\Phi V} = -\frac{R_e^2}{24V} \int d^3\mathbf{r} \left\{ \frac{(\nabla\phi)^2}{\phi} + \frac{(\nabla(1-\phi))^2}{1-\phi} \right\} \quad (7)$$

In this case the interface profile, $\phi(z)$, is also described by the tanh-profile like in Eq.(6), but with a width

$$\frac{w_{SSL}}{R_e} = \frac{1}{\sqrt{6\chi N}} \quad (8)$$

This result can be heuristically understood as follows: The properties of the interface are dominated by excursions of A -polymers into the B -rich phase, and vice versa. The energy cost of each loop is comparable to the thermal energy scale T . Each monomer of a loop contributes to the energy an amount of the order χT , and, consequentially, the number of monomers per loop scales as $1/\chi$. The spatial extension of the loops determines the interfacial width. In the Gaussian chain model the conformational distribution is Gaussian on all length scales, and therefore the spatial extension of loops is on the order of $R_e/\sqrt{\chi N}$. Each monomer in the interfacial region contributes to the interface tension an amount χ , and the interface tension scales as $\chi \Phi N R_e/\sqrt{\chi N}$. The final result (including prefactor of order unity) is:

$$\frac{\sigma_{SSL} R_e^2}{T} = \sqrt{N} \sqrt{\frac{\chi N}{6}} \quad (9)$$

In the strong segregation limit (SSL) loops determine the interface properties, which depend sensitively on the conformational statistics of the molecule. The SCF theory describes both limits, (i) WSL and (ii) SSL, and the crossover between them.

3 Models and Techniques

Simulations of the miscibility behavior in polymer blends are considerably more exacting in computational terms than those of small molecules or magnetic systems. The problems stem from the difficulty of dealing with the widely spread time and length scales caused by the extended structure of the macromolecules. Therefore coarse-grained models are promising candidates for investigating the universal, qualitative characteristics of miscibility. Introducing specific structural modifications or asymmetries, we can highlight their influence on the miscibility behavior. A detailed comparison between different computational models is warranted to investigate the degree of universality of the observed effects and to explore the effects of structure on various length scales. Simulational models of various degree of coarse-graining have been employed, ranging from the representation of polymers as self avoiding walks on a simple cubic lattice – as in the original treatment of Flory and Huggins¹¹ – to simulations of the effect of branching in hydrocarbon melts in the framework of a united atom model.²⁸ The choice of the simulation model is a compromise between computational efficiency and a more faithful representation of the details of molecular architecture.

We shall present MC simulations in the framework of the bond fluctuation model,²⁹ which incorporates the relevant universal characteristics of polymer blends: Connectivity of the monomers along a chain, excluded volume of the segments, and a thermal interaction between monomers. In the framework of this coarse-grained lattice model, a monomer occupies the 8 corners of a unit cell from further occupancy. Monomers along a polymer are connected by one of 108 bond vectors of length 2, $\sqrt{5}$, $\sqrt{6}$, 3, and $\sqrt{10}$. The bond vectors are chosen such that the excluded volume interactions prevents a crossing of bond

vectors during the motion. Therefore the algorithm captures the effect of entanglements. The large number of bond vectors allows for 87 different bond angles – an indication for the rather good approximation of continuous space properties by this complex lattice model. This property also allows for a rather realistic implementation of the bending rigidity.

Here and in the following all length scales are measured in units of the lattice spacing. When atomistically detailed simulations are mapped onto the bond fluctuation model a lattice unit corresponds to roughly 2 Angstrom and a monomer in the bond fluctuation model represents a small number – say 3 to 5 – of chemical repeat units.³⁰ If not noted explicitly, we work at a monomer number density of $\rho \equiv \Phi N = 1/16$, i.e., due to the extended structure of the monomers half of the lattice sites are occupied. These parameters correspond to a concentrated solution or a melt. On the one hand the presence of vacancies allows a reasonably fast equilibration of the chain conformations on the lattice. On the other hand the size disparity between vacancies and extended monomers gives rise to packing effects. Indeed, the monomer–monomer density pair correlation function $g(\mathbf{r})$ exhibits oscillations at small distances indicating a fluid–like packing due to the local compressibility. Moreover, the relation between osmotic pressure and density is well describable via the Carnahan-Starling equation³¹ – an approximation for the equation of state of hard spheres. This exemplifies that this lattice model shares many features with off–lattice models.

The conformations of the polymers on the lattice evolve via local random monomer hopping²⁹ – a randomly chosen monomer attempts to move one lattice constant in a random direction – or slithering snake-like moves^{32,33} – a segment of the chain is removed at one end of the chain and added at the opposite one. While the former allows for a dynamical interpretation of the MC simulations in terms of a purely diffusional dynamics,²⁹ the latter relaxes the chain conformations a factor N faster.

The blends comprise two components – denoted A and B . Monomeric units of the same type attract each other whereas different monomers repel each other via a square well potential

$$\epsilon = -\epsilon_{AA} = -\epsilon_{BB} = \epsilon_{AB} . \quad (10)$$

The potential is extended over the first peak of the pair correlation function, i.e., it incorporates the first 54 neighbors up to a distance $\sqrt{6}$. The form of the potential is chosen by computational convenience; we expect our results to be qualitatively independent from the specific potentials used. However, if we were to model the interactions as (strongly) attractive with (slightly) different strengths between unlike species – a more faithful modeling of interactions in view of the experiment situation – the presence of vacancies would allow for a liquid-vapor phase separation between a concentrated polymer melt and a dilute phase in our ternary system. This liquid-vapor phase separation is common to both lattice based models and models in continuous space. The temperature scale of this liquid-vapor coexistence is set by the Θ temperature, which is chain length independent. This contrasts with the temperature scale of the liquid-liquid phase separation into A -rich and B -rich phases with similar content of vacancies. The latter temperature scale increases linearly with the chain length. Therefore, the two phenomena are well separated in blends of high molecular weight. In this paper we focus on the liquid-liquid phase separation at high temperatures.

In the following, we identify the Flory-Huggins parameter χ via the energy of mixing:

$$\chi N = \Phi N^2 \int d^3\mathbf{r} \left[g_{AB}^{\text{inter}}(\mathbf{r}) v_{AB}(\mathbf{r}) - \frac{g_{AA}^{\text{inter}}(\mathbf{r}) v_{AA}(\mathbf{r}) + g_{BB}^{\text{inter}}(\mathbf{r}) v_{BB}(\mathbf{r})}{2} \right] \quad (11)$$

$$= \frac{2N z_c \epsilon}{T} \quad \text{with} \quad z_c = \Phi N \int_{r \leq \sqrt{6}} d^3\mathbf{r} g^{\text{inter}}(\mathbf{r}) \quad (12)$$

where v_{IJ} ($I, J = A, B$) denote the interaction potentials between the different segments. $g_{IJ}^{\text{inter}}(\mathbf{r})$ denotes the intermolecular paircorrelation function, which describes the probability to find a monomer of type J that belongs to a different chain a distance $\mathbf{r}$ away from a monomer of type I . In analogy to the original treatment of Flory and Huggins we refer to z_c as the effective coordination number. This equation expresses the χ parameter in terms of observables that are directly accessible in the MC simulations. Moreover, we shall often approximate the intermolecular pair correlation functions by their athermal values. It is not that the total number of interactions determine the miscibility behavior, but rather the intramolecular interactions do not contribute. Even in this simple form it is evident that the chain architecture has pronounced effects on the intermolecular paircorrelation function and the miscibility behavior: The more open the chains are, the larger is the number of intermolecular contacts, and the smaller is the miscibility.

Being a lattice model, the bond fluctuation model is highly computationally efficient. It allows for the investigation of rather large chain lengths and large system sizes. The latter is necessary to accurately locate the critical temperature via finite size scaling analysis. As we shall illustrate, the large chain length is crucial for reaching the high molecular weight scaling limit and extrapolating some quantities to experimentally relevant chain lengths. For the present investigation chains with up to 512 (2048 for athermal systems) monomeric units have been employed.

Various simulation techniques have been used for exploring the miscibility properties of polymer blends. The most direct one is the simulation of both the coexisting phases in the simulation cell. This method has been employed by Madden³⁴ and Cifra³⁵ for well segregated blends, and it also yields information about interfacial properties. However, it requires rather large simulation cells in order to extract “bulk” properties. Especially, the scheme is not very well suited to cope with the growing of the correlation length and vanishing of the difference between the phases as the critical point is approached. Computationally more efficient seems the direct estimation of the chemical potential of each individual species as a function of temperature and composition. At coexistence the chemical potentials of the species in both phases are equal and the coexistence curve can be mapped out. This technique has been applied successfully by Kumar³⁶ using the incremental chemical potential method.³⁷ Kumar explored the influence of pressure and compressibility on the miscibility behavior. If one point on the coexistence curve is known, a Gibbs Duhem integration technique can be employed to obtain the coexistence under constant pressure conditions.^{38,39} Both methodologies are particularly useful for blends in which the constituents are characterized by very different chain architecture.

Sariban and Binder¹⁴ employed simulations in the semi-grandcanonical ensemble for investigating the phase behavior at constant volume. In this ensemble, the total monomer density is fixed, the composition of the blend fluctuates, and the chemical potential difference $\Delta\mu$ between the species is controlled. The MC scheme comprises two types of moves. Canonical updates relax the conformation of the macromolecules on the lat-

tice, whereas semi-grandcanonical ones transfer A polymers into B polymers and vice-versa. Sariban and Binder investigated strictly symmetric chains, for which the semi-grandcanonical moves consists in a mere exchange of labels. The algorithm can be extended to some degree of structural asymmetry (e.g., different chain lengths between the species³²). Overall speaking, it is reasonably efficient for a modest degree of structural asymmetry between the different constituents, but the extension to pronounced structural asymmetries is a challenging task. Improvement might be achieved via gradually “mutating” one species into another.⁴⁰ The major advantages of the methodology stem from the fluctuation of the composition of the mixture:

(i) The relaxation times are much smaller than in the canonical ensemble, where the composition is conserved and composition fluctuations decay via the slow diffusion of polymers in a melt. The semi-grandcanonical ensemble allows the straightforward application of finite size scaling techniques known from simple mixtures. Therefore we can measure the critical temperature in symmetric and asymmetric mixtures accurately from the MC simulations of modest system size. (ii) Moreover, the SG-EOS which relates the composition of the mixture to the difference of the chemical potentials of the species is directly accessible with high accuracy. The latter is particularly important to establish a direct contact to analytical approaches outside of the ultimate critical region where 3D Ising critical behavior dominates – i.e., in the region where the mean field theories are applicable. (iii) Additionally, it is possible to determine excess interfacial properties (e.g., interfacial tension,⁴¹ excess energy,⁴² enrichment of a third component⁴²).

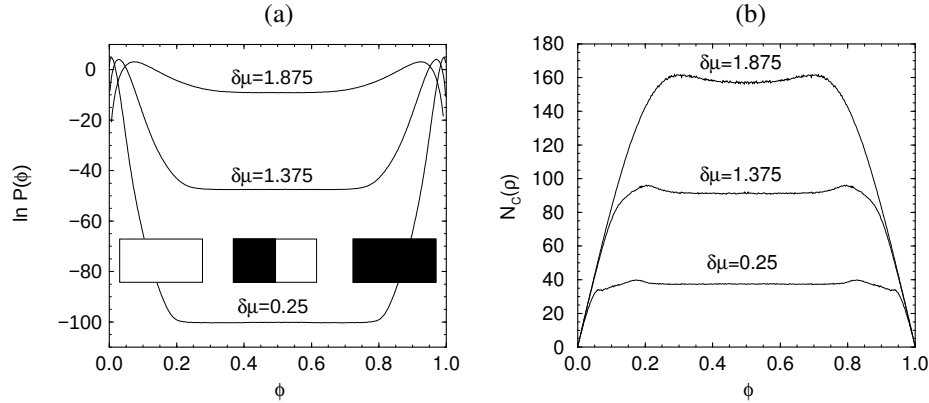


Figure 1. Ternary blend containing two homopolymers A and B and a symmetric AB diblock copolymer. (a) Probability distribution at $\epsilon = 0.054$ and system size $48 \times 48 \times 96$. Upon increasing the chemical potential $\delta\mu$ of the copolymers the “valley” becomes shallower, indicating that the copolymers decrease the interfacial tension. One clearly observes a plateau around $\rho = 1/2$. This assures, that our system size is large enough to neglect interfacial interactions in the measurement of the interfacial tension. (b) Average number of copolymers as a function of the composition. The copolymer number is enhanced in the configuration containing two interfaces. From Müller and Schick.⁴²

An important quantity in the semi-grandcanonical simulations is the probability distribution of the composition $P(\phi)$. At phase coexistence, it exhibits two peaks that correspond to the two coexisting phases. At coexistence, the exchange chemical potential

$\Delta\mu = \mu_A - \mu_B$, conjugated to the composition, has to be chosen such that both peaks have equal weight. (For a symmetric blend, $\Delta\mu_{\text{coex}} = 0$, of course).

At the critical point, the probability distribution adopts a universal shape when scaled to unit norm and variance. The shape characterizes the universality class to the transition. In case of unmixing this is the 3D Ising universality class. Adjusting the temperature (i.e., χN) as to match the probability distribution obtained from simulations and the predetermined universal curve, we accurately determine the critical temperature from the simulation.⁴³

At lower temperatures, the system has to tunnel between the two coexisting states, which are separated by a large free energy barrier (cf. Fig. 1(a)). Therefore the probability of finding the system at $\phi_{\text{middle}} = (\phi_{\text{coex}}^{(1)} + \phi_{\text{coex}}^{(2)})/2$ is very low. In the middle of the miscibility gap, the typical configuration consists of a slab of A -rich phase, which is separated by two interfaces of size L^2 from the B -rich phase. The probability of these configurations is suppressed by an amount $P_{\text{middle}}/P_{\text{coex}} = \exp(-2L^2\sigma/T)$. Thus, we not only obtain information about the coexisting phases (composition, susceptibility, and coexistence chemical potential), but also about interface properties. One can also monitor other quantities (e.g., the excess of a third species, see Fig. 1(b)) as a function of the composition and obtain the difference (excess) between the bulk and the system containing two interfaces.

In order to overcome the free energy barrier, one applies pre-weighting techniques: One adds to the original Hamiltonian $\mathcal{H}$ of the system a function $w(\phi)$ that depends on the composition, but not on the microscopic conformations of the polymers. The probability distribution of the pre-weighted system is given by:

$$P_{\text{pw}}(\phi) \sim P_{\text{orig}}(\phi) \exp(-w(\phi)/T) \quad (13)$$

Choosing $w(\phi) \approx T \ln P_{\text{orig}}(\phi)$ the system samples all compositions with roughly equal probability. The crux is that the probability distribution P_{orig} is not known a priori; it rather is the result of the calculation. Several strategies have been devised to generate estimates to be used in simulation. (i) Histogram-reweighting techniques⁴⁴ alleviate this problem by performing a sequence of weighted simulations and extrapolations starting at a point where barriers are small and the system explores a wide range of composition. More sophisticated methods combine results of multiple histograms⁴⁵. (ii) Multicanonical recursion⁴⁶ conducts a series of short trial runs. After each run, the pre-weighting factors w are adjusted until the simulation can access all pertinent values of composition. The weight factors can also be self-adjusted during the simulation^{47–50}. However, detailed balance is violated in this process and separation of statistical and systematic errors becomes difficult. (iii) Weight factors can also be obtained from the transition probabilities between macrostates^{51,52}. Alternatively, one can use successive umbrella sampling, thereby dividing the pertinent range of composition into smaller subintervals and exploring each successively.⁵³

4 Results

4.1 Phase Behavior: Fluid Structure and Composition Fluctuations

4.1.1 Symmetric blends

The scaling of the critical temperature with chain length was studied by Deutsch and Binder¹⁸ in the framework of the bond fluctuation model. In accord with the predictions of the Flory-Huggins theory,¹¹ the simulations exhibited a linear scaling of the critical temperature with chain length, which was also observed in carefully designed experiments.⁵⁴ This finding stimulated advances in the development of integral equation theories.⁵⁵ The crossover between the mean field behavior away from the critical point and the ultimate 3D Ising behavior at the critical point was unraveled via a sophisticated analysis of the MC data which coped simultaneously with finite size effects and the crossover from mean field to 3D Ising critical behavior.

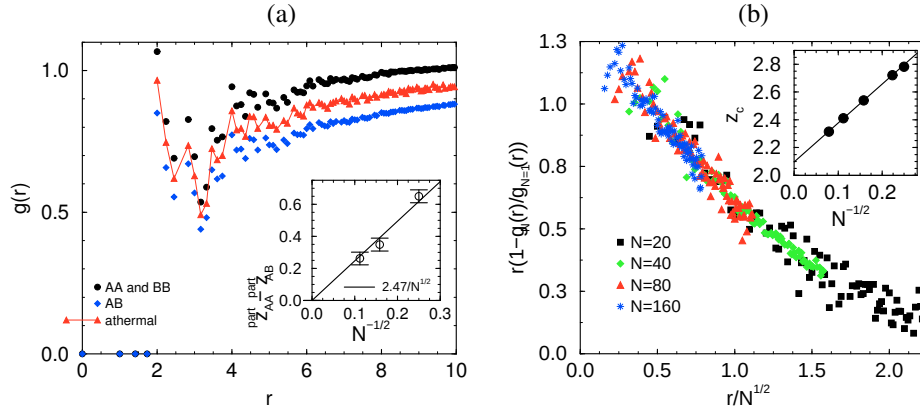


Figure 2.

(a) Intermolecular paircorrelation function for chain length $N = 80$ for the athermal system (triangles) and at criticality (circles and diamonds). The inset presents the scaling of the non-random mixing with increasing chain length. $z_{AA}^{\text{part}} = \langle \phi \rangle \int d^3 \mathbf{r} g_{AA}^{\text{inter}}(\mathbf{r})$. From Müller.⁵⁶ (b) Correlation hole of linear chains: Scaling behavior of the athermal intermolecular pair correlation function with chain length. The inset shows the chain length dependence of the effective coordination number. The line corresponds to $z_c = 2.1 + 2.8/\sqrt{N}$. From Müller and Binder.³²

In order to relate the measured critical temperatures to the structure of the polymeric fluid,³² we present in Fig.2(a) the intermolecular pair correlation function of strictly symmetric polymer blends in the athermal limit and at the critical point. In the athermal case, the distinction between the two species becomes irrelevant. The intermolecular pair correlation function mirrors two effects:³² (i) Due to the extended monomer structure the pair correlation function vanishes for distances $r < 2$. The presence of vacancies introduces local packing effects, which give rise to a highly structured function at short distances. One

can identify several neighbor shells, which are characteristic of the monomeric fluid. These packing effects are, of course, absent in simple lattice models where a monomeric unit occupies a single lattice site and are less pronounced in the bond fluctuation model than in continuum models. The length scale of these packing effects is set by the monomeric extension or the statistical segment length; the detailed shape depends strongly on the model and the degree of structure on local length scales. **(ii)** Furthermore, the extended structure of the macromolecules manifests itself in a reduction of contacts with *other* chains on intermediate length scales.²⁶ The length of this polymeric correlation hole is set by the size of the molecules, R_e , and its shape is characteristic for the large scale conformations of the molecule.

To a first approximation, we assume that the fluid structure is determined by the packing of the hard cubes on the lattice; neither the connectivity of the monomers along a polymer nor the thermal interactions influence the total pair correlation function. Under this assumption, we can separate the fluid-like packing effects on the monomer scale and the polymeric correlation hole effect and approximate the intermolecular pair correlation function by:³²

$$g^{\text{inter}}(r) = g_1(r) \left(1 - \frac{1}{\sqrt{N}} f(r/R_e) \right) \quad (14)$$

where g_1 denotes the pair correlation function of the monomer fluid and the function f parameterizes the structure of the molecule on the scale R_e . The prefactor is determined by requirement that the correlation hole contains one polymer:

$$\Phi \int d^3\mathbf{r} \left(\frac{g^{\text{inter}}(r)}{g_1(r)} - 1 \right) = 1 \quad (15)$$

Indeed, this factorization works excellently for flexible molecules in the bond fluctuation model. The ratio $g^{\text{inter}}(r)/g_1(r)$ is largely independent of packing effects and permits a distinction between monomeric packing effects and polymeric correlation hole effects in the simulations – though the length scales are not clearly separated for short chains. Not surprisingly, the correlation hole becomes deeper and wider as we increase the molecular weight. The scaling behavior of the correlation hole is shown in Fig.2**(b)** for linear athermal chains. It imparts a chain length dependence on the Flory-Huggins parameter or the effective coordination number, respectively. The effective coordination number is related to the short distance behavior of the intermolecular pair correlation function $z_c \sim \Phi N \sigma_e^3 g^{\text{inter}}(\sigma_e)$, where σ_e describes the range of the thermal interactions. More quantitatively, we calculate:

$$z_c = \Phi N \int_{r \leq \sqrt{6}} d^3\mathbf{r} g^{\text{inter}}(r) = z_c^\infty \left(1 + \frac{\text{const}}{\sqrt{N}} \right) \quad (16)$$

The scaling of the effective coordination numbers for flexible linear chains is presented in the inset of the panel **(b)**. The effective coordination number approaches its limiting scaling behavior with a $1/\sqrt{N}$ correction.

In Fig.2**(a)** we also present the intermolecular pair correlation functions g_{AA} and g_{AB} for chain length $N = 80$ close to criticality. In accord with intuition, AA contacts are more likely than AB ones and, hence, $g_{AA}^{\text{inter}} > g_{BB}^{\text{inter}}$. However, note that the sum of AA and

AB correlations can be well approximated by the intermolecular pair correlation function $g_{\text{atherm}}^{\text{inter}}(\mathbf{r})$ in the athermal limit:

$$\frac{g_{AA}^{\text{inter}}(\mathbf{r}) + g_{AB}^{\text{inter}}(\mathbf{r})}{2} \approx g_{\text{atherm}}^{\text{inter}}(\mathbf{r}) \quad (17)$$

This relation shows that the weak difference in interactions between the monomers $\chi \sim 1/N$ does not alter the structure of the monomer fluid. The energy of the system is mainly determined by composition fluctuations. The approximative decoupling between density fluctuations/packing effects and composition fluctuations in our model makes it possible to use the athermal value of the intermolecular pair correlation functions in Eq.(11). This identification corresponds to the high temperature approximation in the framework of the P-RISM theory.⁵⁵ In this sense, the miscibility behavior in this model can be described to a good approximation by a purely enthalpic χ parameter.

The inset of Fig.2(a) presents the integral of the correlation functions over the range of the square well potential, i.e., $z_{AA}^{\text{part}} = \langle \phi \rangle \int_{r \leq \sqrt{6}} d^3\mathbf{r} g_{AA}^{\text{inter}}(\mathbf{r})$ and similarly for z_{AB}^{part} . The MC simulations show that the difference between the AA and AB contacts decreases like $1/\sqrt{N}$, when χN is held constant. This exemplifies that the mean field approximation (or random mixing assumption) is justified in the limit $N \rightarrow \infty$.

The vanishing of composition fluctuations can be rationalized by estimating the scaling of non-random mixing effects with growing chain length. The strength of composition fluctuations in a volume V is of the order $1/\Phi V$. Expressing the composition fluctuations via the correlation functions, we obtain for a symmetric blend:

$$\begin{aligned} 1 \sim \Phi V (\langle \phi^2 \rangle - \langle \phi \rangle^2) &\sim \Phi \int d^3\mathbf{r} \left[\frac{g_{AA}^{\text{inter}}(\mathbf{r}) + g_{BB}^{\text{inter}}(\mathbf{r})}{2} - g_{AB}^{\text{inter}}(\mathbf{r}) \right] \\ &\sim \sqrt{N} \int d^3\mathbf{x} \left[\frac{g_{AA}^{\text{inter}}(\mathbf{x}) + g_{BB}^{\text{inter}}(\mathbf{x})}{2} - g_{AB}^{\text{inter}}(\mathbf{x}) \right] \quad \text{with} \quad \mathbf{x} = \mathbf{r}/R_e \end{aligned} \quad (18)$$

Therefore, we expect the difference in the AA and AB correlations to vanish like

$$g_{AA}^{\text{inter}}(\mathbf{r}) - g_{AB}^{\text{inter}}(\mathbf{r}) \sim 1/\sqrt{N} \quad (19)$$

upon increasing the chain length. These mean field arguments are in agreement with P-RISM calculations by Yethiraj and Schweizer.⁵⁵

The non-random mixing gives rise to a correction of the mean field critical temperature of the order $1/\sqrt{N}$. This behavior is in agreement with the Ginzburg criterium⁵⁷: The neglect of fluctuations is justified when the concentration fluctuations in one ‘‘correlation volume’’ of size ξ^3 are small compared to the composition difference between the two coexisting phases. Using the Flory-Huggins free energy, we obtain for the binodals in the vicinity of the critical point: $\phi = 1/2(1 \pm \sqrt{3(\chi - \chi_c)/\chi_c})$. The strength of composition fluctuations is determined by the second derivative of the free energy of mixing:

$$V\Phi(\langle \phi^2 \rangle - \langle \phi \rangle^2) = \frac{1}{\frac{1}{\phi} + \frac{1}{1-\phi} - 2\chi N} \quad (20)$$

and the correlation length scales like $\xi \sim R_e/\sqrt{\chi N - 2}$. Using the above expressions, one obtains:

$$\chi N - 2 \gg \text{Gi} \sim \frac{1}{N} \quad (21)$$

for the mean field theory to be accurate. The Ginzburg criterion⁵⁷ states that composition fluctuations are important at the critical point, where simulations¹⁴ and experiments^{54,16} find 3D Ising critical behavior. However, unlike the situation in mixtures of small molecules, the temperature interval in which composition fluctuations are dominant is restricted to the ultimate vicinity of the critical point. General arguments⁵⁸ rationalize that fluctuations lead to an overestimation of the “true” critical temperature T_c in the mean field theory T_c^{MF} by an amount $(T_c^{\text{MF}} - T_c)/T_c \sim 1/\sqrt{N}$, in accord with the non-random mixing behavior discussed above.

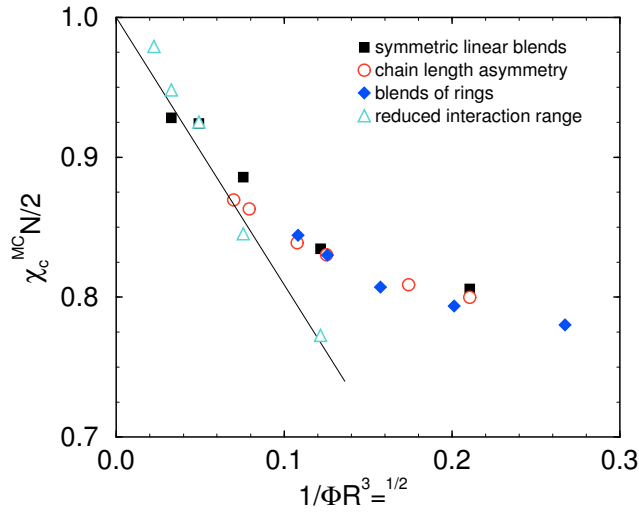


Figure 3. Ratio of the critical temperature (as determined in MC simulations) and the Flory-Huggins estimate for binary blends. Using the scaling variable $\sqrt{N}$ the MC results for blends of linear chains and ring polymers collapse onto a common curve. From Müller.⁵⁶

The linear scaling of the critical temperature at constant density has been observed for symmetric¹⁸ and asymmetric³² chains, and the behavior is also reproduced in off-lattice models.^{59,60} The overestimation of the critical temperature is investigated in more detail in Fig.3. Here, we plot the ratio between the MC results and our simple mean field estimate of the critical temperature. Upon increasing the chain length the difference between the MC result and the mean field estimate decreases. The figure includes the ratios between the critical temperature and the mean field estimate for mixtures of different chain lengths and blends of ring polymers. Using the scaling variable $1/\sqrt{N}$ we achieve a collapse of all data onto a common curve within the accuracy of the MC data which is of the order 1–5%. The figure also displays results for two choices of interactions ranges (for symmetric linear chains).¹⁸ Circles represent the results of a model, where the square well potential is extended over the first 54 lattice sites (as in the remainder of this paper), while triangles denote the results of a model, where the interaction comprises only the 6 nearest lattice sites. For large chain lengths the MC results are consistent with a linear dependence on the scaling variable. The collapse of the ratio T_c/T_c^{MF} with $\sqrt{N}$ for all blends marks the regime of chain lengths where the universal polymeric behavior dominates. This universal behavior is indicated as a straight line.

The bending towards a constant value of T_c/T_c^{MF} for small chains is due to the following effect: The deviations from the mean field behavior depend on the correlation length

ξ . The length scale is set by the amplitude of the square gradient term in the expansion of the free energy per chain with respect to long-ranged composition fluctuations. In a symmetric polymer mixture the prefactor is of the form $R_e^2/36\phi(1-\phi)$. Hence, the free energy cost of an inhomogeneous composition is due to the configurational entropy. However, for finite-ranged interactions there is an enthalpic contribution to the square gradient term with a prefactor of the order $\chi N \sigma_e^2$, where σ_e denotes the range of the monomeric interaction potentials. For long chain lengths the entropic contribution dominates, while the enthalpic term becomes important when the range of the thermal interactions becomes comparable to the radius of gyration of the chains. Consequently, the range of the interactions does not enter the Ginzburg criterion⁵⁷ to leading order. For very small polymers or rather long-ranged monomeric interactions, the interaction range σ_e might increase the correlation length and the shift between mean field and true critical temperature is smaller than estimated by $1/\sqrt{\mathcal{N}}$. When the range of interactions σ_e is decreased this effect should set in at smaller chain lengths. This is consistent with the simulation data: The data with the reduced interaction range show larger deviations at small chain lengths.

In the long chain length limit $\mathcal{N} \rightarrow \infty$, strictly symmetric blends in 3 spatial dimensions are very well describable by mean field theory, if the χ parameter is identified via the intermolecular pair correlation function of the athermal blend. The decoupling of composition and density fluctuations and the temperature independence of the structure of the underlying monomer fluid in the temperature range, where the phase separation in the binary blends occurs, give rise to a purely enthalpic χ parameter. For finite chain length, deviation between the MC results can be traced back to composition fluctuations. At the critical point, however, the mean field theory fails, and we observe 3D Ising critical behavior in accord with general arguments and deviations from mean field theory observed in simulation and experiment.

4.1.2 Films and Twodimensional Systems

Much of the success of the mean field description is related to the strong interdigitation of the polymers in the bulk. This quantity is measured by $\mathcal{N} = (\Phi R_e^d)^2$. It increases like the chain length N in $d = 3$ spatial dimensions. If the polymers are confined into quasi-twodimensional configurations the chain dimensions remain Gaussian ($R_e^2 \sim N$) in a dense melt. In this case, however, the quantity $\mathcal{N}$ is, to the leading order, independent from the chain length, i.e., a given molecule interacts only with a finite number of neighbors.

Let us consider a simple scaling argument for the behavior of the chain conformations upon confining a polymer into a thin film. According to Silverberg's argument⁶² the chain conformations can be conceived simply as random walks reflected at the surface. If a finite stiffness (or bending rigidity) along the chain is considered, parallel and perpendicular chain dimensions are no longer independent, but this short-ranged correlation along the chain is not expected to affect the properties on long length scales.

At some film thickness D , however, the description of the polymer conformations as mutually non-interacting Gaussian chains will fail. When the film thickness becomes very small, the chain folds back many times into its own volume and the density inside of the Gaussian coil increases. The fractal structure of the segments of a single chain gradually becomes compact (i.e., space-filling). When the density inside of the coil finally becomes comparable to the average density of the melt, the parallel chain extension begins to grow

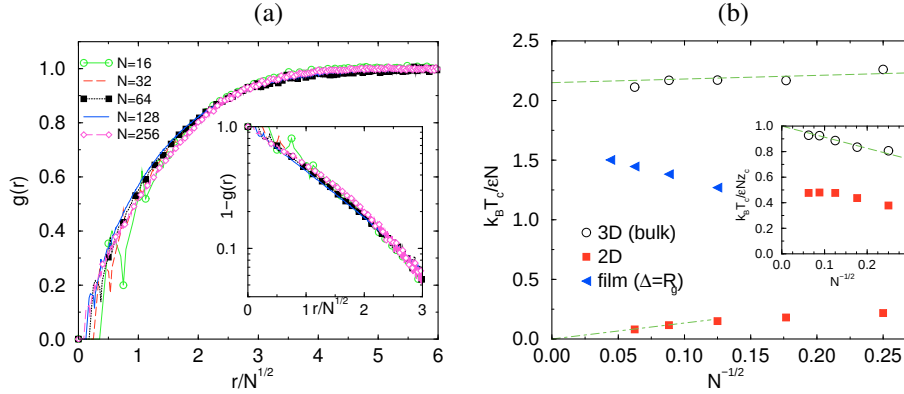


Figure 4. (a) Scaling plot of the intermolecular pair correlation function $g(r)$ versus $r/N^{1/2} \sim r/R$. Data are always taken close to criticality, but refer to all chains independent of their species. Chain lengths $N = 16, 32, 64, 128$, and 256 are included, as indicated in the key. For short chain lengths oscillations on the length scale of a few lattice units are visible and arise from packing effects. The inset shows $1 - g(r)$ on a logarithmic scale (cf. Eq.23) (b) Scaling of T_c/N and $T_c/N z_c$ (inset) vs. $N^{-1/2}$. For comparison the results for the 3d model are included. Adapted from Cavallo et al.⁶¹.

such that the density of the film remains laterally homogeneous. In this limit the chains adopt disk-like compact conformations of polymers in two dimensions. The stretching parallel to the surface is only negligible when $\rho D R_e^2 \ll N$ or

$$\frac{D}{R_e} \gg \frac{1}{\sqrt{N}} \quad (22)$$

where R_e denotes the unperturbed chain extension in the bulk. This reasoning suggests the following behavior of the average chain conformation in a thin film: (i) For $D \gg R_e$ the chain extensions parallel and perpendicular to the surface are unperturbed. Here, we expect also the intermolecular pair correlation function not to deviate strongly from the bulk behavior. (ii) For $R_e \gg D \gg R_e/\sqrt{N}$ the parallel chain extensions are unperturbed, but the chain folds back into the volume of its own coil. Other chains are gradually squeezed out of this volume, i.e., the correlation hole in the intermolecular pair correlation function deepens and the interdigitation of the chains decreases. In this regime, the number of intermolecular contacts z_c is reduced, but it is still a finite number. (iii) For $R_e/\sqrt{N} \gg D$ the chains do not overlap strongly and stretch parallel to the surface as to maintain a laterally uniform density. The lateral extension scales like: $R_{\parallel}^2 \sim R_e^3/(D\sqrt{N})$. Note that this effect occurs, when the film thickness is of the order of the excluded volume screening length $\xi_{ev} \sim R_e/\sqrt{N}$.⁶³ In this quasi-twodimensional limit, the correlation hole is describable by:

$$g^{\text{inter}}(r) = 1 - c \exp\left(-\frac{\text{const} \cdot r}{R_e}\right) \quad \text{for} \quad r > \xi \quad (23)$$

where $c \leq 1$ is a constant. The exponential term corresponds to the correlation function of a Gaussian walk in two dimensions and the functional form suggests that the correlation hole in the intermolecular pair correlation function in a dense melt is exactly canceled by the

density of the monomers of the reference chain, as it is the case in three dimensions. The MC data for the intermolecular paircorrelation function are presented in Fig. 4 (a). The data are compatible with the value $c = 1$. Therefore, the intermolecular paircorrelation function behaves at small distances like $1 - \exp(-\text{const} \cdot r/R_e) \approx \text{const} \cdot r/R_e$ and the number of intermolecular contacts per monomer is proportional to $\rho g^{\text{inter}}(b) \sim 1/\sqrt{N}$. This implies a scaling of the critical temperature like $T_c \sim Nz_c \sim \sqrt{N}$, as it is compatible with the simulation data shown in Fig. 4 (b). Indeed, T_c/N decreases like $N^{-1/2}$ as suggested by the scaling arguments. Unlike the situation in three spatial dimensions the mean field theory does not become quantitatively correct in the limit of long chain lengths. In two dimensions, even long chains interact only with a finite number of neighbors, and therefore $T_c/z_c N$ does not tend to unity for $N \rightarrow \infty$ (cf. inset of panel (b)).

Of course, these considerations highlight the influence of the confined chain conformations on the miscibility. If the surfaces prefer one component of the mixture, one encounters wetting transitions that might also lead to a pronounced alteration of the phase diagram in confined geometry.^{64–66}

4.2 Entropic Contribution to the χ Parameter: Non-Additive Packing

The simple form of the χ parameter is based on the decoupling of density fluctuations or packing effects from molecular architecture, composition of the blend, and temperature. To illustrate the consequences of chain architecture on the miscibility behavior, let us discuss a further example: A blend of polymers with indented monomer shapes. In Fig.5 (a) we sketch a symmetric mixture with indented monomer shapes.⁶⁷ As one can observe, monomers of different types are separated by a spatial distance of at least $\sqrt{3}$, whereas monomers of the same species can approach each other up to a distance 2. The properties of the pure phases are not altered, because the packing constraints only restrict the minimal distance between unlike species. Of course, this extremely simple monomer shape is no faithful representation of realistic monomer packing effects on a microscopic length scale. However, in the spirit that a monomer in the bond fluctuation model corresponds to a small number of chemical repeat units, we expect the model to capture some universal, long-wavelength properties on the coarse-grained length scale of a Kuhnian segment. The shape of the monomers leads to a non-additive packing between monomers of different species. In an A -rich phase an A polymer possesses more conformational freedom as in a B -rich environment. Hence the free energy of mixing acquires also an entropic contribution.

We cannot resort to the simple estimate of the χ parameter (11), but we have to go back a step and approximate the partition function. Let us consider the ratio of the canonical partition functions of an additive mixture Z_{add} and a blend with non-additive monomer shape Z_{nadd} :

$$\begin{aligned} \ln \frac{Z_{\text{nadd}}}{Z_{\text{add}}} &= \ln \frac{\sum \exp(-\beta E) \Pi_{i,j} (1 - \delta(r_{ij} - 2))}{\sum \exp(-\beta E)} \\ &= \ln \langle \Pi_{i,j} (1 - \delta(r_{ij} - 2)) \rangle_{\text{add}} \end{aligned} \quad (24)$$

where the sum comprises all configurations of the additive mixture. The index i (j) runs through all monomers of type A (B) and the factor $(1 - \delta(r_{ij} - 2))$ excludes all configurations violating the non-additivity constraint. Neglecting correlations among the monomer positions, one can factorize the average and get to a first approximation:

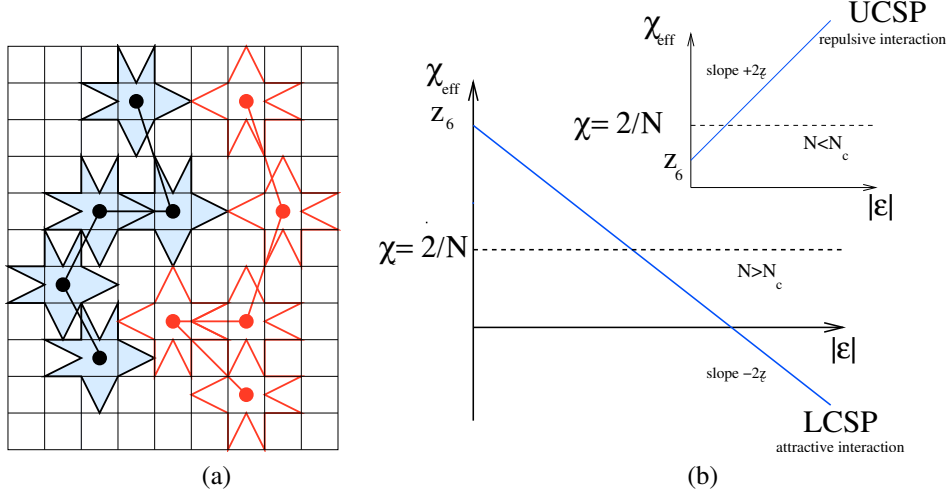


Figure 5. (a) Illustration of the intended monomer structure. A monomers are shaded, B monomers open. The effect of the indentation can be described by a non-additive hard core repulsion between monomers. As alluded, there is a strong entropic packing advantage which leads to phase separation. (b) Qualitative behavior of the χ -parameter for short chains (*UCSP*) and long chains (*LCSP*). From Müller.⁶⁷

$$\begin{aligned}
 \ln \frac{Z_{\text{nadd}}}{Z_{\text{add}}} &\approx \sum_{i,j} \ln \langle (1 - \delta(r_{ij} - 2)) \rangle_{\text{add}} \\
 &\approx - \sum_{i,j} \langle \delta(r_{ij} - 2) \rangle_{\text{add}} \\
 &= -n_A n_B \frac{z_6}{V \Phi N} = -V \Phi N \phi (1 - \phi) z_6
 \end{aligned} \tag{25}$$

where z_6 corresponds to the mean number of particles at a distance 2 in an additive mixture. This quantity is accessible in the simulation via the intermolecular paircorrelation function in an additive mixture:

$$z_6 = \Phi N \sum_{i=1}^6 g_{AB}(\mathbf{x}_i)_{\text{add}} \tag{26}$$

where the pair correlation function is normalized such that $g(r) \rightarrow 1$ for $r \rightarrow \infty$. Neglecting all local packing effects, one gets $z_6 = 0.2625$. Finally the packing induced contribution to the free energy and the effective χ parameter takes the form:

$$\frac{\Delta V}{T \Phi V} \approx N z_6 \phi (1 - \phi) \quad \chi_{\text{eff}} = \chi_0 + z_6 = \frac{2z_c}{T} + z_6 \tag{27}$$

where z_c is the effective coordination number of the thermal interaction, i.e. the mean number of monomers of *other* chains within the range of the square well potential.³² As anticipated the effective χ -parameter contains an enthalpic part and an entropic contribution. The excess entropy of mixing comprises two concurrent terms. The combinatorial

entropy stabilized the mixture whereas the packing contribution favors phase separation. Since the former is reduced by a factor $1/N$ the packing entropy dominates the behavior in the long chain limit.

The consequences for the miscibility behavior are discussed in Fig.5 (b): If the chain length is small enough, the translational entropy, which favors mixing dominates over the positive entropic contribution to the χ parameter. For $N < N_c \equiv 2/z_6$ the athermal blend is completely miscible. Upon lowering the temperature the blend phase separates at an upper critical solution temperature (UCST). For longer chains ($N > N_c$) the athermal blend is only partially miscible and to bring about a phase transition, we have to assume an attractive interaction between unlike species. Keeping with the notation of additive blends the attraction corresponds to negative values of ϵ . Upon increasing the absolute magnitude of the interactions $|\epsilon|$, the blend becomes miscible at a lower critical solution temperature (LCST):

$$\frac{1}{T_c} \sim -\epsilon_c = \frac{z_6}{2z_c} \left(1 - \frac{2}{z_6 N} \right) \quad (28)$$

The scaling of the lower critical solution temperature is in marked contrast to the scaling at the upper critical solution point with temperature. In the latter case the transition temperature is determined by a competition between the translational entropy of a polymer versus the monomeric repulsion; the critical temperature (UCST) increases linearly with chain length. In the former the conformational entropy per segment is balanced against the monomeric interactions and a chain length independent lower critical solution temperature (LCST) is approached from above. When expressed in terms of the χ parameter the Ginzburg criterion⁵⁷ and the critical amplitudes of the magnetization take the same form than for the upper critical solution points of the symmetric, additive mixture. However, when these quantities are written in terms of temperature the combination $(\chi_c - \chi)/\chi_c$ takes the form $Nz_c|T_c - T|/TT_c \approx 0.5Nz_6|T - T_c|/T$. This gives rise to an additional chain length dependence of the critical amplitudes and the Ginzburg number, when the temperature instead of the χ parameter is used in the vicinity of a lower critical solution point. For instance, the binodals in the temperature–composition plane open the wider the larger the chain length and the Ginzburg number which measures the relative distance between the critical temperature and the temperature where mean field behavior sets in becomes proportional to N^{-2} .

This mean field picture has been investigated via MC simulations⁶⁷ in the bond fluctuation model with the indented monomers sketched in Fig.5(a). Due to the additional excluded volume between unlike species the simulations have been performed at a reduced monomer density $\rho = 0.35/8$ in order to allow for a reasonable acceptance ratio of the semi-grandcanonical identity switches. Using a finite size scaling analysis we have determined the critical temperatures accurately. The dependence of the critical temperature on the chain length is summarized in Fig.6. For chain length $N = 10$ we find an upper critical solution temperature while for chain length $N = 16$ and larger we observe lower critical solution points. We find Ising critical behavior at the lower critical solution points and the temperatures approach a limiting value upon increasing the chain length. The critical temperatures are describable by a dependence of the form: $-\epsilon_c = 0.069(2) - 0.81(4)/N$. For chain length $N = 20$ we have measured $z_c = 1.41$ and $z_6 = 0.238$ in the MC simulations. This yields $\epsilon_c = 0.084 - 0.71/N$ as mean field estimate for the critical temperatures which

is also displayed in the Fig.6. It describes the simulation data only qualitatively. Deviations are partially due to the chain length dependence of the parameters z_c and z_6 via the correlation hole effect and composition fluctuations. Moreover, pronounced deviations stem from the crude approximation of the composition dependence of the the configurational entropy which lead to the simple expression(27). Our approximation treats the (rather strong) non additivity only perturbatively and neglects the composition dependence of the chain conformations (cf. Sec. 4.2.3).

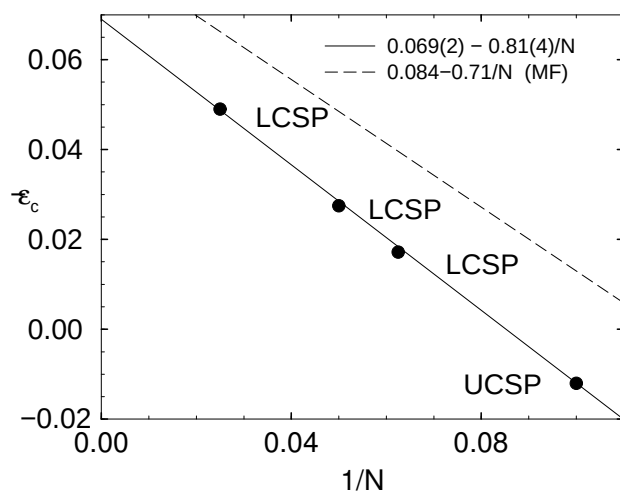


Figure 6. Chain length dependence of the critical temperature for mixtures of non-additive monomers. The straight line is given by $\epsilon_c = 0.069(2) - 0.81(4)\frac{1}{N}$. From Müller.⁶⁷

If monomer of the same species pack less efficiently than monomers of the same species, a negative entropic contribution to the χ parameter results. An entropic contribution to the χ parameter occurs not only for non-additive monomer shapes but also for large enough disparities in the segment size. Mixtures of small and large spheres demix⁶⁸ when the size difference between the species is large enough. We expect the consequences to be even more pronounced for polymers due to the small entropy on mixing. These effects have been explored in the framework of the Lattice Cluster Theory.¹³

4.2.1 Entropic Contribution to χ Parameter: Stiffness Asymmetry

Generally, the constituents of a blend are not symmetric, and asymmetry might have pronounced effects on the miscibility behavior. Since the properties of a blend deviate from the linear superposition of the individual properties of its components, the blend has new, possibly favorable characteristics.

A common asymmetry in polymer blends are differences in the statistical segment length.¹⁵ This effect has attracted much attention recently because of synthesis techniques for saturated hydrocarbon with a controlled degree of branching and their practical applications.⁶⁹ Bates¹⁵ suggested that the degree of branching can be represented on a coarse-grained scale by a difference in statistical segment lengths. Graessley and co-workers⁶⁹ have studied systematically the miscibility behavior of this class of polymers. Many -

though not all blends – were describable in terms of Hildebrands solubility⁷⁰ parameters. This suggests that the incompatibility is chiefly determined by enthalpic effects.

In the framework of the bond fluctuation model the effect of stiffness can be incorporated via an intramolecular bond angle potential of the form:^{71,67,25}

$$E(\theta) = f \cos(\theta) \quad (29)$$

where θ denotes the complementary angle to two successive bonds. Increasing the stiffness parameter f we energetically favor straight bond angles and increase the spatial extension of the molecule. The more open the molecule, the larger the number of intermolecular contacts. Upon increasing the stiffness parameter f from 0 (flexible chains) to 2 (semi-flexible), the chain extension for a polymer of $N = 32$ segments increases about a factor 1.5 and the number of intermolecular contacts z_c increases from 2.65 to 3.29 at $\epsilon = 0.05$ because the bond stiffness makes a folding back of the chain less probable.^{67,25} Unfortunately, the behavior of the intermolecular paircorrelation function cannot be decomposed into packing effects of the monomeric units and polymeric correlation hole: On the one hand the chain structure is not Gaussian on all length scales, the rod-like behavior on short distances becomes more important upon increasing stiffness. On the other hand, the bond angle potential between neighboring monomers influences the packing structure of the liquid, which for semi-flexible chains differs from the packing of the monomer fluid. The interplay between the packing arrangement of the monomers and the local conformations favored by the bond angle potential gives rise to an entropic contribution to the Flory-Huggins parameter.

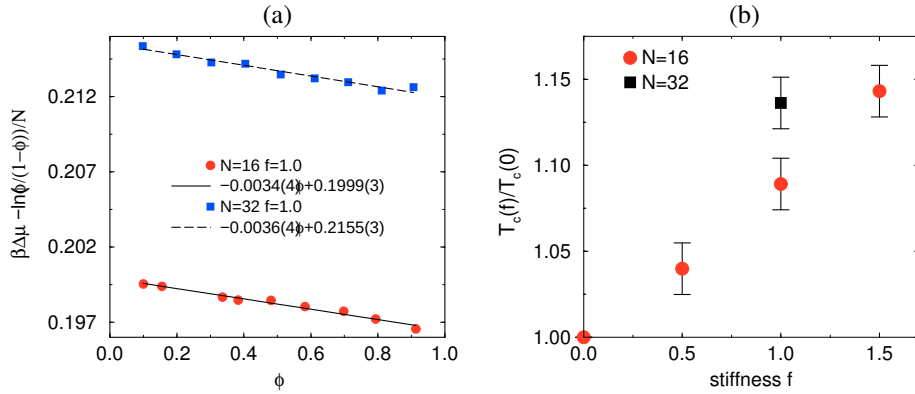


Figure 7. (a) Deviations from the semi-grandcanonical equation of state for blends of polymers with different stiffness. The (negative) slope is proportional to the entropic contribution to the χ parameter. (b) Dependence of the critical temperature on the stiffness f for chain length $N = 16$ and 32. From Müller.⁶⁷

To explore the possibility of entropic contributions to the χ parameter,⁶⁷ we accurately measure the dependence of the chemical exchange potential $\Delta\mu$ per polymer on the composition of the mixture. If there were no entropic contributions to the χ parameter, only the (exactly known) translational entropy would determine the relation between the

exchange chemical potential $\delta\mu$ and the composition ϕ . In Fig.7 (a) we present the deviations $\beta\Delta\mu - \ln[\phi/(1-\phi)]$ from the ideal mixing behavior. An entropic contribution to χ is related to a composition dependence in the form $-\chi N(2\phi - 1)$. Indeed, the MC results do reveal a composition dependence of this form and we accurately extract a small positive entropic contribution $\Delta\chi$ to the Flory-Huggins parameter.⁶⁷ $\Delta\chi = 0.0017(2)$ for $N = 16$ and $f = 1$, $\Delta\chi = 0.0018(2)$ for $N = 32$ and $f = 1$, and $\Delta\chi = 0.0031(3)$ for $N = 16$ and $f = 1.5$. For the parameters investigated, the entropic contribution $\Delta\chi$ increases with stiffness disparity in the blend and is roughly independent of chain length. For the chain lengths considered, it is only a few per cent of the critical value $\chi_c = 2/N$ and we anticipate only a small increase in the critical temperature. If the chain length independence of the entropic contributions remains true in the long chain length limit, the data suggest that for long macromolecules ($N \approx \mathcal{O}(1000)$) the stiffness disparity alone will be sufficient to cause phase separation or LCST behavior.⁶⁷ The positive entropic contribution to the χ parameter is consistent with field theoretical studies by Liu, Fredrickson and Bates,⁷² recent P-RISM calculations by Singh and Schweizer⁷³ and Lattice Cluster Theory by Foreman and Freed.⁷⁴

Unfortunately, the semi-grandcanonical MC moves rapidly become less efficient as the chain length or the stiffness disparity is increased, because the typical conformations of the two species differ strongly. For the chain lengths accessible in the simulations the shift in the critical temperature is small and not chiefly determined by the entropic contribution to the χ parameter. Most notably, the stiffness increases the size of the molecules and the number of intermolecular contacts. Hence, the enthalpic contribution to the χ parameter increases as well. The increase of the enthalpic and entropic contributions are both of the order of a few percent. We expect both effects to persist in the long chain length limit.

For the short chain length considered in the simulation the mean field theory overestimates the critical temperature by about 20%. According to the Ginzburg criterion the deviation between the critical temperature and the mean field estimate decreases with the chain extension (cf. Eq.(21)). We expect composition fluctuations to shift the critical temperature down the less the higher stiffness. Even if the χ parameter remained unaltered, the critical temperature in the MC simulations (for short chains) would increase.

The measured shift of the critical temperature is presented in Fig.7 (b). Upon increasing the stiffness or the chain length the ratio of critical temperatures between the strictly symmetric blend and the blend with stiffness disparity increases. For $N = 16$ and $f = 1$ we obtain a relative shift of 9% for T_c . The entropic contribution in the athermal system amounts to $2\Delta\chi/N = 0.014$ while the relative increase of the effective coordination number is 6% as measured in the MC simulations. The remaining deviation is consistent with the dependence of the ratio T_c^{MF}/T_c on chain extension.

A similar study on the consequences of stiffness disparity was pursued in an off-lattice model by Weinhold et al.⁷⁵ Using the increment chemical potential method³⁷ the authors explored the miscibility behavior. Upon blending the chemical potential of the flexible chains increases, while the stiffer component lowers its free energy. This effect was rationalized via equation of state effects: At constant density the blend has a higher osmotic pressure than the pure flexible component and a lower osmotic pressure than the stiff component. The authors stated that the behavior is in almost quantitative agreement⁷⁵ to the simulations in the bond fluctuation model. Again this indicates that the lattice structure in the bond fluctuation model is a good approximation for the continuum space proper-

ties. The net excess free energy per monomer obtained from the simulations is essentially zero to within the $\pm 0.005T$ statistical error.⁷⁵ This is consistent with the small values $\Delta\chi$ found in our simulations. They also illustrated the significance of the intermolecular pair correlation function for the χ parameter.⁷⁶

Structural asymmetries can lead also to a composition dependence of the Flory-Huggins parameter χ . Dudowicz et al⁷⁷ studied a dependence of the form: $\chi = a + (b + c\phi)/T$. Depending on the coefficients a, b , and c different phase behaviors can be observed, and the coefficients can be approximatively related to the structure of the monomeric units via the Lattice Cluster Theory.

4.2.2 Effects of Pressure or Solvent Density

The effect of pressure on the miscibility behavior has attracted much interest. There is an interesting interplay between equation of state effects and the phase behavior. This has also a practical importance when a blend is mixed in an extruder or during injection molding of plastics.⁷⁸ In the framework of the bond fluctuation model the local fluid structure is mainly determined by athermal packing effects. In the athermal melt, the osmotic pressure is independent of the chain length.⁷⁹ This chain length independence of the pressure at high densities is a universal property of polymer melts.²⁶ The excess free energy change upon mixing per monomer is of the order $\chi \sim 1/N$. The free energy cost for a density fluctuation (on the length scale of a monomer) is proportional to the compressibility and, hence, chain length independent. Therefore, in the one phase region, the interactions lead only to a small excess volume change upon mixing for high molecular weight (compatible) polymer blends. These kind of compressibility effects have been observed, e.g., at polymer-polymer interfaces, where the energetic unfavorable interactions at the interface result in a slight decrease of the density. However, even for rather strongly segregated blends⁴¹ ($\chi N < 20$) in the bond fluctuation model the effect yields only a density reduction by a few percent. This is also in agreement with the decoupling of composition and density fluctuations. The insignificance of compressibility effects in weakly interacting blends and the consequences for the analysis of neutron scattering data has been explored in ref.^{80,81} Gromov and de Pablo⁸² found in MC simulations of a model of Lennard-Jones chains of length $N = 16$ a volume change of approximately 10% at constant pressure.

For our specific choice of interactions the total energy density per monomer is also of the order χ . Hence, the fluid structure corresponds to that of an athermal melt in the limit of long chain lengths and $\chi N = \text{const}$. This observation is in accord with the temperature independence of the packing and the effective coordination number in the temperature range where the phase separation occurs. If we were to simulate at constant pressure, the density around the phase transition would be chiefly determined by the value of the athermal system in the limit of long chains. Therefore, we expect not much change in the chain length dependence of the miscibility at constant volume or at constant pressure.

However, we would like to emphasize that the approximate decoupling between the fluid structure/density and the temperature (at constant pressure) is not a universal property and does depend on the specific choice of the interactions. Unlike the excess free energy change upon mixing, the total energy density per monomer needs not to be small for long polymers at $\chi N = \text{const}$. In many experimental instances, concentrated polymer solutions and melts exhibit a temperature dependent equation of state. Therefore the

density Φ is a function of temperature T at a given fixed pressure. If the Flory-Huggins parameter χ is mainly enthalpic, we still can use Eq.(11) to calculate the χ parameter. In a very crude approximation the intermolecular pair correlation function is independent of the density. Then the effective coordination number is proportional to the monomer density $\rho = \Phi N$. The critical temperature of the blend at constant pressure scales like $T_c \sim \rho(T_c)N$. For a blend with an UCSP the critical temperature increases with chain length and the monomer density $\rho(T_c)$ at the critical temperature decreases. This leads to a weaker dependence of the critical temperature with chain length. Such an effect has been observed in various experiments. Escobedo and de Pablo⁶⁰ have investigated the scaling of the critical temperature in a symmetric blend under constant pressure. They found that the critical temperature increases effectively like $\sqrt{N}$ for the range of chain length investigated. Moreover, recent experiments on polyolefin blends by Lohse and co-workers find evidence for a temperature-pressure superposition:⁷⁸ Far from the UCST the interaction energies depend on the pressure P only via $\rho(P)$. However, deviations from this scaling are found for blends which demix upon heating. This might indicate an additional dependence of the local packing arrangements with the density.

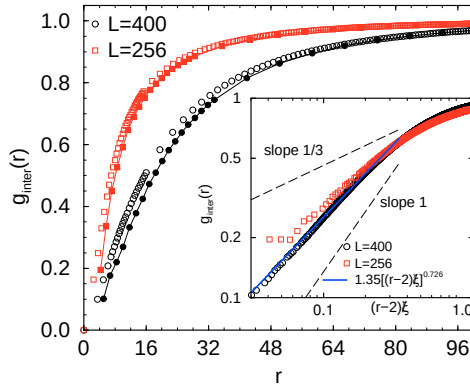


Figure 8. Real space dependence of the intermolecular correlations: (a) Intermolecular pair correlation function $g^{\text{inter}}(r)$ for chain length $N = 2048$ and monomer number densities $\rho = 0.0032$ ($L=400$) and $\rho = 0.0122$ ($L=256$). Open symbols correspond to the MC results, whereas filled one represent the RG calculations. Inset: Power-law dependence of $g^{\text{inter}}(r)$ at small distance. The data are compatible with a dependence of the form $g^{\text{inter}}(r) \sim r^\alpha$ with an exponent $0.7(3)$, close to the predicted value $\alpha = 0.8$. From Müller et al.⁸³

If the density becomes lower one obtains a semi-dilute solution instead of a melt: In a 3-dimensional melt the chain statistics is Gaussian up to microscopic length scales and the correlation hole in the intermolecular pair correlation function has only a finite depth. In a semi-dilute solution, the chain statistics is Gaussian for distances larger than the excluded volume screening length ξ_{ev} , but self-avoiding walk-like for smaller distances, i.e., inside the excluded volume blob. The size of the excluded volume blob ξ_{ev} can be determined by requiring that the monomer density inside the blob is mainly created by monomers of the same chain. Let g denote the number of monomeric units inside the blob, then $\rho = g/b_{\text{ev}}^3 g^{3\nu_{\text{ev}}}$, where $\nu_{\text{ev}} = 0.588$ characterizes the chain extension of a self-avoiding walk $R \sim b_{\text{ev}} N^{\nu_{\text{ev}}}$. Therefore the size of the blob decreases with density like $\xi \sim b_{\text{ev}} (\rho b_{\text{ev}}^3)^{-\nu_{\text{ev}}/(3\nu_{\text{ev}}-1)}$. Thus, the number of monomers of other chains inside the excluded volume blob is small and chains do not interdigitate on the length scale ξ_{ev} or smaller. The intermolecular pair correlation function exhibits a deep correlation hole and shows a power-law behavior $g^{\text{inter}}(r) \sim (r/\xi_{\text{ev}})^\alpha$ for $r < \xi_{\text{ev}}$ (cf. Fig. 8). The exponent adopts the value $\alpha = 3 - 2/\nu_{\text{ev}} - \omega_{12} \approx 0.8$,^{83,84} where $\omega_{12} \approx 0.4$ is the

correction to scaling exponent that characterizes the contacts ($\sim N^{\nu\omega_{12}}$) of two mutually interdigitating self-avoiding walks. The number of intermolecular contacts scales like $\chi \sim z_c \sim \rho g^{\text{inter}}(\sigma_e) \sim (\rho b_{\text{ev}}^3)^{(1+\nu\omega_{12})/(3\nu_{\text{ev}}-1)} \sim \rho^{0.616}$. This effect – the Joanny renormalization of the Flory-Huggins parameter – has been worked out in detail by Leibler⁸⁵ and Schäfer.⁸⁶

If the two components differ strongly in their equation of state, more complex phase diagrams occur (cf. Konynenburg and Scott⁵ for a classification of possible types of miscibility behavior in compressible blends). In addition to liquid-liquid phase separation one encounters liquid-vapor phase coexistence and three phase coexistence regions where two liquids and a vapor phase coexist.

4.2.3 Single Chain Conformations.

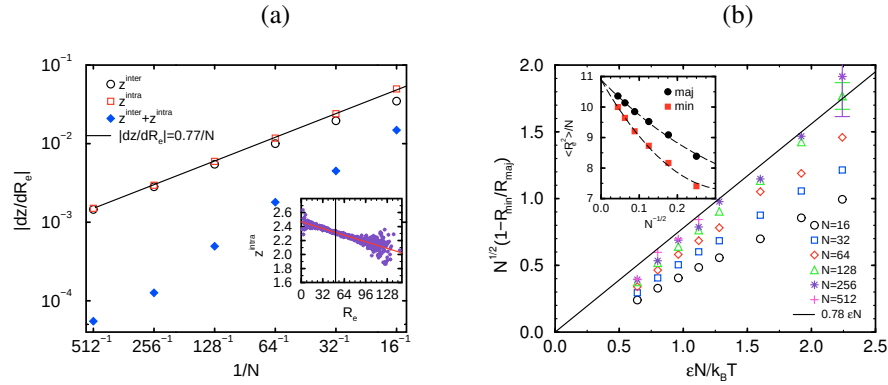


Figure 9. (a) Correlation between the chain extension and the number of intermolecular and intramolecular contacts. The straight line marks the prediction of the scaling considerations $dz/dR \sim 1/N$. (b) Shrinking of the end-to-end extension below the critical temperature. The solid line marks the expected behavior for long chain lengths. The inset in (a) shows the chain length dependence of end-to-end distance of the majority and minority component at constant $\epsilon N = 0.64$. Lines in the inset are only guides to the eye. From Müller.²²

Sophisticated theoretical approaches have been developed to study the dependence of chain conformations on the environment: self-consistent P-RISM theory^{87,88}, SCF theory for clusters of chains²², and ellipsoid models.^{20,89} In the following we constrain ourselves to MC simulations and simple scaling arguments.

MC simulations suggest that one possible mechanism of conformational changes in blends is associated with exchanging energetically unfavorable intermolecular contacts for attractive intramolecular contacts upon reducing the spatial extension of the molecule. Attributing this shrinking of the minority component to a balance between the entropy loss due to deviations from the unperturbed conformations and energy gain upon shrinking, we can estimate the magnitude of conformational changes:²² Within the Gaussian chain model a deviation from the unperturbed chain extension R_e gives rise to an entropic force of the form

$$\frac{dS}{dR} \sim \frac{(R - R_e)}{R_e^2} \quad (30)$$

This is opposed to an enthalpic force dE/dR , where E denotes the single chain energy. E comprises energetically favorable interactions Nz^{intra} among monomers of the same chain and Nz^{inter} interactions with monomers of other polymers. The exchange of an intermolecular interaction with an intramolecular one lowers the single chain energy by an amount of the order χ . The number of intramolecular interactions per monomer z^{intra} is given by the density of monomers of the same chain inside of its volume $z^{\text{intra}} \sim 1/\Phi R_e^3 \sim 1/\sqrt{N}$. Under the assumption that the reduction of the chain extension does not affect the total number of interactions, but merely exchanges intermolecular interactions into energetically favorable intramolecular ones, we estimate the chain length dependence of the energy change as

$$\frac{dE}{dR} \sim -\frac{\chi N}{\Phi R_e^4}. \quad (31)$$

Balancing the entropic force against the enthalpic one, we obtain:

$$\frac{R_e - R}{R_e} \sim \frac{\chi N}{\Phi R_e^3} = \frac{\chi N}{\sqrt{N}}. \quad (32)$$

These scaling arguments are similar to those for a chain in a marginal solvent and suggest that the perturbation of the chain conformations decreases upon increasing the chain length at $\chi N = \text{const.}$ The conformations in high molecular weight blends are only very mildly perturbed in the minority phase.

The derivation of the scaling arguments relies on the number of intramolecular contacts and its dependence of the chain extension. Obviously, this estimate excludes contributions from the neighbors along the polymer, which give rise to $z^{\text{intra}} \sim N$. Though the neighbor contribution is important for the scaling of the number of intramolecular contacts with chain length, we assume them to be independent from the instantaneous shape/extension of the polymer.^a

Using the scaling estimate above, the dependence of the z^{intra} on the instantaneous extension R at fixed chain length is given by: $dz^{\text{intra}}/dR \sim 1/\Phi R_e^4 \sim 1/N$. Clearly, a detailed verification of this scaling behavior by MC simulations is warranted. Such a test is presented in Fig.9 (a). The inset presents the average number of intramolecular contacts at fixed end-to-end distance for an athermal melt of chain length $N = 256$. At the mean end-to-end distance $\sqrt{\langle R^2 \rangle_0}$ we determine the slope dz/dR as indicated by linear regression. The chain length dependence of the derivatives of the number of inter- and intramolecular contacts with respect to the chain extension decreases like $1/N$ for large chain lengths. This confirms the scaling predictions. Moreover, the sum of intermolecular and intramolecular contacts depends much weaker on the spatial extension R_e and the dependence decreases faster than $1/N$. This indicates that the fluid structure of the monomers is mainly determined by packing and approximately decouples from the chain conformations. For long chains, the conformational changes merely result in an exchange of inter- and intramolecular contacts. From the MC data we estimate $dz/dR = 0.77(7)/N$ for the bond fluctuation model at density $\rho \equiv \Phi N = 1/16$.

The scaling predictions can be made more quantitatively in the framework of the Gaussian chain model. Let $P(\mathbf{R})$ denote the probability distribution of the end-to-end vector $\mathbf{R}$ which incorporates the dependence of the single chain energy E on the chain extension.

^aThe fact that the number of intermolecular contacts per chain Nz^{intra} scales with N , but its derivative with

$$E(\mathbf{R}) \approx E(\sqrt{\langle R_e^2 \rangle_0}) + \frac{dE}{dR} \left[|\mathbf{R}| - \sqrt{\langle R_e^2 \rangle_0} \right] \quad (33)$$

where $\langle R_e^2 \rangle_0 = b^2 N$ denotes the end-to-end distance in the athermal limit. Assuming Gaussian statistics for the unperturbed chain, we can write the probability distribution for a B polymer in the form

$$P(\mathbf{R}) \sim \exp \left(-\frac{3\mathbf{R}^2}{2\langle R_e^2 \rangle_0} - \frac{1}{T} \frac{dE}{dR} \left[|\mathbf{R}| - \sqrt{\langle R_e^2 \rangle_0} \right] \right). \quad (34)$$

The total energy change associated with the transfer of two intermolecular contacts of a B polymer $[\epsilon(2\langle\phi\rangle - 1)]$ into an intermolecular contact $[-\epsilon]$ and a contact between monomers not belonging to this B polymer $[-\epsilon(2\langle\phi\rangle - 1)^2]$ amounts to $\Delta E = 4\epsilon\langle\phi\rangle^2$. The number of intramolecular contacts per monomer increases by 2, and the number of intermolecular contacts decreases by the same amount. Therefore dE/dR equals $2\epsilon\langle\phi\rangle^2 dz/dR$. Using this estimate and assuming that the conformational changes are small, we calculate the mean square end-to-end distance:

$$\langle R_e^2 \rangle \approx \langle R_e^2 \rangle_0 \left(1 - \sqrt{\frac{8}{27\pi}} \frac{1}{\sqrt{N}} \left[\frac{\Phi R_e^4 dz^{\text{intra}}}{z_c dR} \right] \chi N \langle\phi\rangle^2 \right) \quad (35)$$

We expect this asymptotic expression to hold only for very small values of $\chi N/\sqrt{N}$, where the conformational changes can be treated perturbatively. This expression predicts a quadratic dependence of the chain extension on the composition of the mixture. The effect

respect to the chain dimension R is independent of the chain length can be rationalized in the framework of the Gaussian chain model. Let $R_0^2 = b^2 N$ denote the unperturbed end-to-end distance. For convenience, we replace the square well contacts by Gaussian contacts of widths σ_e . Then, the number of intramolecular contacts can be written in the form:

$$N z^{\text{intra}}(R) \sim \frac{1}{2} \int ds \, ds' \, d^3\xi \left(\frac{3}{2\pi\sigma_e^2} \right)^{3/2} \left(\frac{3}{2\pi b^2 |s - s'|} \right)^{3/2} \left(\frac{3}{2\pi b^2 (N - |s - s'|)} \right)^{3/2} \left(\frac{3}{2\pi b^2 N} \right)^{-3/2} \\ \exp \left(-\frac{3\xi^2}{2\sigma_e^2} \right) \exp \left(-\frac{3\xi^2}{2b^2 |s - s'|} \right) \exp \left(-\frac{3(R - \xi)^2}{2b^2 (N - |s - s'|)} \right) \exp \left(+\frac{3R^2}{2b^2 N} \right)$$

In the limit $\sigma_e \ll b^2 N$ this expression can be simplified to:

$$N z^{\text{intra}}(R) \sim \int_0^N d\Delta s \, (N - \Delta s) \left(\frac{3}{2\pi(\sigma_e^2 + b^2 \Delta s)} \right)^{3/2} \left(\frac{N}{N - \Delta s} \right)^{3/2} \\ \exp \left(-\frac{3R^2}{2b^2 N} - \frac{\Delta s}{N - \Delta s} \right)$$

The major contribution to this integral comes from $\Delta s \approx 0$, i.e., the neighbors along the chain dominate the number of intramolecular contacts. $N z^{\text{intra}}$ scales like N . The derivative with respect to R is given by:

$$\frac{dN z^{\text{intra}}}{dR} \sim \int_0^N d\Delta s \, (N - \Delta s) \left(\frac{3}{2\pi(\sigma_e^2 + b^2 \Delta s)} \right)^{3/2} \left(\frac{N}{N - \Delta s} \right)^{3/2} \\ \frac{3R}{b^2 N} - \frac{\Delta s}{N - \Delta s} \exp \left(-\frac{3R^2}{2b^2 N} - \frac{\Delta s}{N - \Delta s} \right)$$

The integral converges in the limit $\sigma_e \rightarrow 0$ and scales like N^0 at $R = R_0$.

increases linearly with the χ parameter, but decreases at fixed χN like $1/\sqrt{N}$. Hence, the conformations of long macromolecules are only very weakly dependent on composition. This is also in qualitative agreement with field theoretical calculations of Vilgis and Holyst,⁹⁰ and Garas and Kosmas.⁹¹ A decreasing dependence of the single chain conformations on the environment is also found in simulations and self-consistent P-RISM calculations for atomistic models²¹.

In Fig.9(b) we explore the scaling of the shrinking of the chains in the minority phase at two phase coexistence. The simple estimate (35) predicts that the chains in the majority phase are unperturbed, while the minority component reduces its size in a well segregated blend. The relative shrinking $1 - R_{\min}/R_{\text{maj}}$ increases linearly with ϵN and is at fixed ϵN of the order $1/\sqrt{N}$. This is in qualitative agreement with the simulation data for the end-to-end distance presented in Fig.9(b). The straight line represents the prediction of Eq.(35). Surprisingly, the simulation data approach the asymptotic behavior very slowly. Only for chain length $N \geq 128$ the simulations reach the scaling limit; for smaller chain length the estimate (35) overpredicts the shrinking.

There are at least two possible reasons for the pronounced small chain length corrections to the scaling behavior: First, the considerations hold only in the regime $2 \ll \chi N \ll \sqrt{N}$. The first limit is set by the condition that the blend is well segregated, i.e., $\langle \phi \rangle_{\min} \ll 1$. The second requirement corresponds to small conformational changes. This temperature regime is experimentally relevant because the concentration of the minority component is small but does not vanish for long chain lengths. However, for short chain lengths these conditions are rather restrictive. For very strong segregation the linear decrease of the chain dimensions with $\chi N/\sqrt{N}$ will certainly break down and the changes cannot be treated perturbatively.

A second source of corrections to asymptotic scaling behavior might be deviations from the Gaussian chain statistics upon shrinking. The ratio between the end-to-end distance and the radius of a completely collapsed coil $(3/4\pi b^3 \rho \sqrt{N})^{1/3} = (3/4\pi \sqrt{N})^{1/3}$ decreases only very weakly with chain length. Even for the chain length $N = 256$ the end-to-end distance exceeds the radius of the densely packed coil only by a factor of 5. If the extension of the shrunken chain becomes comparable to the size of the completely collapsed coil, it cannot reduce its size much further. In this case the data would not scale as a function of $\chi N/\sqrt{N}$, but they would crossover to a temperature independent end-to-end distance the earlier the smaller the chain length.

The conformational changes result in a composition dependence of the segmental entropies and, hence, give rise to a composition dependence of the χ parameter. The scaling arguments suggest that the conformational changes alone might only produce an effect of the order $1/N$ per chain. Using the relation between the exchange chemical potential and the composition we can measure a χ parameter. In the one phase region far above the critical temperature we observe an upward parabolic dependence of the χ parameter on the composition. The composition dependence is about 4%.

4.3 Interfaces

A symmetric binary blend will partition into two coexisting phases already at high temperatures. Domains of the two phases will form and they will be separated by interfaces. Comparing the properties of interfaces as obtained from simulations or experiments to the

predictions of mean field theory one has to distinguish two types of properties: Excess quantities which do not make any explicit reference to the interface position - like e.g., the interface tension or the surface excess of a component - can be directly compared. Profiles across the interfaces or observables that depend on the local position of the interfaces are affected by capillary waves.

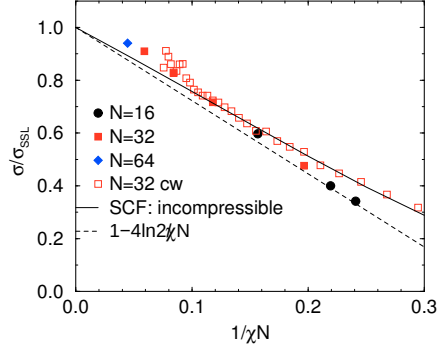


Figure 10. Interface tension in units of $\sigma_{SSL}R_e^2/T = \sqrt{N}\sqrt{\chi N}/6$. The symbols correspond to MC simulations for three chain length as indicated in the key. The interfacial tension has been extracted from the probability distribution of the composition in the semi-grandcanonical ensemble, and the spectrum of capillary waves (for $N = 32$). Data are compared to SCF results and the exact asymptotic correction in the limit $\chi N \rightarrow \infty$ by Semenov.⁹² Adapted from Schmid and Müller.⁹³

In Fig.10 we compare the interface tension extracted from the simulations to the prediction of the SCF theory using the identification of the Flory-Huggins parameter χ as in the previous section. The simulation data confirm, that σ/σ_{SSL} is only a function of χN and they agree quantitatively with the SCF theory.

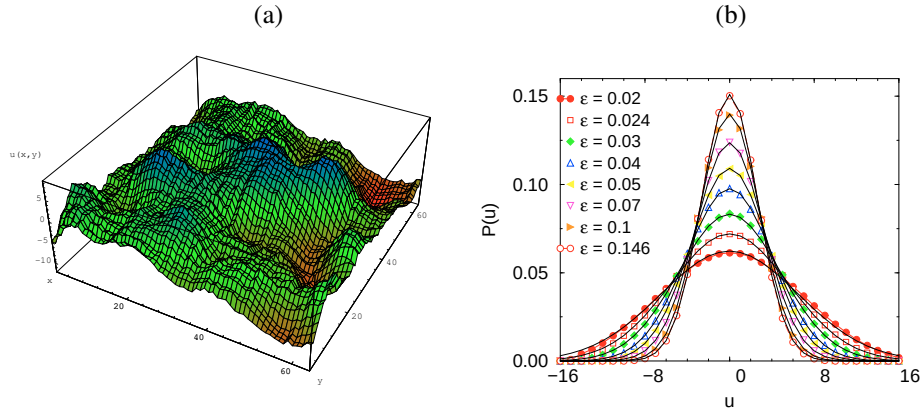


Figure 11. (a) Fluctuations of the local interface position in a binary polymer blend at $\epsilon = 0.03$ ($\chi N \approx 5.1$). The position has been averaged over a lateral size of $B = 8$. (b) Distribution of the local interface positions as a function of the incompatibility ϵ . From Werner et al.⁹⁴

A comparison of profiles between simulations/experiments and SCF calculations has to take due account of the fluctuations of the local interface position. Interfaces are not perfectly flat, as assumed in the SCF calculations, but there are thermal fluctuations. A snapshot of the interface position in the MC simulations of a binary blend is shown in

Fig.11(a). To a first approximation the effect of these fluctuations is to increase the effective area of the interface. Let $u(\mathbf{r}_{\parallel})$ denote the local interface position. Then, the free energy cost of deviations from perfectly flat configurations are described by the Hamiltonian $\mathcal{H}$:

$$\mathcal{H}[u(\mathbf{r}_{\parallel})] = \int d^2\mathbf{r}_{\parallel} \left\{ \frac{\sigma}{2} |\nabla u|^2 + \frac{\kappa}{2} |\Delta u|^2 \right\} \quad (36)$$

This is an expansion in terms of small u and its derivatives, σ denotes the interfacial tension and κ the bending rigidity. For interfaces between not too long homopolymers κ is very small; for copolymer laden interfaces, however, the second term becomes important as we shall discuss. This capillary wave Hamiltonian is diagonal and quadratic in terms of the Fourier components $u(q)$ and the equipartition theorem yields for the spectrum of fluctuations in thermal equilibrium:

$$\langle u^2(q) \rangle = \frac{T}{\sigma q^2 + \kappa q^4} \quad (37)$$

The local positions $u(\mathbf{r}_{\parallel})$ are also Gaussian distributed $P(u)$ with variance s :

$$s^2 = \frac{1}{4\pi^2} \int d^2\mathbf{q}_{\parallel} \langle u^2(\mathbf{q}_{\parallel}) \rangle = \frac{T}{2\pi\sigma} \ln \left(\frac{q_{\max}}{q_{\min}} \right) \quad (38)$$

where a short and long length scale cut-off $q_{\max}$ and $q_{\min}$ have to be introduced to avoid the divergence at $q \rightarrow 0$ and $q \rightarrow \infty$. The bending rigidity κ has been neglected, it would make the cut-off at small distance obsolete. The MC result for the distribution $P(u)$ of the local positions is presented in Fig.11(b). Upon decreasing the incompatibility ϵ , we increase the strength of the fluctuations.

These capillary waves broaden the apparent profiles p_{app} . Laterally averaged profiles, as obtained in experiments or MC simulations, are describable via the convolution of an intrinsic profiles p_{int} of an ideally flat interfaces and the distribution of the local positions

$$p_{\text{app}}(z) = \int du P(u) p_{\text{int}}(z - u), \quad (39)$$

where z denotes the coordinate perpendicular. When applied to a erfc-shape profile, one obtains:^{95,94,96}

$$\frac{w_{\text{app}}^2}{R_e^2} = \frac{w_{\text{int}}^2}{R_e^2} + \frac{T}{4\sigma R_e^2} \ln \left(\frac{q_{\max}}{q_{\min}} \right) \quad (40)$$

$$\stackrel{\chi N \gg 2}{=} \frac{1}{\sqrt{6\chi N}} \left(1 + \frac{3}{2\sqrt{N}} \ln \frac{q_{\max}}{q_{\min}} \right) \quad (41)$$

where we have used the expression for the strong segregation limit in the last line. The apparent width is broader than the intrinsic one and depends via the two cut-offs on the system geometry.^{95,94} For a free interface the lower cut-off $q_{\min}$ is set by the lateral block size B on which the interface is observed. This might be set by the size of the simulation cell or the coherence length of the neutron beam by which the interfacial structure is investigated. Gravitation or interactions with walls/surfaces also limit long-wavelength

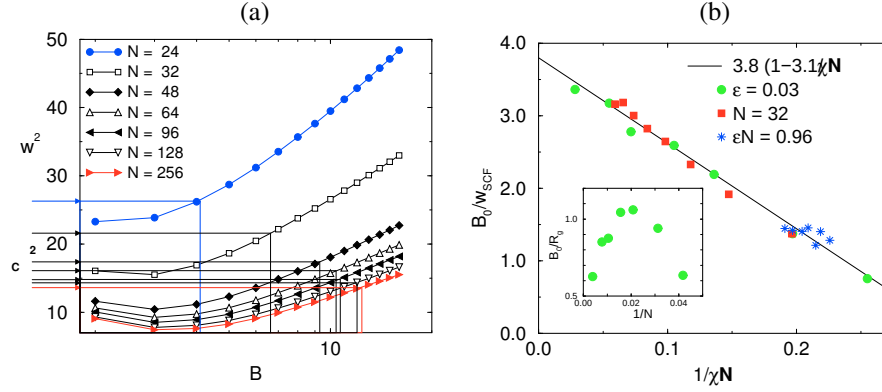


Figure 12. (a) Dependence of the apparent interfacial width on the lateral block size B at constant incompatibility $\epsilon = 0.03$ for different chain lengths. From the intersections of the simulation data with the SCF predictions (presented as horizontal lines) we determine the small length scale cut-off B_0 . (b) The ratio B_0/w_{scf} approaches a constant value 3.8. Leading corrections are of the order $1/\chi N$. Note that data sets at constant N , χN and χ collapse in this representation. The inset shows the ratio B_0/R_e for fixed incompatibility $\epsilon = 0.03$. The ratio tends to zero for large N . From Werner et al.⁹⁴

fluctuations. The upper cut-off q_{max} describes the crossover from capillary waves (on large distances) to “intrinsic” fluctuations, which build up the smooth profile of the ideally flat interface. Measuring the width of an interface in the experiments or simulations one cannot extract both, the “intrinsic” width w_{int} and the upper cut-off q_{max} individually.

Polymer blends are well suited to examine this crossover, because the strong interdigitation of the molecules makes SCF calculations describe the properties of interfaces accurately except for capillary wave fluctuations. Taking the SCF prediction as the “intrinsic” width of an hypothetical flat interface, we use Eq. (40) to define the length scale $B_0 = 2\pi/q_{max}$ at which the crossover between “intrinsic” fluctuations and capillary waves occurs. This procedure is illustrated in Fig.12(a). There are three possible candidates for B_0 : A microscopic length scale (e.g., the bond length) independent from temperature or chain length; the width of the interface, which depends on temperature but not on N ; or the radius of gyration/correlation length which depends both on χN and R_e . The simulation data⁹⁴ in panel (b) indicate a behavior of the form $B_0 = 3.8w_{scf}(1 - 3.1/\chi N)$, i.e. the intrinsic width of the interface sets the crossover length; a result compatible with calculations of Semenov.⁹⁷

The spectrum of interfacial fluctuations is an alternative route for measuring the interfacial tension in MC simulations. This is illustrated for a blend with a structural asymmetry in Fig.13. When we increase the stiffness disparity, the semi-grandcanonical identity switches become increasingly inefficient and σ cannot be obtained via pre-weighting techniques. Stiffness increases the incompatibility and this, in turn, results in a larger interfacial tension (c.f. Fig.13(a)). This effect is quantitatively captured by the SCF calculations which account for the detailed chain architecture. Qualitative agreement is also obtained within the Gaussian chain model. The deviations from the prediction of Helfand and Sapse⁹⁸ are mainly due to chain end corrections. The situation is qualitatively different for the intrinsic width of the interface (c.f. panel (b)). MC simulations and SCF calculations, which

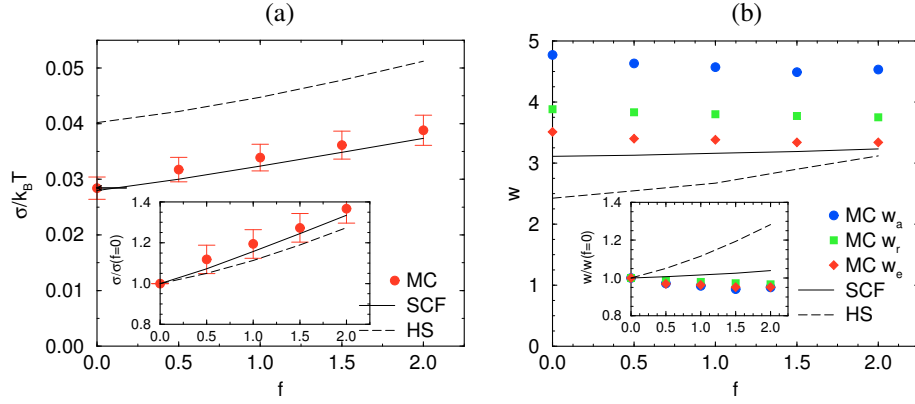


Figure 13. Interfacial tension **(a)** and interfacial width **(b)** in a blend of flexible ($f = 0$) and stiff (f as indicated) polymers at rather strong segregation $\epsilon = 0.05$. Comparisons with detailed SCF calculations, which take account of the chain architecture on all length scales, and predictions of the Gaussian chain model by Helfand and Sapse (HS) are shown. In panel **(b)** w_a denotes the apparent width, which is averaged over the whole lateral system size $L = 64$, w_r represents the width on the block size $B = 16$, and $w_e = 2\Delta e/\chi T\rho$ denotes the width extracted from the excess energy Δe of the interface per unit area, respectively. From Müller and Werner.²⁵

enumerate explicit single chain conformations extracted from the MC simulations, predict no increase or even a reduction of the width for larger stiffness, while the intrinsic width increases in the Gaussian chain model. The break down of the Gaussian chain model can be qualitatively rationalized as follows: The width of the interface is determined by loops of the polymers into the other phase. For large incompatibility the width of the interface becomes comparable with the persistence length and the conformation of a loop differs from the Gaussian statistic of the chain on large length scales. Likewise, MC simulations, SCF calculations and predictions of the Gaussian chain model agree for smaller χ , where $w \gg b$.

The ability of the SCF calculations to provide a detailed description of the intrinsic interface structure is illustrated in Fig.14. We present the MC and SCF results for the orientations as measured by the 2nd Legendre polynomial of the angle between the bond vector or end-to-end distance $\mathbf{R}_e$ with respect to the interface. Both vectors align parallel to the interface, but the effect is much stronger for $\mathbf{R}_e$ and more stiffness dependent for the bond vectors.

Diblock copolymers are model surfactants for the AB homopolymer blend. They adsorb at the interface as to extend both halves into the corresponding homopolymer phases. This decreases their enthalpy, but the localization at the interface reduces the translational entropy and the conformational entropy due to chain stretching at high copolymer excess at the interface. Upon increasing the chemical potential $\delta\mu$ (or concentration) of the copolymers in the bulk, we observe the adsorption of copolymers at the interface and the concomitant reduction of the interfacial tension in Fig.15 (cf. also Fig.1). Both MC simulations and SCF calculations agree at high segregation. However, rather than forming a dense copolymer brush at the interface, a phase separation into a homopolymer-rich phase and a lamellar phase (swollen by homopolymers) is encountered.

For small ϵ the addition of copolymers drives the system to compatibility (c.f.

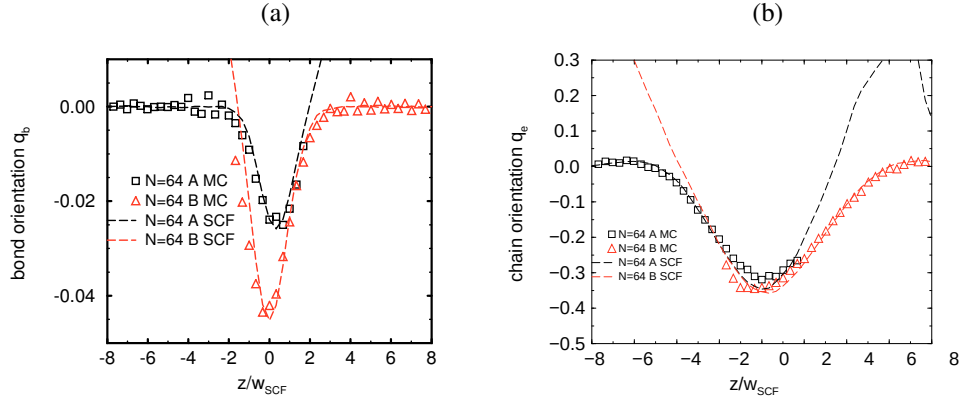


Figure 14. Orientations of the bond vector **(a)** and $\mathbf{R}_e$ **(b)** in a blend of polymers with stiffness $f = 0$ (left side) and $f = 1$ (right side) at rather strong segregation $\epsilon = 0.05$ and chain length $N = 64$. Symbols denote the results of the MC simulation on a lateral length scale $B = 16$, while lines represent the SCF calculations. From Müller and Werner.²⁵

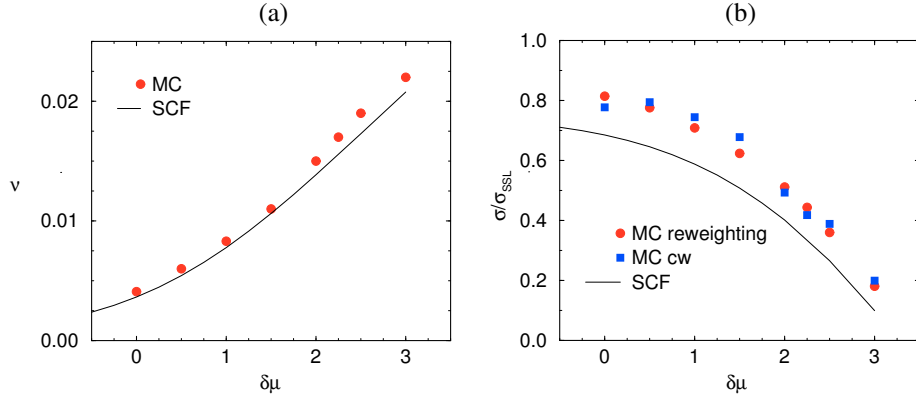


Figure 15. **(a)** Adsorption of diblock copolymers at a homopolymer/homopolymer interface as a function of the chemical potential of the copolymer at $\epsilon = 0.1$ and chain length $N = 32$. **(b)** Reduction of the interfacial tension upon adding copolymers. From Werner et al.⁹⁹

Fig.16(a)). At intermediate segregation we find a three phase coexistence between two homopolymer-rich phases and a copolymer-rich disordered phase. The latter has a structure of a microemulsion (as revealed, e.g., by snapshots). SCF calculations by Janert and Schick¹⁰⁰ rather predict highly swollen lamellar phases in this region. Some insight into this discrepancy can be gained from the spectrum of interface fluctuations. Upon adding copolymers to the systems at $\epsilon = 0.054$ we decrease the interfacial tension and deviations from a simple q^2 dependence become apparent (see panel (b)). We can obtain a rough estimate of the bending rigidity κ of the copolymer-loaden interface according to Eq.(37). The bending rigidity turns out to be much smaller than $T/2\pi$. It is this bending rigidity, however, which stabilizes the liquid-crystalline order of the lamellar phase. De Gennes and Taupin¹⁰¹ argued that a small value of κ leads to the formation of a microemulsion. Indeed this is observed in the simulation¹⁰² and experiments.¹⁰³ If we were to increase the

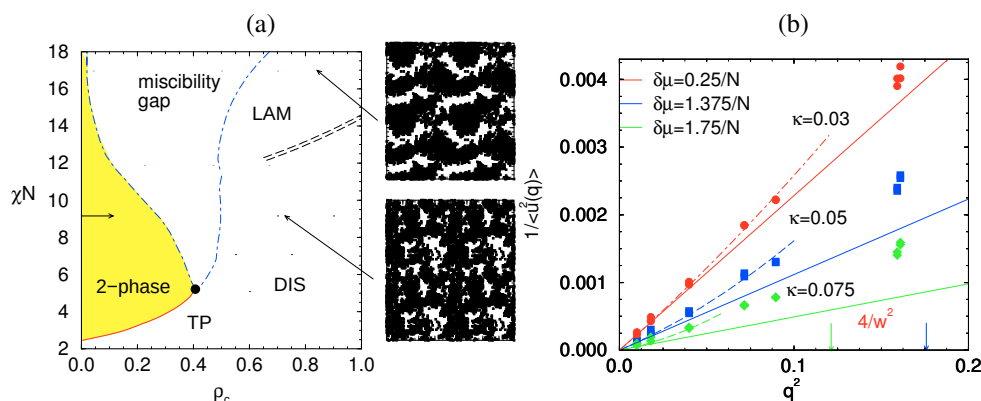


Figure 16. (a) Isopleth cut through the phase triangle of a ternary blend. (b) Spectrum of interface fluctuations in a ternary blend of two homopolymers and a diblock copolymer at $\epsilon = 0.054$. Estimates for the bending rigidity of the interfaces are indicated. From Müller and Schick.⁴²

chain length we would increase the bending rigidity $\kappa \sim \sqrt{N}^{104}$ and stabilize the lamellar phases predicted by the SCF theory. This reasoning agrees with the Ginzburg criterium at the (tricritical) Lifshitz point.¹⁰⁵

Acknowledgments

It is a great pleasure to thank Anna Cavallo for a critical reading of the manuscript. Fruitful and enjoyable collaborations with K. Binder, A. Cavallo, M. Fuchs, L. Schäfer, M. Schick, F. Schmid, A. Werner, financial support from the DFG under grants Mu1674/1 and access to the CRAY T3Es at the NIC Jülich and the HLR Stuttgart are acknowledged.

References

1. C. Solc, (edt.) *Polymer Compatibility and Incompatibility, Principles and Practices*, Haarwood Academic, Chur, (1980). C.D. Han (edt.) *Polymer Blends and Composites in Multiphase Systems*, ACS, Washington D.C. R.W. Cahn, P. Haasen, and E.J. Kramer, *Materials Science and Technology, A Comprehensive Treatment*, Vol 12, VCH, Weinheim (1993). D.R. Paul and S. Newman, *Polymer blends*, AP, New York (1978). D.S. Walsh, J.S. Higgins and A. Maconnachie, *Polymer blends and mixtures*, Martinus Nijhoff Publishers, Dordrecht (1985).
2. C. Creton, E.J. Kramer, and G. Hadziioannou, *Macromolecules* **24**, 1846 (1991).
3. G.I. Taylor, *Proc.R.Soc. London*, **A 138**, 41 (1932).
4. S.T. Milner, *MRS Bull.* **22**, 38 (1997).
5. P. Van Konynenburg and R.L. Scott, *Philos Trans Soc London Series A* **298**, 495 (1980).
6. E. Helfand, Y. Tagami, *J.Chem.Phys.* **56**, 3592 (1972); E. Helfand, *J.Chem.Phys.* **62**, 999 (1975).
7. J. Noolandi, K.M. Hong, *Macromolecules* **14**, 727 (1981); *ibid* **15**, 483 (1982).

8. K.R. Shull, *Macromolecules* **26**, 2346 (1993).
9. J.M.H.M. Scheutjens, G.J. Fleer, *J.Phys.Chem.* **83**, 1619 (1979); *ibid* **84**, 178 (1979); *Macromolecules* **18**, 1882 (1985).
10. M.W. Matsen and M. Schick, *Phys.Rev.Lett.* **74**, 4225 (1995).
11. M.L. Huggins, *J.Chem.Phys.* **9**, 440 (1941); P.J. Flory, *J.Chem.Phys.* **9**, 660 (1941).
12. K.S. Schweizer and J.G. Curro, *Advances in Chemical Physics*, **XCVIII**, 1 1997. Prigogine and S.A.Rice (eds.), Wiley.
13. K.W. Foreman and K.F. Freed, *Advances in Chemical Physics* **103**, 335 (1998); K.F. Freed and J. Dudowicz, *Macromolecules* **31**, 6681 (1998).
14. A. Sariban and K. Binder, *Macromolecules* **21**, 711 (1988); **24**, 587 (1991); *J.Chem.Phys.* **86**, 5859 (1987); *Makromol.Chem.* **189**, 2357 (1988).
15. F.S. Bates, M.F. Schultz, J.H. Rosedale, and K. Almdal, *Macromolecules* **25**, 5547 (1992); M.D. Gehlsen and F.S. Bates, *Macromolecules* **27**, 3611 (1994); F.S. Bates and G.H. Fredrickson, *Macromolecules* **27**, 1065 (1994).
16. D. Schwahn, G. Meier, K. Mortensen, and S. Janssen, *J.de.Phys. II (France)* **4**, 837 (1994); H. Frielinghaus, D. Schwahn, L. Willner, and T. Springer, *Physica B* **241**, 1022 (1998).
17. P. Cifra, E. Nies, and J. Broersma, *Macromolecules* **29**, 6634 (1996); P. Cifra, F.E. Karasz, and W.J. MacKnight, *Macromolecules* **25**, 192 and 4895 (1992).
18. H.-P. Deutsch and K. Binder, *Europhys. Lett.* **18**, 667 (1992); *Macromolecules* **25**, 6214 (1992); *J.Phys. II (France)* **3** 1049 (1993).
19. M. Müller and K. Binder, *J.Phys. II (France)* **6**, 187 (1996).
20. M. Murat and K. Kremer, *J.Chem.Phys.* **108**, 4340 (1998).
21. D. Heine, D.T. Wu, J.G. Curro, and G.S. Grest, *J.Chem.Phys.* **118**, 914 (2003).
22. M. Müller, *Macromolecules* **31**, 9044 (1998).
23. K.A. Peterson, A.D. Stein, and D.M. Fayer, *Macromolecules* **23**, 111 (1990).
24. I. Szleifer, A. Ben-Shaul, and W.M. Gelbhart, *J.Chem.Phys.* **1986**, **85**, 5345 (1986). *ibid.* **86**, 7094 (1987). I. Szleifer and M.A. Carignano, *Adv. Chem. Phys.* **94**, 742 (1996).
25. M. Müller and A. Werner, *J.Chem.Phys.* **107**, 10764 (1997).
26. P.G. de Gennes, *Scaling concepts in polymer physics*, Cornell University Press, **1979**, Ithaca.
27. I.M. Lifshitz, A.Y. Grosberg, and A.R. Khokhlov, *Rev.Mod.Phys* **50**, 683 (1978).
28. J.K. Maranas, M. Momdello, G.S. Grest, S.K. Kumar, P.G. Debenedetti, W.W. Graessley, *Macromolecules* **31**, 6991 and 6998 (1998); J.D. Londono, J.K. Maranas, M. Mondello, A. Habenschuss, G.S. Grest, P.G. Debenedetti, W.W. Graessley, S.K. Kumar, *J.Poly.Sci. B* **36**, 3001 (1998).
29. I. Carmesin and K. Kremer, *Macromolecules* **21**, 2819 (1988); H.-P. Deutsch and K. Binder, *J.Chem.Phys.* **94**, 2294 (1991).
30. V. Tries, W. Paul, J. Baschnagel, and K. Binder, *J.Chem.Phys.* **106**, 738 (1997).
31. N.F. Carnahan and K.E. Starling, *J.Chem.Phys.* **51**, 635 (1969). H.-P. Deutsch and R. Dickman, *J.Chem.Phys.* **93**, 8983 (1990).
32. M. Müller and K. Binder, *Macromolecules* **28**, 1825 (1995); M. Müller and K. Binder, *Computer Phys. Comm.* **84**, 173 (1994).
33. K. Kremer and K. Binder, *Comp.Phys.Rep.* **7**, 261 (1988).
34. W.G. Madden, A.I. Pesci, and K.F. Freed, *Macromolecules* **23**, 1181 (1990).

35. P. Cifra, K.F. Karasz, and W.J. McKnight, *Macromolecules* **25**, 4895 (1992); E. Nies and P. Cifra, *Macromolecules* **27**, 6033 (1994).
36. S.K. Kumar, *Macromolecules* **27**, 260 (1994); S.T. Cui, H.D. Cochran, P.T. Cummings, and S.K. Kumar, *Macromolecules* **27**, 3375 (1997).
37. S.K. Kumar, I. Szleifer, and A.Z. Panagiotopoulos, *Phys.Rev.Lett.* **66**, 2935 (1991) and *Phys.Rev.Lett.* **68**, 3456 (1992).
38. F.A. Escobedo and J.J. dePablo, *Europhys. Lett.* **40**, 111 (1997); F.A. Escobedo and J.J. dePablo, *J.Chem.Phys.* **106**, 2911 (1997).
39. D.A. Kofke, *J.Chem.Phys.* **98**, 4149 (1993).
40. R.D. Kaminski, *J.Chem.Phys.* **101**, 4986 (1994).
41. M. Müller, K. Binder, and W. Oed, *J.Chem.Soc.Faraday Trans.* **91**, 2369 (1995).
42. M. Müller and M. Schick, *J.Chem.Phys.* **105**, 8885 (1996).
43. M. Müller and N.B. Wilding, *Phys.Rev.* **E 51**, 2079 (1995).
44. A.M. Ferrenberg and R. Swendsen, *Phys.Rev.Lett.* **61**, 2635 (1988).
45. A.M. Ferrenberg and R. Swendsen, *Phys.Rev.Lett.* **63**, 1195 (1989).
46. B. Berg, *J.Stat.Phys.* **82**, 323 (1996).
47. F. Wang and D.P. Landau, *Phys.Rev.Lett.* **86**, 2050 (2001).
48. Q. Yan and J.J. de Pablo, *Phys.Rev.Lett.* **90**, 035701 (2003).
49. U. Hansmann and L. Wille, *Phys.Rev.Lett.* **88**, 068105 (2002).
50. B. Schulz, K. Binder, and M. Müller, *J.Mod.Phys.* **C 13**, 477 (2002).
51. J.S. Wang, T. Tay, and R. Swendsen, *Phys.Rev.Lett.* **82**, 476 (1999).
52. G. Smith and A. Bruce, *J.Phys.* **A 28**, 6623 (1995).
53. P. Virnau and M. Müller (preprint, 2003).
54. M.D. Gehlen, J.H. Rosedale, F.S. Bates, G.D. Wignall, K. Almdal, *Phys.Rev.Lett.* **68**, 2452 (1992).
55. A. Yethiraj and K.S. Schweizer, *J.Chem.Phys.* **97**, 5927 (1992) and **98**, 9080 (1993); K.S. Schweizer and A. Yethiraj, *J.Chem.Phys.* **98**, 9053 (1993); C. Singh, K.S. Schweizer, and A. Yethiraj, *J.Chem.Phys.* **102**, 2187 (1995).
56. M. Müller, *Macromol.Theory Simul.* **8**, 343 (1999).
57. V.L. Ginzburg, *Sov.Phys.Solid State* **1**, 1824 (1960); P.G. de Gennes, *J.Phys.Lett. (Paris)* **38**, L-441 (1977); J.F. Joanny, *J.Phys.A* **11**, L-117 (1978); K. Binder, *Phys.Rev.A* **29**, 341 (1984).
58. M.E. Fisher, *Rev.Mod.Phys.* **46**, 587 (1974); K.K. Mon and K. Binder, *Phys.Rev.* **E 48**, 2498 (1993).
59. G.S. Grest, M.-D. Lacasse, K. Kremer, A.M. Gupta, *J.Chem.Phys.* **105**, 10583 (1996).
60. F.A. Escobedo and J.J. dePablo, *Macromolecules* **32**, 900 (1999).
61. A. Cavallo, M. Müller, and K. Binder, *Europhys. Lett.* **61**, 214 (2003).
62. A.J. Silverberg, *Colloid Interface Sci* **90**, 86 (1982).
63. Since we focus on semi-dilute solutions (and melts), we express all quantities for $r > \xi_{ev}$ in terms of R_e and $\mathcal{N}$. Conceiving a chain as a random walk of blobs of size ξ_{ev} , we can write: $R^2 = b^2 N \sim \xi_{ev}^2 N/g$ where g is the number of monomeric units per blob. To a first approximation, blobs do not interdigitate and $g \sim \rho \xi_{ev}^3$. Then we obtain $\xi \sim 1/\rho b^2 = R_e/(\sqrt{\mathcal{N}})$. To make contact with the dilute limit, we use $R_0 \sim b_{ev} N^{\nu_{ev}}$ with $\nu_{ev} = 0.588$ for the chain extension in the limit $\rho \rightarrow 0$, and we define the overlap $s = \rho R_0^3/N$. The extension in semi-dilute solutions is related to the properties in the dilute limit via: $R_e^2 \sim R_0^2 \mathcal{F}(s) \xrightarrow{s \rightarrow \infty}$

- $R_0^2 s^{(1-2\nu_{ev})/(3\nu_{ev}-1)} \sim b_0(\rho b_0^3)^{(1-2\nu_{ev})/(3\nu_{ev}-1)} N$. Using this limiting scaling behavior we find for the screening length $\xi_{ev} \sim R_e/\sqrt{N} \sim b_0(\rho b_0^3)^{\nu_{ev}/(1-3\nu_{ev})}$ and for the overlap $s \sim N^{3\nu_{ev}-1}$.
64. M. Müller and K. Binder, *Macromolecules* **31**, 8323 (1998).
 65. M. Müller, K. Binder, E.V. Albano, *Int.J.Mod.Phys.* **B15**, 1867 (2001).
 66. K. Binder, D.P. Landau, M. Müller, *J.Stat.Phys.* **110**, 1411 (2003).
 67. M. Müller, *Macromolecules* **28**, 6556 (1995).
 68. M. Dijkstra and D. Frenkel, *Phys.Rev.Lett.* **72**, 292 (1994); M. Dijkstra, D. Frenkel, and J.P. Hansen, *J.Chem.Phys.* **101**, 3179 (1996).
 69. W.W. Graessley, R. Krishnamoorti, N.P. Balsara, L.J. Fetters, D.J. Lohse, D.N. Schulz, and J.A. Sissano, *Macromolecules* **27**, 2574, 3073 and 3896 (1994); R. Krishnamoorti, W.W. Graessley, G.T. Dee, D.J. Walsh, L.J. Fetters, and D.J. Lohse, *Macromolecules* **29**, 367 (1996); G.C. Reichart, W.W. Graessley, R.A. Register, R. Krishnamoorti, and J.D. Lohse, *Macromolecules* **30**, 3036 (1996); W.W. Graessley, R. Krishnamoorti, G.C. Reichart, N.P.Balsara, L.J. Fetters, and D.J. Lohse, *Macromolecules* **28**, 1260 (1995); R. Krishnamoorti, W.W. Graessley, L.J. Fetters, R.G. Garner, and D.J. Lohse, *Macromolecules* **28**, 1252 (1995).
 70. J.H. Hildebrand and R.L. Scott, *The solubility of nonelectrolytes*, Dover Publications, New York (1964); J.H. Hildebrand, J.M. Prausnitz, R.L. Scott, *Regular and related solutions*, Van Nostrand Reinhold Co., New York (1970).
 71. J.P. Wittmer, W. Paul, and K. Binder, *Macromolecules* **25**, 7211 (1992).
 72. G.H. Fredrickson, A.J. Liu, and F.S. Bates, *Macromolecules* **27**, 2503 (1994). G.H. Fredrickson and A.J. Liu, *J.Polym.Sci.* **B33**, 1203 (1995).
 73. C. Singh and K.S. Schweizer, *Macromolecules* **28**, 8692 (1995); *J.Chem.Phys.* **103**, 5814 (1995); *Macromolecules* **30**, 1490 (1997); K.S. Schweizer and C. Singh, *Macromolecules* **28**, 2063 (1995).
 74. K.W. Foreman and K.F. Freed, *Macromolecules* **30**, 7279 (1997); K.W. Foreman, K.F. Freed, and I.M. Ngola, *J.Chem.Phys.* **107**, 4688 (1997); K.F. Freed and J. Dudowicz, *Macromolecules* **29**, 625 (1996).
 75. J.D. Weinhold, S.K. Kumar, C. Singh, and K.S. Schweizer, *J.Chem.Phys.* **103**, 9460 (1995).
 76. S.K. Kumar and J. Weinhold, *Phys.Rev.Lett.* **77**, 1512 (1996).
 77. J. Dudowicz, K.F. Freed, and J.F. Douglas, *J.Chem.Phys.* **116**, 9983 (2002); *Phys.Rev.Lett.* **88**, 095503 (2002).
 78. M. Rabeony, D.J. Lohse, R.T. Garner, S.J. Han, W.W. Graessley, and K.B. Migler, *Macromolecules* **31**, 6511 (1998).
 79. M. Müller and W. Paul, *J.Chem.Phys.* **100**, 719 (1994).
 80. J.K. Taylor-Maranas, P.G. Debenedetti, W.W. Graessley, and S.K. Kumar, *Macromolecules* **30**, 6943 (1997).
 81. J.F. Joanny and H. Benoit, *Macromolecules* **30**, 3704 (1997); J.F. Joanny, H. Benoit, and W.H. Stockmayer, *Macromol.Symp.* **121**, 95 (1997).
 82. D.G. Gromov and J.J. dePablo, *J.Chem.Phys.* **109**, 10042 (1998).
 83. M. Müller, K. Binder, and L. Schäfer, *Macromolecules* **33**, 4568 (2000).
 84. M. Fuchs and M. Müller, *Phys. Rev.* **E 60**, 1921 (1999).
 85. J.-F. Joanny, L. Leibler, and R. Ball, *J.Chem.Phys.* **81**, 4640 (1984).
 86. L. Schäfer and C. Kappeler, *J.Chem.Phys.* **99** 6135 (1993); L. Schäfer and C. Kap-

- peler, J.Phys. (France) **46**, 1853 (1985).
87. C.J. Grayce, A. Yethiraj, K.S. Schweizer, J.Chem.Phys. **100**, 6857 (1994). C.J. Grayce, K.S. Schweizer, J.Chem.Phys. **100**, 6846 (1994).
 88. J.D. Weinhold, J.G. Curro, A. Habenschuss, J.D. Londono, Macromolecules **32**, 7276 (1999).
 89. F. Eurich and P. Maass, J.Chem.Phys. **114**, 7655 (2001).
 90. R. Holyst and T.A. Vilgis, Macromol.TheorySimul, **5**, 573 (1996).; R. Holyst and T.A. Vilgis, J.Chem.Phys. **99**, 4835 (1993); Phys.Rev.E **50**, 2087 (1994); M.G. Brereton and T.A. Vilgis, J.Phys. (France) **50**, 245 (1989); A. Aksimentiev and R. Holyst, Macromol. Theory Simul. **7**, 447 (1998).
 91. G.E. Garas and M.K. Kosmas, J.Chem.Phys. **103**, 10790 (1995); *ibid* **105**, 4789 (1996); *ibid* **108**, 376 (1998); M.K. Kosmas, Macromolecules **22**, 720 (1989).
 92. A.N. Semenov, J.Phys. II **6**, 1759 (1996).
 93. F. Schmid and M. Müller, Macromolecules **28**, 8639 (1996).
 94. A. Werner, F. Schmid, M. Müller, and K. Binder, Phys.Rev. **E 59**, 728 (1999).
 95. A. Werner, F. Schmid, M. Müller, and K. Binder, J.Chem.Phys. **107**, 8175 (1997).
 96. M.D. Lacasse, G.S. Grest, and A.J. Levine, Phys.Rev.Lett. **80**, 309 (1998).
 97. A.N. Semenov, Macromolecules **27**, 2732 (1994).
 98. E. Helfand and A.M. Sapse, J.Chem.Phys. **62**, 1329 (1975).
 99. A. Werner, F. Schmid, and M. Müller, J.Chem.Phys. **110**, 5370 (1999).
 100. P.K. Janert and M. Schick, Macromolecules **30**, 3916 (1997).
 101. P.G. de Gennes and C. Taupin J.Phys.Chem. **86**, 2294 (1982).
 102. M. Müller and M. Schick, J.Chem.Phys. **105**, 885 (1996).
 103. G.H. Fredrickson and F.S. Bates, Polym.Sci. **B 35**, 2775 (1997).
H.S. Jeon, J.H. Lee, and N.P. Balsara, Phys.Rev.Lett. **79**, 3274 (1997).
 104. M.W. Matsen, J.Chem.Phys. **110**, 4658 (1999).
 105. M. Müller and G. Gompper, Phys.Rev. **E 66**, 041805 (2002).

Inhomogeneous Coarse-Graining of Polymers and Polymer/Metal Interfaces

Cameron F. Abrams

Department of Chemical Engineering, Drexel University,
Philadelphia, Pennsylvania 19104, U.S.A.,
E-mail: cfa22@drexel.edu

A “molecule-fixed” inhomogeneous coarse-graining scheme for bisphenol-A polycarbonate (BPA-PC) represents the first attempt to use both atomistic and coarse-grained resolution simultaneously to model a specific polymer [Abrams, et. al, *Phys. Rev. E* **67**:021807]. In this model, non-terminal BPA-PC chemical repeat units are each represented by four spherical beads, corresponding to four comonomeric functional groups. At chain ends, the terminal carbonate groups are represented atomistically. Interactions between the respective beads and a nickel surface were parameterized using *ab initio* calculations. This model is reviewed here, and we discuss its application in molecular dynamics simulations of dense liquids of BPA-PC adjacent to (111) nickel surfaces. We observe that the structure of the liquid near the wall is sensitively affected by (a) the strong attraction of the chain ends to the surface, and (b) the strong orientational dependence of this attraction. On scales larger than the bulk unperturbed radius of gyration, the solid/liquid interface has two distinct liquid layers: in the innermost layer near the wall, most chains have both ends adsorbed, while in the outermost layer, most chains have a single end adsorbed. Chain conformations are generally flattened in the innermost layer and stretched in the outermost layer. The overlap between these two layers is more diffuse in the case of the longer chains. Because the transferability of “molecule-fixed” inhomogeneous models to other polymer/metal interface systems seems unlikely, we conclude by speculating on the more general notion of “domain-fixed” inhomogeneously resolved models for polymeric systems. These speculations are based upon preliminary simulations of a domain-fixed dual-resolution model of liquid methane.

1 Introduction

Simulation of specific polymer systems has recently been greatly advanced by employing systematically derived coarse-grained molecular models.^{1–4} These models vary in the details of their development and implementation, yet all seek to construct particle- or lattice-based chain-like objects whose constituents represent at most a few chemical repeating units of any specific polymer. This is meant to circumvent the great difficulty in producing equilibrated samples of atomically-resolved polymers due to the $O(10)$ orders of magnitude spread between the resolution required in time for standard atomistic molecular dynamics (MD) simulation (10^{-15} s) and the slow molecular relaxation times ($\sim 10^{-5}$ s) which must be achieved. Monte Carlo simulation fares no better as an alternative approach, especially in dense multimolecular systems, because local moves are constrained to impractically small values by steep bonded potentials, and enormous numbers of successful moves have to accumulate in order to move chain sections larger than a repeat unit or so. Integrating out these fast motions, or put another way, averaging over their underlying steep potentials, allows us in principle to construct models that can produce equilibrated configurations with relatively much less computational effort.

Coarse-graining aims to guarantee that the chain conformations in a simulation sample represent true equilibrium conformations of the specific polymer considered. A central

feature of coarse-graining is that the models retain only as much unique and relevant information as needed about the specific polymer(s) under investigation, using significantly fewer degrees of freedom (*i.e.*, particles) than required for full atomistic detail. The coarsened degrees of freedom must be constrained within ensembles of configurations which represent an appropriate average over the microscopic atomic-scale potential energy surface of the fully resolved system. Once an equilibrated sample at the coarse-grained level is generated using an appropriate simulation technique which samples these ensembles, atomic details can be inverse mapped to study atomic-scale properties and processes on then appropriately short length and time scales.⁵

Our focus in recent years has been developing a method for coarse graining melts of bisphenol A-polycarbonate (BPA-PC)^{6,5,7-9} (Fig. 1). BPA-PC not only provides a challenging test case for this type of modeling, it is also by far the most utilized and intensively studied variety of polycarbonate, thanks to its many valuable material properties, such as high impact strength, ductility, glass transition and melting temperatures¹⁰. Coarse-grained simulations can potentially fill a gap that now exists between traditional atomistic simulation and experiments, by addressing longer length and time scales, allowing one to test hypotheses regarding inter- and intramolecular interplay in packing and entanglement that are otherwise more difficult to approach.

Metal surfaces are particularly relevant for industrial processing of polycarbonate, and indeed many other synthetic polymers. The interplay between specific local adsorption of organic chain segments and the much larger scale chain conformations is difficult to capture with traditionally simple molecular and atomic-scale models. Our contributions so far have presented a technique involving appropriate coupling of information from detailed *ab initio* calculations of small molecule/surface interactions into coarse-grained molecular dynamics simulations of liquids of comparatively much larger macromolecules.^{8,11} For the purposes of these proceedings, we are particularly interested in understanding the adhesion behavior of BPA-PC liquids next to crystalline nickel surfaces, and how chain length affects this behavior.

Polymer/surface interactions are especially challenging to capture with coarse-grained models due to the specificity of the interactions for various comonomeric groups in complicated polymers, such as BPA-PC. The challenge lies in the fact that the averaging over the microscopic potential energy surface which gives coarse-grained potentials for bulk BPA-PC is strictly not valid when chains strongly interact with a surface, which is the case for the BPA-PC/Ni. For this reason, we have constructed a dual-resolution representation of BPA-PC— combining both coarse-grained and atomic-scale resolution in the same representation – in order to capture as much of the specificity of the polymer/surface interaction as possible, while maintaining the computational advantages of coarse-graining¹¹. In a recent publication, we presented the details of this dual-resolution scheme (which are briefly recapitulated below) and presented preliminary results on the structure of a 10-repeat-unit BPA-PC liquid near a Ni (111) surface¹¹. Here, we expand upon the results presented in Ref.¹¹ to consider the effects of longer chains on the structure of the melt near the surface.

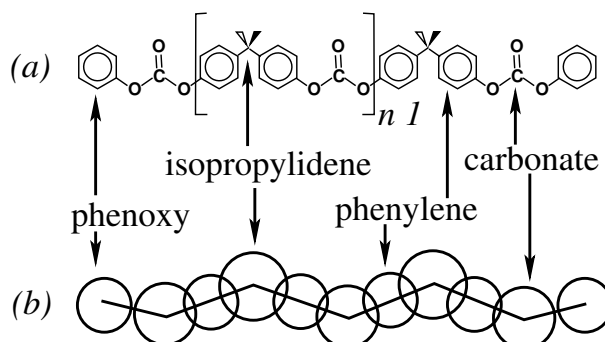


Figure 1. (abrams/a) Schematic representation of the structure of bisphenol A-polycarbonate. The relevant comonomeric groups are labeled. The n th repeat unit is outside the brackets because we distinguish *terminal* carbonate groups from *internal* carbonate groups in the dual-resolution model. (b) Schematic representation of the 4:1 coarse-grained model of BPA-PC, detailed in Ref.⁹.

2 The “Molecule-Fixed” Dual-Resolution Coarse-Grained Model of BPA-PC

2.1 The 4:1 Single-Resolution Model

For a complete presentation of the dual-resolution model of BPA-PC and simulation techniques, we refer the reader to Ref.¹¹. Fig. 1 shows the chemical structure of a BPA-PC molecule composed of n repeat units. The repeat unit is constructed of four sequential subunits: a phenylene, an isopropylidene, a second phenylene, and a carbonate. The ends of a BPA-PC molecule are phenyls connected to carbonates. Because the terminal phenyls are each bound to an oxygen of their respective carbonate, we term these “phenoxy” ends. Other end groups apart from phenoxies are currently being considered in work that is not discussed here.¹²

The 4:1 coarse-grained representation of BPA-PC⁹ replaces the atomic structure of the repeat unit, which contains 33 atoms, with four spherical beads, two of which are centered on the backbone carbons of the carbonate and isopropylidene groups, respectively. The other two beads “float” at a fixed distance along a line connecting adjacent pairs of carbonate-isopropylidene (“c”–“i”), and are meant to roughly occupy space ascribed to the phenylene (“p”) groups. The diameters of these beads are chosen so that they have proper relative sizes and as a group occupy the proper van der Waals excluded volume of the repeat unit.¹³ These diameters parameterize repulsive pairwise Weeks-Chandler-Andersen 12–6 potentials¹⁴ for the non-bonded interactions among beads.

The beads adjacent along the chain are connected by stiff harmonic bonds whose lengths are determined from the atomic-scale geometry. In addition to this connectivity, and the non-bonded pair potentials, the overall potential also operates on three types of bond angles: c-p-i, c-i-c, and i-c-i. All c-p-i angles are held at 180° by stiff harmonic potentials. The angles defined by c-i-c and i-c-i triples (skipping the bridging phenylene beads) are constrained according to Boltzmann-inverted potentials tabulated via Monte-Carlo averaging over the atomic scale potential. At a specified temperature, the value

of any c-i-c or i-c-i angle is thus constrained to the correct distribution consistent with a canonical average over the atomic scale potential.

The philosophy behind this method of coarse-graining has been discussed previously^{6,9}. Briefly, there are three major assumptions used. (1) The coarse-grained degrees of freedom are uncorrelated, allowing us to write a coarse-grained potential with additive contributions from each degree of freedom. (2) The intramolecular potential at the atomic level is not influenced by intermolecular degrees of freedom. This is a safe assumption for most synthetic (carbon-based) polymers at temperatures below those at which thermal decomposition can occur. (3) The intermolecular interactions at the coarse-grained level are approximated to an acceptable level of accuracy by pair-wise repulsive potentials, at least for simulating equilibrium structures.

2.2 Adapting the 4:1 Model for Surface Studies: *Ab Initio* Calculations

The coarse-graining procedure described above yields models which are most appropriately used in bulk simulations. However, it is of interest presently to understand how polymer molecules behave at metal interfaces. We have chosen to consider the specific system of BPA-PC on nickel (111), for the reasons alluded to in the Introduction. What changes must be made to the current 4:1 model, if any, to allow us to use it to study the BPA-PC-Ni interface at the molecular level?

To answer this question, it would be desirable to perform *ab initio* calculations of at least a single BPA-PC molecule in the vicinity of a semi-infinite slab of nickel atoms with a (111) surface arrangement. Were such a simulation tractable, we could use it to determine the exact nature of any specific interactions between any group of atoms in the chain and any group of atoms on the surface. Appropriately averaging over all possible interaction geometries would then yield probability distributions which could be inverted into potentials for use in coarse-grained simulation.

There are two problems here. The first is that the *ab initio* calculations described are impossible based purely on the number of atoms required, with current or even next-generation computing power. The second is, even if they were possible, and even if we could gain enough statistical strength to perform the averages appropriately, we would have to greatly expand the parameter space to consider how specific interactions with the surface alter the *intramolecular* potential of the molecule. Probing all possible conformations and orientations of a chain of even one repeat unit next to a sufficiently large slab of nickel using *ab initio* techniques would be a monumental effort.

Our strategy has been to consider first small molecules which correspond to the chemical subunits of BPA-PC^{8,11}. This is motivated not only by the need of computational tractability, but also has the important advantage that the subunits we study are also important for polymers other than BPA-PC. We performed Car-Parrinello MD calculations¹⁵ to predict the interaction energies of each of these molecules with a finite slab of nickel comprised of four atomic layers. Here, we summarize the results. We observed that at all relevant orientations, carbonic acid and propane are strongly repelled from nickel (111) when the distance between center of mass and first layer Ni atoms is below 3 Å. Both benzene and phenol strongly adsorb in a horizontal orientation at a distance of 2 Å from the top layer, with a strength of about 1 eV. This interaction is short-ranged, dying off above 3 Å, and strongly orientationally dependent, decaying nearly to zero if the angle between

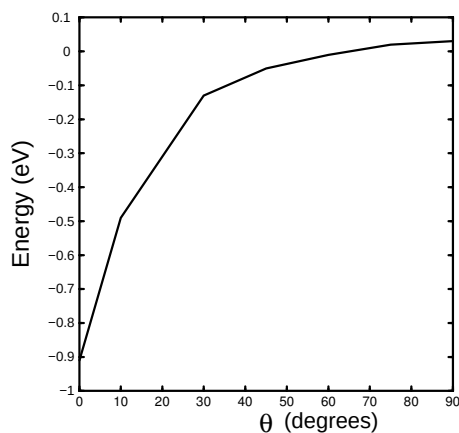


Figure 2. Adsorption energy of phenol at the Ni(111) bridge site as a function of the inclination of the carbon ring with respect to the surface.

the normal of the surface of the ring and the normal of the nickel surface plane deviate in direction by more than about 45° , as depicted in Fig. 2.

The crucial step in incorporating this information into the 4:1 coarse-grained model is to realize two things. First, although internal phenylenes would apparently prefer to stick to the nickel surface, thereby lowering the system potential energy, they are sterically hindered by the strong repulsion felt by the adjacent isopropylidene. Second, the phenoxies at the ends of the chains are not sterically hindered because the carbonate can orient with one bridging oxygen at the surface and the other away, allowing the terminal phenoxy group to strongly bind without steric hindrance. However, the orientation of the ring is a sensitive determinant of the interaction strength, and this degree of freedom does not exist in the bulk 4:1 model.

2.3 The Dual Resolution BPA-PC Model

The solution to this problem is the so-called molecule-fixed dual-resolution model, depicted in Fig. 2. The essential feature of the dual-resolution model is that the terminal *carbonates* are resolved atomically. This is true of all molecules in the system, regardless of their proximity to the surface. We therefore refer to this as a “molecule-fixed” dual-resolution scheme. Instead of using a single c-type bead to represent the outermost carbonates, we use explicitly one carbon atom, denoted “C,” two bridging oxygen atoms, denoted “O_b,” and a single carboxylic oxygen, denoted “O_C.” The orientational degree of freedom we must access is determined by a vector joining the outer bridging oxygen of the terminal carbonate group and the center of the terminal phenyl ring. The direction of this vector is the “tilt” of the phenyl ring. Because we know the torsional barriers around this bond to be on the order of thermal energy at $T = 570$ K, we assume that rotation of the phenyl ring optimizes instantaneously relative to the tilt, and hence, we need only to resolve the carbonate group to access the tilt degree of freedom. Atomic-scale potentials taken from all-atom simulations of liquid diphenyl carbonate¹⁶ are used to enforce all

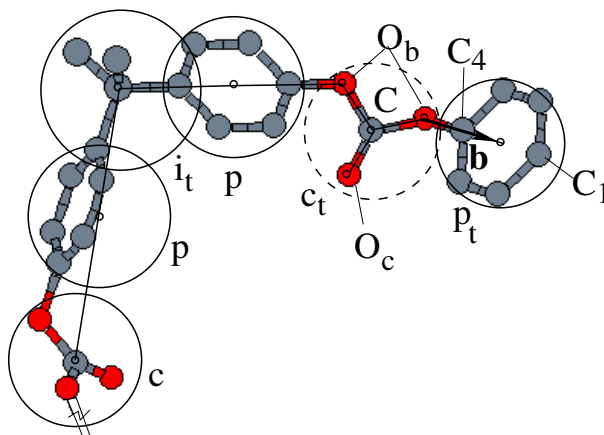


Figure 3. Atomistic structure of a representative chain-end conformation of BPA-PC illustrating the dual-resolution scheme. The small black circles are mapping points in the scheme, and the large circles represent the excluded volume diameters of those points. The C_4 - C_1 orientation vector, $\mathbf{b}$, is shown. Lower-case letters denote coarse-grained bead designations: “p” is phenylene, “ c_t ” is the terminal carbonate, “ p_t ” is the terminal phenoxy, and “ i_t ” is the terminal isopropylidene. Capital letters denote atom type designations (see text).

bonds, angles, and the one improper dihedral arising from this group.

The challenge in this dual-resolution scheme, as in any scheme that simultaneously resolves a single system with two or more levels of resolution, is handling the interface between the two levels of resolution. Here, this requires specifying how the inner O_b is bonded to its adjacent phenylene bead and how the outermost O_b is bonded to its adjacent, chain-terminating phenyl, forming the phenoxy end. For this reason, the outermost phenylidene (p in Fig. 2 owns a mapping point tethered to its center of mass, unlike other phenylidenes. Then, the p - O_b bond is enforced at its natural length by a stiff harmonic bond. The same applies for the bond between the outermost O_b and the terminal phenyl p_t .

The p - O_b -C and p_t - O_b -C angles are treated using harmonic potentials as well, as if both the p-type and p_t beads were carbon atoms. The two O_b -C torsional potentials also use the p-type beads as carbon atoms. The angle defined by the outermost isopropylidene, i_t , the inner bridging oxygen, and the carbonate carbon is constrained at its geometrically determined value of 180° using a stiff harmonic potential. Because the outermost phenylene now owns a tethered mapping point, the angle at the outermost isopropylidene is determined by the sequence c - i_t - O_b obeys a different distribution than the normal c - i - c angle, which is computed ahead of time using the same type of Monte-Carlo averaging as in the previous case.

Finally, new intermolecular interactions must be accounted for relative to the previous 4:1 case when the carbonate groups are resolved atomically. However, as a first step, we have chosen the simplest route where the carbon atom is treated in an excluded volume sense as a c-type bead, and the oxygens experience no excluded volume.

The benefit of this new scheme is that the vector connecting the outermost O_b and the terminal phenyl bead determines the orientation of the terminal phenoxy *if* we make the

assumption that a negligible atomic-scale torsional barrier on the O_b-C_4 exists (C_4 is the carbon atom of the phenyl that is bound to the bridging oxygen of the carbonate group). Indeed, this barrier is less than $1 kT$ at our temperature of interest, which is $570 K$ ¹⁶. Thus, we can include the orientation of this vector as a degree of freedom upon which we can apply a potential energy, and this allows to accurately account for both the short range and strong orientational dependencies of the phenoxy-nickel interaction. Most importantly, because we are not forced to resolve the entire molecule atomically, we know that no matter how quickly the system responds to the presence of the surface, the chains more than one repeat unit distance from the surface are locally equilibrated, and global equilibration is feasible.

The simulation system is a slit pore, and particles are confined to a central region by the walls. Particle-wall interactions are purely repulsive for internal beads, and are (111)-site-specific and orientationally dependent for the terminal phenoxy beads. The MD simulations are preformed in reduced Lennard-Jones units ($1 \sigma = 4.41 \text{ \AA}$) using a time step of 0.005τ and a Langevin-type thermostat¹⁷ with a friction of $0.5 \tau^{-1}$. Time is measured in units of $\tau \equiv \sigma \sqrt{m/k_B T}$, where m is the unit mass. Because of the stochastic forces introduced by the thermostats, the masses of the explicit C and O atoms are artificially increased to 1.0 (which is the mass of all beads in the system) to reduce their bond oscillation frequencies. In this sense, using “heavy” carbon and oxygen in the dual resolution scheme is an acceptable alternative to limiting the integration to much smaller time-steps. Our assumption is that using such heavy atoms does not greatly affect the equilibrium structure of the chains or the liquid. The systems were initialized by first growing the chains as phantom random walks in a box with periodic boundaries in all three dimensions. Periodicity was then turned off in the z -dimension, and particle position z -components were unfolded. Walls are then placed at extremes in z and slowly, over 5000 MD steps, brought toward one another until the desired density is achieved.

The results presented in Ref.¹¹ were extracted from a simulation system of 240 chains, each chain having 10 chemical repeat units. The simulations were run long enough to achieve equilibration, which in this case was assessed by considering the total phenoxy-surface interaction energy. The decay of this energy to its equilibrium value of about $12 k_B T$ per phenoxy required more than 80,000 τ of integration time for the 10 repeat unit chains. Here, we compare a selection of these results to those from a system of 400 chains, each having 20 chemical repeat units.

3 Dual-Resolution BPA-PC: Results and Discussion

Here we compare a selection of results for chains of 10 and 20 repeat units. The first result we consider is the equilibrium structure and orientation of the chains. As presented in Ref.¹¹, the chain-end sticking for 10-repeat-unit BPA-PC resulted in a unique two-layered structure in chain packing and orientation as a function of distance from the wall. The dominant length scale in this structure is the unperturbed root-mean-square radius of gyration, R_g , which was about $20 \text{ \AA}$ for the 10-repeat-unit chains. It was observed that the chains with centers of mass less than about $2R_g$ from the surface were divided clearly into two populations: those with centers of mass between 0 and R_g from the surface (layer I) had both ends adsorbed, while those between R_g and $2R_g$ (layer II) had only one end each adsorbed and were strongly stretched along the direction normal to the wall. This

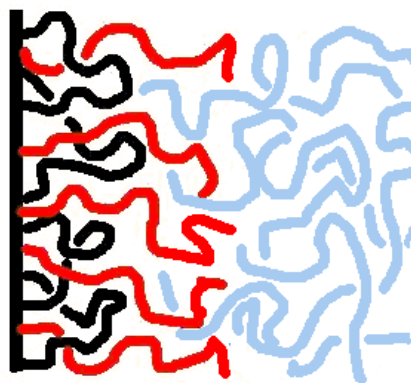


Figure 4. Schematic representation of the molecular structure of a BPA-PC melt adjacent to a Ni wall, showing double-ended adsorbed, single-ended adsorbed, and non-adsorbed chains.

is depicted schematically in Fig. 3. Moreover, the chains in layer I were flattened in the direction z normal to the wall, and chains in layer II were strongly stretched in z . The shapes of the chains are quantified in terms of the zz components of each chain's gyration tensor $\mathbf{S}$, where the trace of this tensor is R_g .

In the case of chains of 20 repeat units, the same general conclusions hold, as shown in Fig. 4. For these chains, R_g is about 29 Å. In Fig. 5(a), we plot both the density of chain centers-of-mass and S_{zz} as functions of distance from the wall, z . At a particular value of z , S_{zz} is the average of all chain S_{zz} 's for chains with centers of mass between z and $z + \delta z$. This data was averaged over the equilibrated configurations with a lateral resolution of $\delta z = 0.031 \sigma$. Flattened chains in layer I and stretched chains in layer II are evident. In Fig. 5(b), we show the breakdown of the chains by the number of ends adsorbed. Here we see that most chains in layer I have both ends adsorbed, while most in layer II have only a single end adsorbed. The delineation between populations of chains with single- and double-ended adsorption is not as clear as for the shorter chains, but the same length scaling applies: Those chains with centers of mass less than one bulk R_g from the wall have both ends stuck and are somewhat flattened, while those between one and two R_g have a single end and are stretched. The degree of flattening and stretching is not as pronounced for the 20-repeat-unit chains as for the 10-repeat-unit chains, because each chain has twice as many internal degrees of freedom upon which to call to relieve the stress associated with adsorbing a chain end to the surface while packing densely among other chains.

Finally, because the dual-resolution scheme allows us to ask where the individual C and O atoms of the carbonate groups are located, we can examine atomically resolved depth profiles involving these atoms. Fig. 6 shows the depth profiles of C and O atoms in the terminal carbonate groups, as well as the terminal phenoxy centers of mass, for the chains

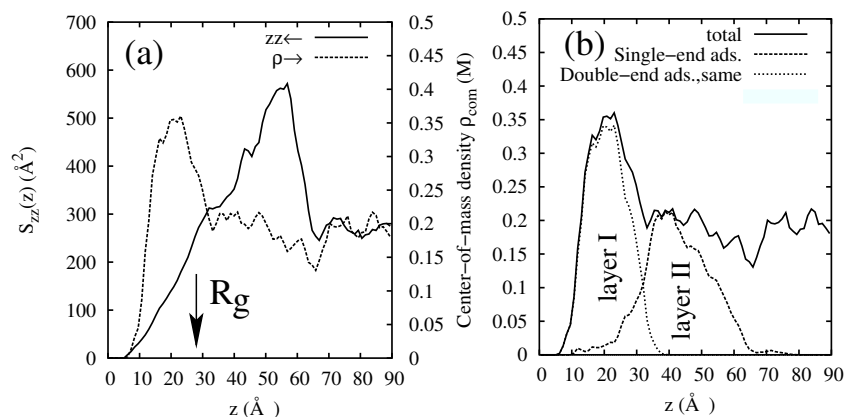


Figure 5. Simulation results for 20-repeat-unit BPA-PC: (a) zz component of the lab-frame mean depth-resolved gyration tensor vs. distance from wall, and chain center of mass density vs. distance from wall. (b) Breakdown of chains into those with single- double-ended adsorption.

with 20 repeat units. We can see that the phenoxy chain ends are tightly localized near the surface. This peak is closely shadowed by one of the two oxygen peaks, populated with those oxygen atoms which are bonded to the phenoxies. Note that the oxygen peak does not lie directly atop the phenoxy peak, as we would expect if all phenoxies are perfectly horizontal. The separation in these two peaks represents the finite temperature of the system; because of the inhomogeneity in the domain caused by the presence of the impenetrable surface, thermal fluctuations can only result in a mean orientation that is off-horizontal.

4 On the Development of Domain-Fixed Inhomogeneous Coarse-Graining Schemes: Initial Attempts with Liquid Methane

In the presentation of the molecule-fixed dual-resolution scheme, it should be noted that we have taken advantage of the unique interplay between sterics and energetics for BPA-PC-Ni, requiring that only chain ends be atomically resolved. In principle, one would rather have a technique which would decide for itself what molecular fragments must be atomically resolved and what fragments may remain coarsened, when applied to any particular system with an interface or other interesting structure with atomic resolution in at least one dimension. This would, in effect, be a means to conduct embedded full-blown atomic MD simulations with realistic boundary conditions that optimize on-the-fly. Such ideas are not new in the area of simulations of crystalline materials, where one has the luxury of a reference crystal structure upon which a coarsened mesh can be constructed (for example, Ref.¹⁸). It is also not new in the area of hybrid quantum/molecular mechanics models (QM/MM), in which a small region of a few atoms is treated quantum mechanically while being embedded in a larger system of atoms treated empirically (for example, Ref.¹⁹). The idea, however, of “domain-fixed inhomogeneous coarse-graining” for molecular *liquids* is still much more immature. Many unique considerations not yet encountered in other multiresolution techniques must be addressed.

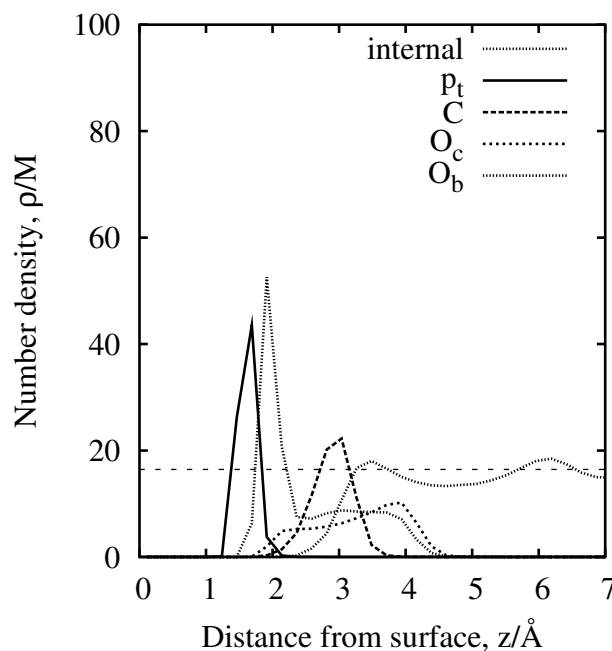


Figure 6. Simulation results for 20-repeat-unit BPA-PC. Depth profiles for atomic constituents in the dual-resolution model.

In this contribution, we discuss attempts to implement a domain-fixed inhomogeneously resolved model of liquid methane. Methane was chosen primarily because it is among the simplest polyatomic molecules (CH_4), and has been modeled in simulation studies with both fully atomic resolution (the so-called four- and five-center models) and at a coarsened molecular level (the so-called single-center models)²⁰. We wish to examine the following situation. Imagine a box, filled with liquid methane, and divided into two regions whose boundaries are *imaginary and static*; that is, they are not material boundaries. All methane molecules in region I are represented fully atomistically. All methane molecules in region II are represented as single-center coarse-grained particles. Molecules are allowed to move freely between regions; each time a five-center molecule crosses the boundary into region II, it is instantaneously mapped into a single-center molecule with a velocity vector equal to the center-of-mass velocity vector of the molecule. Likewise, each time a single-center particle crosses the boundary into region I, it is “remapped” onto a five-center molecule, each of whose atoms obtains a velocity vector equal to that of the coarse-grained particle. Such a scenario is our attempt at the simplest possible test of domain-fixed inhomogeneous coarse-graining of a dense molecular liquid.

Our first question is, what effects do the presence of these imaginary boundaries between regions of different resolution have on the structure of the liquid? To begin to answer this question, we conducted molecular dynamics simulations of the system described above, composed of 1,000 methane molecules in a box roughly 39 Å on a side, at a temperature of 149.9 K. Reduced Lennard-Jones units are again employed, and the relevant

conversion factors are $\sigma = 3.733 \text{ \AA}$, $\epsilon = 149.9 \text{ K}$, $m = 16 \text{ amu}$ (mass of one CH_4). The box is divided in two equal “slabs” by planes normal to the z -axis (the system is periodic, so we need two planes to create two regions). All molecules in slab I are atomistic, and all molecules in slab II are coarse-grained. The single-resolution potentials for methane are taken directly from Ref.²⁰. Interactions through boundaries are parameterized using standard Lennard-Jones mixing rules, except that hydrogens in the atomistic region are invisible to coarse-grained particles. The Newtonian equations of motion are augmented by the dissipative particle dynamics thermostat^{21,22}, with a friction $\zeta = 1.0 \tau^{-1}$, except for pairs interacting across a boundary, for which we chose $\zeta = 3.0 \tau^{-1}$. A single time step of 0.01τ is employed, although it would have been slightly more efficient to resolve the motion of the coarse-grained particles with larger time steps than the atomistic particles. Several runs from different initial conditions were performed, all exceeding several tens of thousands of time-steps.

We discuss a selection of results here that motivates our introductory discussion of domain-fixed inhomogeneous coarse-graining of molecular liquids. Because the imaginary boundaries impart a directionality to the system, it is logical to ask, is the distribution of mass along this direction affected? We show in Fig. 7 that there is indeed a measurable influence of the interface on the distribution of mass in the z -direction. Fig. 7 shows the number density of carbon centers (particle positions in the coarse-grained region, and methane carbons in the atomic scale region) as a function of z . The figure shows the two resolution domains. Clearly, the number density of carbons in both domains should be the same. Note, however, that the density of carbons in the atomic scale region is lower than that in the coarse-grained region. There appears to be some driving force that “pushes” more methanes into the coarse-grained region than should be allowed. The imbalance amounts to about 20 molecules, which is reached within the first few hundred time steps of the simulation.

To understand this apparent artifact, we note first that the two sets of potentials used actually predict different equations of state for methane. The coarse-grained potential, originally due to Saager and Fischer²³, predicts an equilibrium pressure of about 1040 bar (in excellent agreement with experiment) for the state point simulated here. In contrast, the 5-center atomic-scale potential, originally due to Murad and Gubbins²⁴, predicts an equilibrium pressure almost 60% higher. This higher pressure is likely the reason why particles apparently prefer to occupy the coarse-grained region. Presumably, particles flow into the coarse-grained region until the chemical potential (which to a large degree reflects the pressure) on both sides of the interface is equal. Choosing from among the many other 5-center methane potentials might yield a situation in which this density inhomogeneity disappears. It will be interesting to test whether two potentials which each correctly predict the appropriate state point properties and have the same chemical potential with respect to molecule insertion can be developed and whether they are strictly necessary for this kind of domain-fixed dual-resolution simulation.

5 Conclusions

We have reviewed the “molecule-fixed” dual-resolution representation of BPA-PC used to predict the atomic- and molecular-scale liquid structure at a BPA-PC-Ni interface. The results presented herein complement those that appeared together with the original descrip-

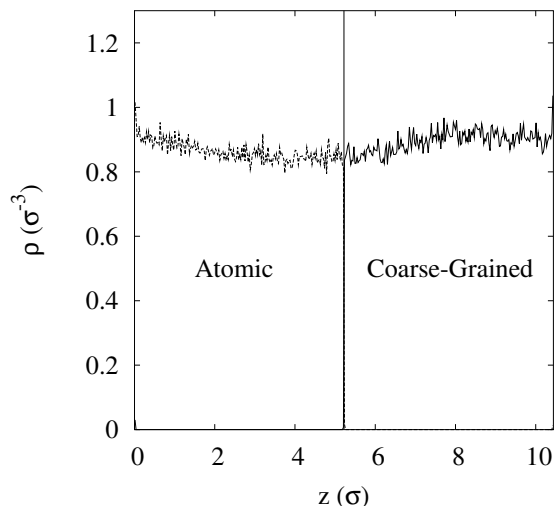


Figure 7. “Depth profiles” of carbon (local number density vs. z position) in both the coarse-grained and atomic scale regions of a domain-fixed dual-resolution simulation of liquid methane. Vertical line denotes boundary between atomically resolved and coarse-grained domains. Average carbon number density is $0.879 \sigma^{-3}$.

tion of the model¹¹. Equilibrium chain structure and orientation is qualitatively the same for the two chain lengths considered (10 and 20 chemical repeat units). The liquid structure near the wall is divided into two layers: the outermost layer (immediately adjacent to the wall) contains flattened chains with both ends adsorbed, while the innermost layer (immediately adjacent to bulk liquid BPA-PC) contains stretched chains with single-ends adsorbed.

Our technique of using insight from *ab initio* to construct a minimalistic yet realistic coarse-grained model for a synthetic polymer interacting with a specific metal surface appears to be a promising technique which could be applied to many other specific systems. However, it is recognized that a truly transferable model is one in which the dual-resolution scheme is domain-fixed, and not dependent upon details of the atomic-scale energetics. To this end, we examined some preliminary results of a domain-fixed dual-resolution model of liquid methane as a test case for this kind of simulation. Artifacts due to the presence of the imaginary boundaries between regions of different resolution were discussed.

Acknowledgments

The author is indebted to L. Delle Site and K. Kremer of the Max-Planck-Institute for Polymer Research. This work was supported by the German Ministry of Education and Research under Grant No. 03 N 6015, the Bayer Corporation, and the United States Office of Naval Research under award No. N00014-03-1-0655.

References

1. J. Baschnagel, K. Binder, P. Doruker, A. A. Gusev, O. Hahn, K. Kremer, W. L. Matice, F. Müller-Plathe, M. Murat, W. Paul, S. Santos, U. W. Suter, and V. Tries. Bridging the gap between atomistic and coarse-grained models of polymers: Status and perspectives. *Adv. Polym. Sci.*, 152:41–156, 2000.
2. W. Paul, K. Binder, K. Kremer, and D. W. Heermann. Structure property correlation of polymers, a Monte-Carlo approach. *Macromolecules*, 24(23):6332–6334, 1991.
3. K. Kremer and F. Müller-Plathe. Multiscale problems in polymer science: Simulation approaches. *MRS Bulletin*, 26(3):205–210, 2001.
4. F. Müller-Plathe. Coarse-graining in polymer simulation: From the atomistic to the mesoscopic scale and back. *ChemPhysChem*, 3(9):754–769, 2002.
5. W. Tschöp, K. Kremer, O. Hahn, J. Batoulis, and T. Bürger. Simulation of polymer melts. II. From coarse-grained models back to atomistic description. *Acta Polymer.*, 49:75–79, 1998.
6. W. Tschöp, K. Kremer, J. Batoulis, T. Bürger, and O. Hahn. Simulation of polymer melts. I. Coarse-graining procedure for polycarbonates. *Acta Polymer.*, 49:61–74, 1998.
7. Oliver Hahn, Luigi DelleSite, and Kurt Kremer. Simulation of polymer melts: From spherical to ellipsoidal beads. *Macromol. Th. Sim.*, 10(4):288–303, 2001.
8. L. Delle Site, C. F. Abrams, A. Alavi, and K. Kremer. Polymers near metal surfaces: Selective adsorption and global conformations. *Phys. Rev. Lett.*, 89(15):156103, 2002.
9. C. F. Abrams and K. Kremer. Combined coarse-grained and atomistic simulation of liquid bisphenol-A-Polycarbonate: Liquid packing and intramolecular structure. *Macromolecules*, 36:260–267, 2003.
10. L. Morbitzer and U. Grigo. Die wichtigsten polymerphysikalischen Aspekte des Polycarbonats. *Ange. Makromol. Chem.*, 162(2753):87–107, 1988.
11. C. F. Abrams, L. Delle Site, and K. Kremer. Dual-resolution coarse-grained simulation of the bisphenol-a-polycarbonate/nickel interface. *Phys. Rev. E*, 67:021807, 2003.
12. L. Delle Site, S. Leon, and K. Kremer. Bpa-pc on a ni(111) surface: The interplay between adsorption energy and conformational entropy for different chain end modifications. *J. Amer. Chem. Soc.*, 2004 (in press).
13. Isaac C. Sanchez and Junhan Cho. A universal equation of state for polymer liquids. *Polymer*, 36(15):2929–2939, 1995.
14. John D. Weeks, David Chandler, and Hans C. Andersen. Role of repulsive forces in determining the equilibrium structure of simple liquids. *J. Chem. Phys.*, 54(12):5237–5247, 1971.
15. J. Hutter, A. Alavi, T. Deutsch, M. Bernasconi, S. Goedecker, D. Marx, M. Tuckerman, and M. Parrinello. *CMPD version 3.4.1*. Max-Planck-Institut für Festkörperforschung and IBM Zürich Research Laboratory, 1995-1999.
16. Oliver Hahn, Damian A. Mooney, Florian Müller-Plathe, and Kurt Kremer. A new mechanism for penetrant diffusion in amorphous polymers: Molecular dynamics simulations of phenol diffusion in bisphenol-A-polycarbonate. *J. Chem. Phys.*, 111(13):6061–6068, 1999.

17. Kurt Kremer and Gary S. Grest. Dynamics of entangled linear polymer melts: A molecular-dynamics simulation. *J. Chem. Phys.*, 92(8):5057–5086, 1990.
18. Jeremy Q. Broughton, Farid. F. Abramam, Noam Bernstein, and Efthimios Kaxiras. Concurrent coupling of length scales: Methodology and application. *Phys. Rev. B*, 60(4):2391–2403, 1999.
19. M. Eichinger, P. Tavan, J. Hutter, and M. Parrinello. A hybrid method for solutes in complex solvents: Density functional theory combined with empirical force fields. *J. Chem. Phys.*, 110(21):10452–10467, 1999.
20. H. Stassen. On the pair potential of liquid methane. *J. Molec. Str. (Theochem)*, 464:107–119, 1999.
21. T. Soddemann, B. Dünweg, and K. Kremer. Dissipative particle dynamics: A useful thermostat for equilibrium and nonequilibrium molecular dynamics simulations. *Phys. Rev. E*, 68(4):046702, 2003.
22. P. Nikunen, M. Karttunen, and I. Vattulainen. How would you integrate the equations of motion in dissipative particle dynamics simulations? *Comput. Phys. Comm.*, 153(3):407–423, 2003.
23. B. Saager and J. Fischer. Predictive power of effective intermolecular pair potentials - md simulation results for methane up to 1000 mpa. *Fluid Phase Equilbr.*, 57(1-2):35–46, 1990.
24. S. Murad and K. E. Gubbins. Molecular dynamics simulation of methane using a singularity-free algorithm. In *Computer Modelling of Matter*, pages 62–71, Washington, D. C., 1978. American Chemical Society.

Liquid Crystal Systems

Michael P. Allen

Centre for Scientific Computing and Department of Physics,
University of Warwick, Coventry CV4 7AL, United Kingdom
E-mail: m.p.allen@warwick.ac.uk

This chapter gives an overview of Monte Carlo and molecular dynamics simulations of liquid crystals. Over the last fifteen years, most effort has focused on the development of methods for measuring, by simulation, the essential defining properties of liquid crystals: the phase behaviour, elastic constants, and surface anchoring coefficients. For this purpose, simple molecular models which reflect the essential physics are sufficient. The results of these simulations are frequently compared with theoretical predictions based on Oseen-Frank orientational elasticity, the Landau-de Gennes free energy, or density functionals. More recently, simulations using simple models have been extended to systems involving colloids and defects. Some references will also be made to the development of simulations using more realistic molecular models, which allow comparison with experiment.

1 Introduction

The computer simulation of liquid crystals is a rapidly expanding field: many of the techniques needed to measure the properties of interest, and to study the relevant phase transitions, are still being developed. In principle, computer simulation combined with liquid-state theories should give us an insight into the link between molecular structure and liquid crystalline behaviour. In practice, this approach complements the very well established experimental method of synthesizing and testing large numbers of compounds. The sensitivity of liquid crystal properties to molecular details means that simulation faces a tough challenge. Nonetheless, the first elements of a coherent picture are beginning to emerge from simulations, and the prospects look good for further progress. Onsager's theory is an early example of the now-popular density-functional theories of the liquid state: such theories directly link molecular properties with bulk-phase behaviour. Computer simulation can make a valuable contribution by testing theories of this kind, perhaps pointing the way to improving them. The first part of this chapter gives examples of this kind of study.

Continuum models of liquid crystals, based on phenomenological elastic constants and hydrodynamic transport coefficients, have been outstandingly successful in modelling behaviour on the length and time scales of most interest to device manufacturers. Simulation can contribute by providing methods of calculating the relevant coefficients for a given molecular model. Moreover, interest is developing in molecular-scale effects, such as behaviour near surfaces and defects, which cannot be modelled properly in the continuum picture. To bridge the gap, it is necessary to simulate very large systems, and some recent work in this area will be described in the second part of this chapter.

2 Molecular Models

Liquid crystals may be modelled at various levels of detail. Atomistic simulations, based on empirical potentials^{1,2} and/or *ab initio* calculations³⁻⁵, are intended to calculate the

properties of specific molecular systems, but current computer power limits the sample sizes to a few thousand molecules with simulation times up to a few tens of nanoseconds. By using more coarse-grained models, in which the molecule is represented by a single non-spherical unit, these figures may be extended by an order of magnitude or more, at the cost of blurring the connection between molecular structure and liquid crystalline properties. Liquid crystalline properties and phase behaviour are extremely sensitive to details of molecular structure and interactions, so it is a great challenge for either simulation or theory to contribute to the design and synthesis of new materials. Moreover, interest in the liquid crystal community is increasingly focusing on highly inhomogeneous situations (behaviour near interfaces and solid surfaces, domain walls and defects, for example) so some consideration of atomic-scale effects is highly desirable.

Most liquid crystals of technological interest are formed by highly elongated molecules. In our simulation work we use a range of elementary molecular models. Typical examples are *ellipsoids of revolution* of length A , width B , elongation $e = A/B$; and *spherocylinders* of overall length $L + D$, width D . They are sufficient to describe the basic physics of liquid crystals; they all have a *nematic* phase, which will be the main object of study here. Given these basic shapes, the interaction potentials are divided into ‘hard’ and ‘soft’ classes: for *hard* potentials, the energy is infinite when two molecules overlap, and zero otherwise; for *soft* potentials, the energy varies continuously with relative position and orientation. We may improve these models by including attractive forces^{6–10}, biaxiality^{11,12}, nonlinearity¹³, flexibility^{14,15}, dipolar forces¹⁶ and hydrogen bonds¹⁷. Space permits us to touch on only one or two of these aspects in the present chapter.

Basic simulation methods are described in other chapters^{18,19} and elsewhere^{20,21}, so only a brief summary of particular simulation strategies will be given in the following sections.

3 Theoretical Approaches

We shall be comparing some simulation results with theoretical predictions that may be expressed by writing down an approximate free energy *functional*. The functional may be assumed to depend on simple functions of particle positions and orientations, for instance the local single-particle density, the orientational order tensor, the direction of orientational ordering. All of these descriptions are coarse-grained: specification of the state of the system does not require knowledge of all of the N -particle microscopic coordinates. Instead, we must input the dependence of a handful of functions on the spatial and orientational coordinates, and minimize the free energy with respect to all possible variations of these functions.

The elastic theory due to Oseen and Frank^{22,23} is based on writing part of the free energy $\mathcal{F}$ as a functional of the position-dependent liquid crystal director $\mathbf{n}(\mathbf{r})$:

$$\mathcal{F}_{\text{OF}}[\mathbf{n}(\mathbf{r})] = \int_V d^3\mathbf{r} f_b(\mathbf{r}) + \int_S d^2\mathbf{s} f_s(\mathbf{s}) \quad (1a)$$

$$f_b(\mathbf{r}) = \frac{1}{2}K_{11}(\nabla \cdot \mathbf{n})^2 + \frac{1}{2}K_{22}(\mathbf{n} \cdot \nabla \times \mathbf{n})^2 + \frac{1}{2}K_{33}(\mathbf{n} \times \nabla \times \mathbf{n})^2 \quad (1b)$$

$$f_s(\mathbf{s}) = \frac{1}{2}W_\theta \sin^2 \theta \quad \text{where } \cos \theta = \mathbf{n} \cdot \mathbf{n}_s. \quad (1c)$$

where S is the surface bounding the sample volume V . There are two classes of terms: an integral, over V , of a bulk free energy density which depends on the splay (K_{11}), twist

(K_{22}) and bend (K_{33}) elastic constants; and an integral, over S , of a surface free energy density, which has been written here simply as a *surface anchoring* term of strength W_θ depending on the angle θ between $\mathbf{n}$ and the preferred anchoring orientation $\mathbf{n}_s$ at the surface. Omitted here are the surface-like elastic terms K_{13} and K_{24} which have been of recent experimental and theoretical interest^{24–26}.

The Frank free energy describes a nematic liquid crystal in regions where the *direction* of ordering is the only important variable, and where, moreover, it varies smoothly with position (hence it takes the form of a gradient expansion). A theory which allows the *degree* as well as the direction of ordering to vary is due to Landau and de Gennes²⁷, and is based on the 2nd-rank order tensor $\mathbf{Q}(\mathbf{r})$:

$$\mathcal{F}_{\text{LdG}}[\mathbf{Q}(\mathbf{r})] = \int_V d^3\mathbf{r} f_b(\mathbf{r}) + \int_S d^2\mathbf{s} f_s(s) \quad (2a)$$

$$f_b(\mathbf{r}) = \kappa |\nabla \mathbf{Q}|^2 + a \text{Tr} [\mathbf{Q}^2] - b \text{Tr} [\mathbf{Q}^3] + c \{ \text{Tr} [\mathbf{Q}^2] \}^2 \quad (2b)$$

$$f_s(s) = w \text{Tr} [\mathbf{Q} - \mathbf{Q}_s]^2 \quad (2c)$$

In the bulk free energy density, we see a one-constant representation of elastic behaviour (the squared gradient term), plus the leading terms in a free energy expansion, which would give a phenomenological description of the bulk I-N transition in the absence of gradients. To this may be added a surface anchoring term, the simplest form of which is exemplified by Eq. (2c), a square term involving the preferred order tensor $\mathbf{Q}_s$ at the surface.

The final theory to be considered here is that of Onsager²⁸ in which the free energy is written as a functional of the single-particle density $\varrho(\mathbf{r}, \hat{\mathbf{a}})$:

$$\begin{aligned} \mathcal{F}_{\text{Ons}}[\varrho(\mathbf{r}, \hat{\mathbf{a}})] = & k_B T \int d^3\mathbf{r} d^2\hat{\mathbf{a}} \varrho(\mathbf{r}, \hat{\mathbf{a}}) [\ln \varrho(\mathbf{r}, \hat{\mathbf{a}}) \Lambda^3 - 1 + u(\mathbf{r}, \hat{\mathbf{a}})/k_B T] \\ & - \frac{1}{2} k_B T \iint d^3\mathbf{r}_A d^2\hat{\mathbf{a}} d^3\mathbf{r}_B d^2\hat{\mathbf{b}} \varrho(\mathbf{r}_A, \hat{\mathbf{a}}) \varrho(\mathbf{r}_B, \hat{\mathbf{b}}) f(\mathbf{r}_{AB}, \hat{\mathbf{a}}, \hat{\mathbf{b}}). \end{aligned} \quad (3)$$

Here, $\hat{\mathbf{a}}, \hat{\mathbf{b}}$ are the molecular orientation vectors; we assume uniaxial particles for convenience. The details of molecular shape and structure enter through the Mayer function $f(\mathbf{r}_{AB}, \hat{\mathbf{a}}, \hat{\mathbf{b}})$ which may be related directly to the intermolecular potential. Onsager's theory is a fore-runner of what we now call density functional theories. These have proved extremely successful in recent years when applied to simple atomic fluids^{29–31}, and have been used to explore liquid crystal phase diagrams³². In general, the functional $\mathcal{F}[\varrho]$ is unknown and nonlocal; the above expression is valid for low densities, but note that it is not restricted to smooth variations of $\varrho(\mathbf{r}, \hat{\mathbf{a}})$. An empirical improvement of Onsager's theory, due to Parsons and Lee^{33–35}, introduces a density-dependent prefactor into the two-body term, and extends its range of validity to higher densities.

4 Liquid Crystal Phase Diagrams

4.1 Hard-Particle Systems

It is twenty years since the phase diagram of the hard ellipsoid system was determined by Frenkel and Mulder³⁶ and the role of their follow-up paper³⁷ in shaping the future development of the field is concisely summarised by Evans³⁸. Their results are shown

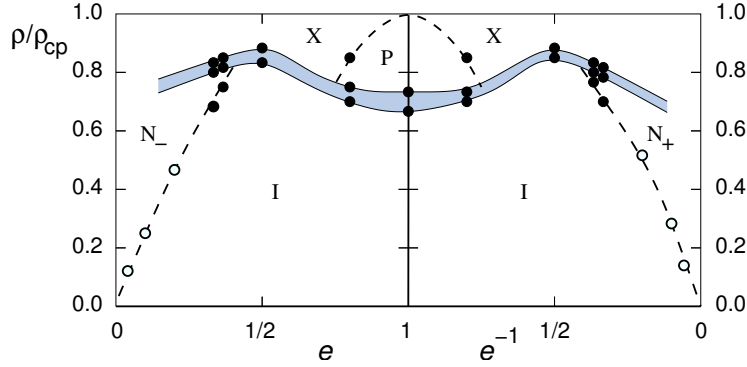


Figure 1. Schematic hard ellipsoid phase diagram as a function of elongation e (on a reciprocal scale for $e > 1$), and density ρ (as a fraction of the close-packed density ρ_{cp}). The isotropic (I), crystalline (X), plastic crystalline (P), nematic (N_+) and discotic (N_-) phases are indicated. The solid-fluid coexistence region is shaded, and the I-N and P-X transitions indicated by dashed lines. The state points studied in Refs. 36, 37 are indicated by filled circles.

schematically in Fig. 1. This work involved determination of free energies and chemical potentials by thermodynamic integration with respect to an external orienting field. Recall, to locate a phase transition, it is necessary to determine the state points which satisfy that conditions for thermodynamic equilibrium: $T_1 = T_2$, $P_1 = P_2$, $\mu_1 = \mu_2$. For particles interacting only with infinitely repulsive potentials, the temperature is not a significant thermodynamic variable, and the first equality is always satisfied. The last equation is the most difficult to determine.

Subsequent studies of ellipsoids at higher elongations³⁹ involved a range of techniques: (a) test particle insertion plus thermodynamic integration; (b) simulation in the Gibbs ensemble; (c) thermodynamic integration with respect to changes in particle shape; (d) tracing phase coexistence boundaries using Clapeyron-like equations; and it is instructive to consider these in more detail.

The chemical potential at a point (T, P) in each phase may be found by Widom test particle insertion. Since $\mu^{ex} = F_{N+1}^{ex} - F_N^{ex}$,

$$\beta\mu^{ex} = -\ln \langle \exp(-\beta\mathcal{U}_{test}) \rangle$$

where $\mathcal{U}_{test}$ is the potential energy of interaction of a randomly inserted test particle with the N particles in the system (in this case, $\mathcal{U}_{test}$ is either infinity or zero, so the exponential is either zero or one). Then one may use thermodynamic integration

$$\mu(P) = \mu(P_0) + \int_{P_0}^P \rho^{-1} dP$$

to solve $\mu_1(P_1) = \mu_2(P_2)$.

Automatic equilibration of two phases, with each phase in its own simulation box, may be achieved using the Gibbs ensemble method of Panagiotopoulos^{40,41}. Briefly, this technique involves standard Monte Carlo moves in each box; volume exchange moves between boxes, which guarantee $P_1 = P_2$; and particle transfer moves arranged to ensure

$\mu_1 = \mu_2$. Both the Widom method and the Gibbs ensemble method work well at low-to-moderate densities, and this is the case for the I-N transition for highly elongated, $e = 20$, hard ellipsoids, as demonstrated in Fig. 2.

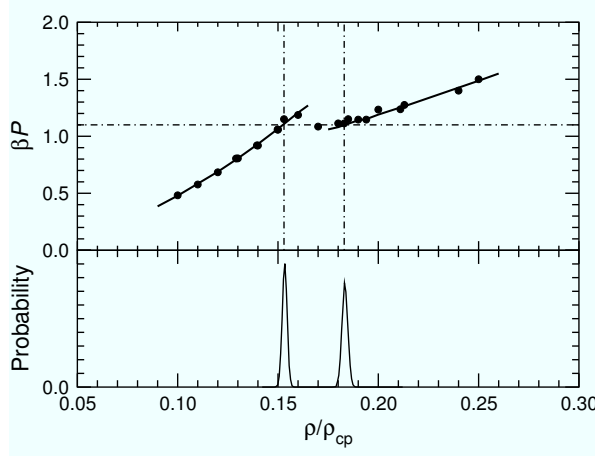


Figure 2. I-N coexistence for $e = 20$ ellipsoids. The data points with solid lines are the equations of state $P(\rho)$ for isotropic and nematic phases, with the coexistence value P_{NI} determined by Widom test-particle insertion plus thermodynamic integration, indicated by dot-dashed lines. Below this is the probability distribution function for density ρ averaged over both boxes in a Gibbs simulation.

For dense fluids, both the Widom method and the Gibbs ensemble become less reliable, due to the low acceptance rates for particle insertion. One solution is to use an *expanded ensemble* in which N -particle and $(N + 1)$ -particle systems are linked by intermediate fractional-particle states^{42,43}. One particle is scaled in size by a parameter κ , taking values $0 \leq \kappa \leq 1$. In the Monte Carlo procedure, we allow transitions between species $\kappa \rightleftharpoons \kappa'$. We determine the probability histogram $\mathcal{P}(\kappa)$, and hence the relative free energies of species $\mathcal{F}(\kappa)$; then $\mathcal{F}(1) - \mathcal{F}(0)$ gives μ^{ex} . To ensure uniform sampling of species, we apply an (iteratively refined) weighting function $\mathcal{W}(\kappa)$ to the κ -moves. It is essential that the scaled particle samples fluid configurations efficiently. To assist, we introduce two additional types of move which dramatically relocate it. Firstly, we attempt to move the scaled particle to a completely random position, as in the (random-insertion) Widom method. Such moves are accepted with high probability when κ is small. Secondly, we attempt to exchange the scaled particle with full-size particles. This works best when κ is large. Details are given elsewhere³⁹.

An essential part of a study of this kind is to determine how the transition pressure changes as we slowly vary molecular shape, or some other parameter. The idea is to avoid computing chemical potentials from scratch, for each case. A useful approach, due to Kofke⁴⁴, is to derive and then use a variant of the Clapeyron equation. Suppose we characterize the molecular shape by a parameter λ ; here we choose $\lambda = \ln a$ (a = length) or $\ln b$ (b = width) for convenience. Then we define a thermodynamic variable $\Gamma = (\partial\mu/\partial\lambda)_{TP}$ conjugate to λ ; this means that we may write $d\mu = \Gamma d\lambda + v dP$. It follows

that

$$\frac{dP}{d\lambda} = -\frac{\Delta\Gamma}{\Delta v} \quad \text{or} \quad \frac{d \ln P}{d\lambda} = -\frac{\Delta\Gamma}{P\Delta v}.$$

Kofke's method consists of evaluating the terms on the right, and integrating step-by-step with respect to λ along the coexistence curve. Applying these techniques to the hard ellipsoid system³⁹, we were able to show the extent of applicability of both Onsager and Parsons theories, and the expected decrease in strength of transition on decreasing elongation, as shown in Fig. 3. The effects of molecular biaxiality on the phase diagram have recently been investigated in the same way using a hard spheroid model¹².

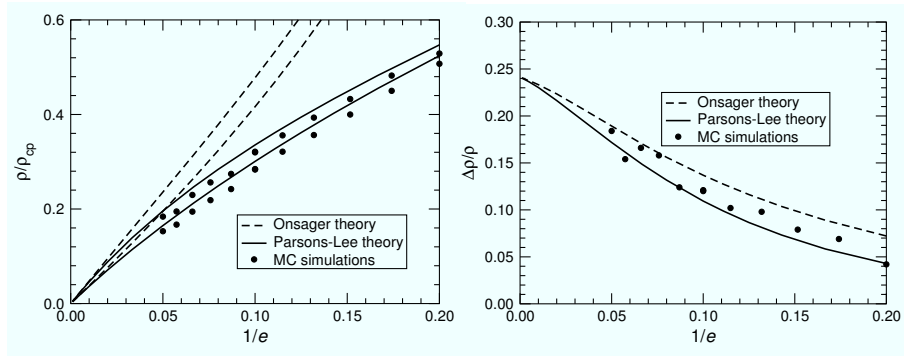


Figure 3. Coexistence densities (left) and fractional density change (right) at the I-N transition vs elongation for hard ellipsoids of revolution. For details see Ref. 39.

4.2 Rod-Plate Mixtures

Mixtures of rod-like and plate-like molecules may form isotropic, uniaxial and biaxial phases as shown in Fig. 4. The biaxial phase competes with demixing into two uniaxial phases: this is an example of phase separation due to solely repulsive interactions. We have studied mixtures of ellipsoids of the same molecular volume, and two conjugate pairs of shapes: $e = 15, 1/15$ and $e = 20, 1/20$ to investigate this competition. Once again it was possible to compare with Onsager theory and the Parsons-Lee rescaling. The phase transitions were located by performing Gibbs ensemble simulations, and observing the behaviour of order parameters characterizing both phases; full details are given elsewhere⁴⁵. The phase diagram for the $e = 15, 1/15$ system is shown in Fig. 5 (that for $e = 20, 1/20$ is similar). A striking feature is the asymmetric line separating the biaxial phase from the two-phase demixing region. For these conjugate shapes, a symmetry of the second virial coefficient under the exchange $e \leftrightarrow 1/e$ means that the phase diagram predicted by the Onsager or Parsons theory is symmetric about the 50:50 composition. Most of the phase boundaries exhibit this symmetry, approximately, but the dividing line between biaxial and two-phase regions is clearly very sensitive to three-body, or perhaps higher, interactions.

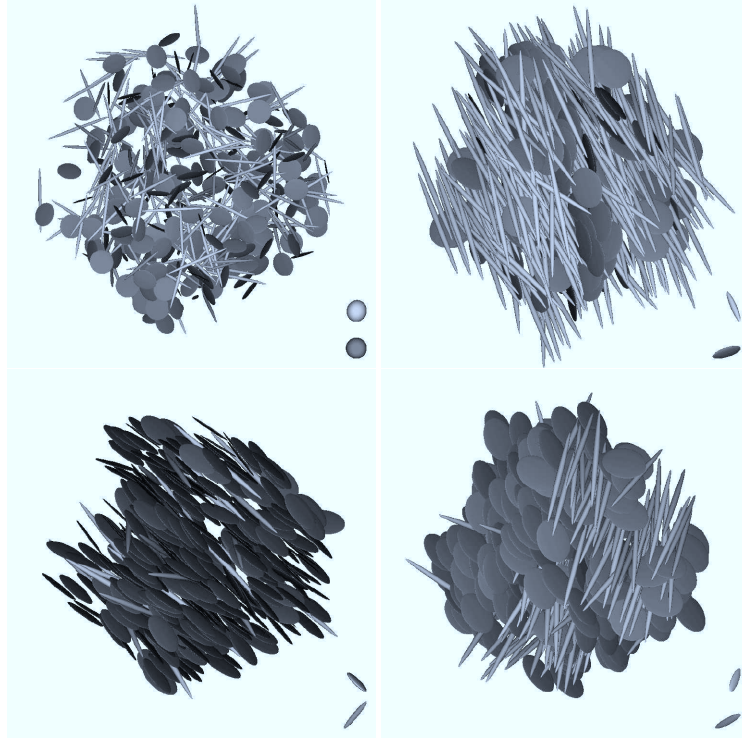


Figure 4. Rod and plate configurations for $e = 20, 1/20$ ellipsoid mixtures, showing isotropic, uniaxial nematic, discotic and biaxial phases. The biaxial configuration also shows some demixing. At the bottom right corner of each configuration, two spheroids give a visual representation of the orientational ordering of rods (upper) and plates (lower).

4.3 The Gay-Berne Model

The introduction of attractive forces into the molecular model makes the phase diagram a little more realistic, but more complicated. The well-known Gay-Berne potential has become a standard model for the study of liquid crystalline phases. The molecular pair potential

$$v_{AB} = 4\varepsilon(A, B) \left[\left(\frac{\sigma_0}{r_{AB} - \sigma(A, B) + \sigma_0} \right)^{12} - \left(\frac{\sigma_0}{r_{AB} - \sigma(A, B) + \sigma_0} \right)^6 \right]$$

depends upon the molecular axis vectors $\hat{\mathbf{a}}$ and $\hat{\mathbf{b}}$, and on the direction $\hat{\mathbf{r}}_{AB}$ and magnitude r of the centre-centre vector $\mathbf{r}_{AB} = \mathbf{r}_A - \mathbf{r}_B$. The parameter σ_0 determines the smallest molecular diameter and there are two orientation-dependent quantities: a diameter $\sigma(A, B) = \sigma(\hat{\mathbf{r}}_{AB}, \hat{\mathbf{a}}, \hat{\mathbf{b}} \mid \kappa, \kappa', \mu, \nu)$ and an energy $\varepsilon(A, B) = \varepsilon(\hat{\mathbf{r}}_{AB}, \hat{\mathbf{a}}, \hat{\mathbf{b}} \mid \kappa, \kappa', \mu, \nu)$. Each quantity depends in a complicated way (not given here) on the shape anisotropy parameter κ , the energy anisotropy parameter κ' , and two exponents μ, ν . The original ‘reference’ model has $\mu = 2, \nu = 1, \kappa = 3, \kappa' = 5$.

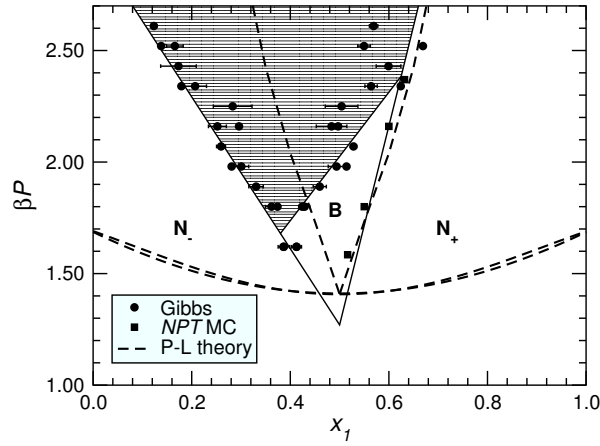


Figure 5. Phase diagram for $e = 15, 1/15$ rod-plate mixtures in the pressure-composition plane: x_1 is the mole fraction of rods. The points are simulation results, the solid lines mark approximate phase boundaries, the two-phase region is shaded, and the dashed-lines are the predictions of Parsons-Lee theory. For further details see Ref. 45.

The effects of varying the attractive anisotropy parameter κ' in the range $1 \leq \kappa' \leq 25$, with fixed $\kappa = 3$, have been reported in detail^{46,9}. Perhaps the most interesting effect is the growth of the liquid-vapour coexistence envelope on reducing κ' , as illustrated in Fig. 6. The envelope was determined by Gibbs ensemble simulation, but this became

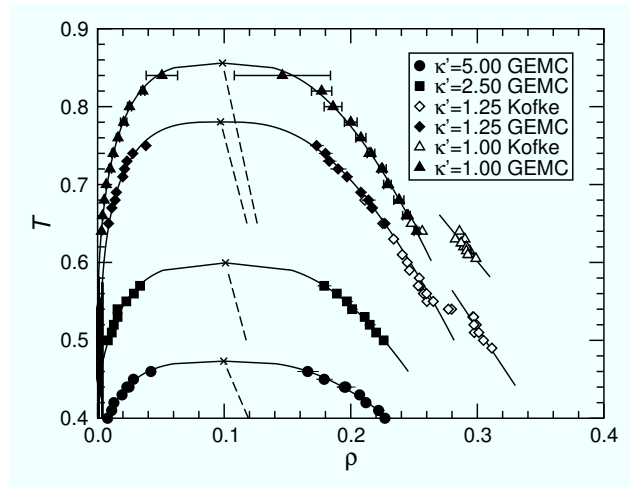


Figure 6. The effects of varying the anisotropy of the attractive forces in the Gay-Berne potential on the liquid-vapour coexistence envelope. Shown are the results of Gibbs simulations and Kofke (Gibbs-Duhem) integration. There is an I-N transition on the liquid branch for $\kappa' = 1.00$ and 1.25.

less efficient at lower temperatures as the liquid density became higher; just at this point,

evidence emerged of an I-N transition on the liquid side, raising the prospect of being able to simulate a nematic-vapour interface. The coexisting vapour density is very low at these temperatures, so a constant-pressure simulation with $P = 0$ provides a good guide to the behaviour of the liquid with changing temperature; however, it was also possible to conduct thermodynamic integration using the Kofke method, to determine both vapour and liquid densities at coexistence. Studies of the nematic-vapour interface for this model have been undertaken by two groups^{46–48}.

Values of the elongation κ in the range $3 \leq \kappa \leq 4$ with fixed $\kappa' = 5$ have also been studied¹⁰. The effects of increasing κ from 3 to 4 are indicated in the phase diagrams of Fig. 7. The liquid-vapour envelope contracts and disappears, while the layered smectic-A phase becomes stable, but bounded at high and low T . Further details, including a discussion of methods for identifying the various smectic phases and distinguishing them from the crystalline state, may be found in Ref. 10.

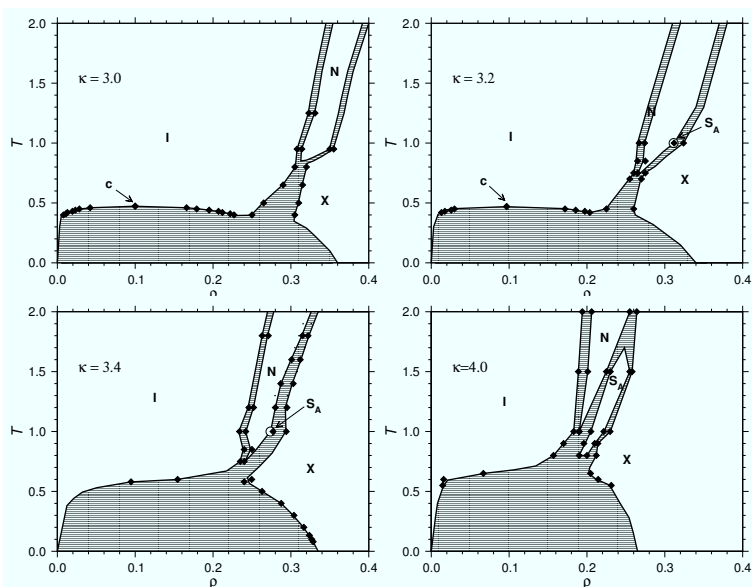


Figure 7. Schematic phase diagrams for Gay-Berne systems at various elongations. For more details see Ref. 10.

5 Dynamics of Liquid Crystals

An early and surprising result of molecular dynamics simulation of hard ellipsoids, was that for some elongations, in the vicinity of the isotropic-nematic transition, the component of the single-particle diffusion coefficient parallel to the director would *increase* with increasing density⁴⁹. Subsequently, it proved possible to correlate this with the rapidly changing order parameter near the transition point⁵⁰: roughly speaking, as the system becomes more ordered, diffusion of a molecule along its length becomes less and less hindered by collisions with its neighbours.

Collective dynamics, and the growth of large, highly-ordered regions on the approach to the isotropic-nematic transition temperature T_{NI} are also of interest. From the isotropic side, the correlation length for second-rank orientational fluctuations ξ_2 diverges and collective orientational dynamics become slow, rather like density fluctuations near the liquid-vapour critical point. The extrapolated divergence temperature T^* is very slightly lower than the thermodynamic transition temperature, and the transition is weakly first-order rather than being continuous. Although a proper system-size dependent study of this phase transition would be very expensive, and is only really feasible for a lattice model^{51,52}, there is some interest in observing pre-translational fluctuations using a molecular model and comparing with simple theoretical predictions. We studied a system of $N = 8000$ molecules, interacting with a version of the Gay-Berne potential⁸, using typical run lengths ~ 20 ns. For temperatures just above the transition temperature we have studied structural correlations and dynamical correlation functions⁵³. The static pair correlation function, $h(1, 2)$, and the direct correlation function, $c(1, 2)$, may be expanded in a set of rotationally invariant functions⁵⁴ to obtain some insight into the slowing down. Our results, shown in Fig. 8 suggest that, as expected, the appropriate components of $c(1, 2)$ remain short-ranged (at least, for the temperatures investigated here) while the corresponding components of $h(1, 2)$ develop long-ranged correlations.

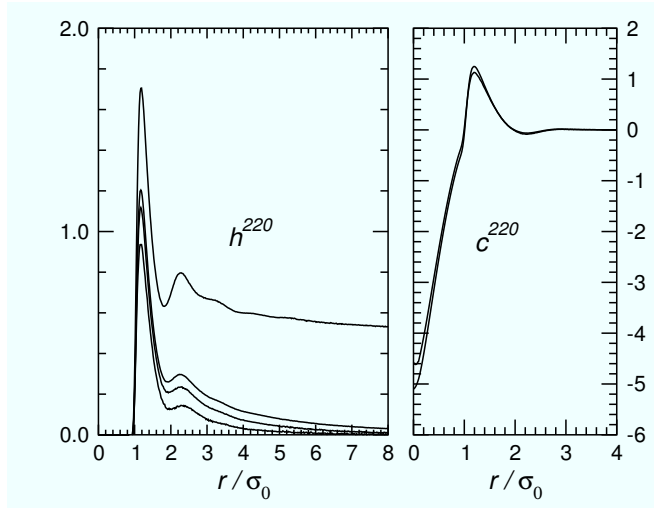


Figure 8. On the left we show the component of the total correlation function $h^{220}(r)$ which shows pretranslational growth of orientational correlations, approaching the I-N transition at $T = 4.0, 3.6, 3.5$ and at $T = 3.45$, below the transition, showing long-ranged order. On the right we show the corresponding component of the direct correlation function, $c^{220}(r)$, approaching the transition at $T = 4.0, 3.5$.

To study the collective dynamics we defined a variable

$$Q_{\alpha\beta}(\mathbf{k}, t) = \sum_i \left(\frac{3}{2} u_{i\alpha}(t) u_{i\beta}(t) - \frac{1}{2} \delta_{\alpha\beta} \right) e^{i\mathbf{k} \cdot \mathbf{r}_i(t)} \quad \alpha, \beta = x, y, z \quad (4)$$

and measured the time correlation function

$$C(k, t) = \sum_{\alpha\beta} \langle Q_{\alpha\beta}(-\mathbf{k}, 0) Q_{\alpha\beta}(\mathbf{k}, t) \rangle \propto \sum_{ij} \left\langle P_2(\hat{\mathbf{a}}_i(0) \cdot \hat{\mathbf{a}}_j(t)) e^{i\mathbf{k} \cdot (\mathbf{r}_i(0) - \mathbf{r}_j(t))} \right\rangle$$

From the simulation results we observed roughly exponential decay

$$C(k, t) = A(k) \exp\{-t/\tau(k)\}$$

with $\tau(k) \sim k^{-2}$. At $k = 0$, we found $\tau^{-1} \propto \xi_2^{-2} \propto (T - T^*)$, in agreement with a simple Landau-de Gennes description²⁷. For further details see Ref. 53.

6 Elastic Properties

The simulation of truly large-scale phenomena associated with elastic deformations of the director field requires parallel computing resources, and the programs used for the work described in the following sections were developed as a consortium project within the U.K. High Performance Computing Initiative. The parallel MD code developed in this project is described elsewhere⁵⁵.

6.1 Bulk Elastic Constants

Long-wavelength deformations of the director field $\mathbf{n}(\mathbf{r})$ are described by the Frank free energy (1). The question to be addressed here is ‘Can simulation provide values of the coefficients in this expression?’.

In the nematic phase, for sufficiently long wavelengths that the elastic picture holds, equilibrium thermal fluctuations of the Fourier-transformed orientation density $\hat{Q}_{\alpha\beta}(\mathbf{k})$ (defined by Eq. (4)) are determined by the elastic constants. It is convenient to define the following:

$$\begin{aligned} \mathcal{W}_{13}(k_1, k_3) &\propto \left\langle \hat{Q}_{13}(\mathbf{k}) \hat{Q}_{13}(-\mathbf{k}) \right\rangle^{-1} \equiv \langle |\hat{Q}_{13}(\mathbf{k})|^2 \rangle^{-1} \propto K_1 k_1^2 + K_3 k_3^2 \\ \mathcal{W}_{23}(k_1, k_3) &\propto \left\langle \hat{Q}_{23}(\mathbf{k}) \hat{Q}_{23}(-\mathbf{k}) \right\rangle^{-1} \equiv \langle |\hat{Q}_{23}(\mathbf{k})|^2 \rangle^{-1} \propto K_2 k_1^2 + K_3 k_3^2 \end{aligned}$$

Here, the director is chosen to lie in the 3-direction; $\mathbf{k} = (k_1, 0, k_3)$ is in the 1-3 plane. These quantities may be determined in a standard computer simulation, and a fit made to $\mathcal{W}_{13}$ and $\mathcal{W}_{23}$ as functions of k_1^2 and k_3^2 . Taking the low- k limit is the crucial part of this process, so it is essential to have a large enough simulation box size L to guarantee the accuracy of the extrapolation. It is convenient to constrain the 123 axis system to coincide with the xyz coordinate system of the simulation box, otherwise the components k_1, k_3 will change as the director drifts; this can be done using Lagrangian constraints in molecular dynamics. We have studied⁵⁶ the Gay-Berne potential^{57,8}. Slow fluctuations of the long-wavelength modes necessitated long simulation runs, while system sizes up to $N = 8000$ particles were used to guarantee the correctness of the low- k extrapolation. Typical slices through the fitted surface, with the corresponding data points, are shown in Fig. 9.

The key conclusions of this work were that elastic constants can indeed be measured quite accurately by this method, and that the absolute magnitudes of the elastic constants,

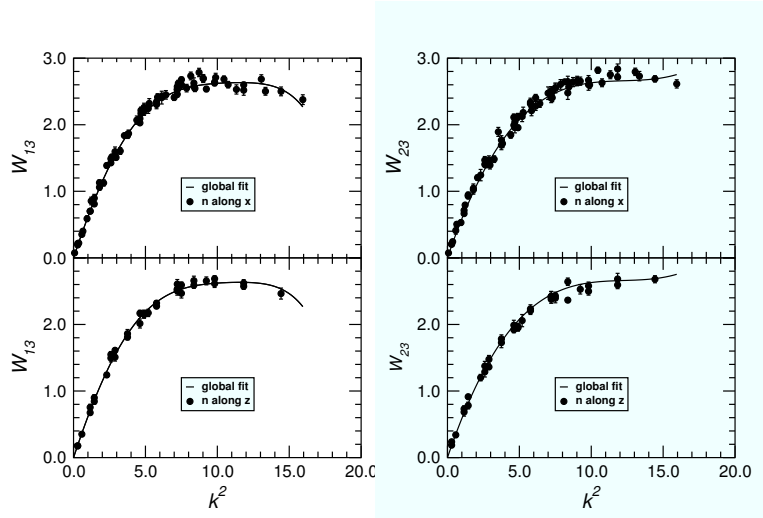


Figure 9. Splay (left) and twist (right) fluctuations as functions of wavelength for the Gay-Berne model with $\mu = 2$, $\nu = 1$, $\kappa = 3$, $\kappa' = 5$ at $\rho = 0.32$, $T = 0.90$. Here the wave-vector component along the director k_3 is taken to be zero, and the inverse fluctuations quantities $\mathcal{W}_{13}$ and $\mathcal{W}_{23}$ are plotted as functions of k_1^2 . The solid lines show sections through the globally-fitted surface. For more details see Ref. 56.

and their ratios, are consistent with those expected for rod-like molecules. In particular the bend elastic constant K_{33} was typically rather larger than the other two, and the values were all reasonably close to those found for hard ellipsoids of similar shape, at comparable densities.

6.2 Surface Anchoring Strengths

Simulation and theory have been used^{58,59} to explore the anchoring of a liquid crystal at a solid surface. To begin with, we consider a simple ‘experimental’ geometry, in which the liquid crystal is confined between two walls: at one wall an orienting field is applied, and the response of the director alignment across the cell is measured, with the aim of calculating the strength of anchoring at the opposite wall. In keeping with the simple molecular model, namely hard ellipsoids of elongation $A/B = 15$, the wall-molecule interaction is a very idealized hard potential, defined simply by the condition that the centres of the molecules may not pass through it. Purely entropic considerations, having their origin in the way molecules pack together, generate homeotropic (normal) anchoring. Experimentally, it is quite common to anchor liquid crystals at a treated interface, in which a layer of organic molecules is adsorbed on the solid surface. The liquid crystal molecules partially penetrate the organic layer: the effect of this is modelled (albeit in a simple way) by the chosen boundary condition.

The wall anchoring coefficient W measures the strength with which the surface resists director deformation away from the preferred direction. The approach used in the simulation to measure this is illustrated in Fig. 10. The key quantity turns out to be the *extrapolation length* $\lambda = K_{33}/W$, measurable by fitting director profiles to the elastic theory. For $A/B = 15$ ellipsoids, λ is of the order of one molecular length, as shown in Fig. 11.

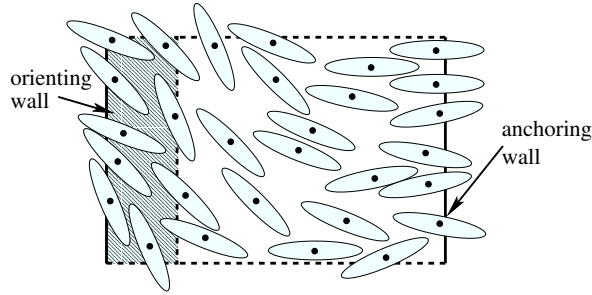


Figure 10. Schematic picture of simulation to measure wall anchoring coefficient W . At the left wall, a strong, but short-ranged, orienting field is applied in the shaded region to fix the nematic director in a chosen direction. The director profile is measured across the cell: in the elastic theory, it is determined by the surface anchoring coefficient at the right wall, W , and the bulk splay and bend elastic constants K_{11} , K_{33}

Direct minimization of the Onsager free energy functional, with essentially no adjustable parameters, reproduces these director profiles very well, *even in the wall regions where the elastic theory is inaccurate*. In the above treatment, it was assumed that the positions of

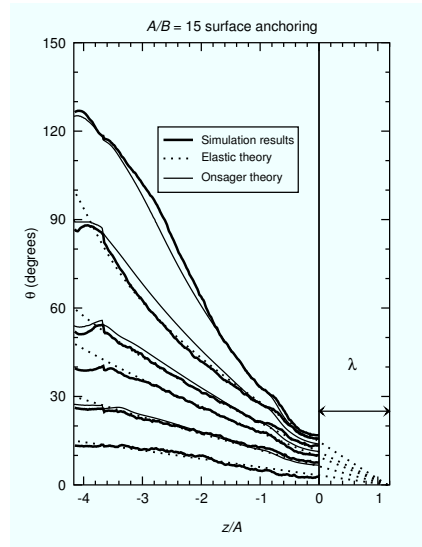


Figure 11. Simulation results, elastic theory fits, and predictions of Onsager theory to determine the extrapolation length λ , and hence the surface anchoring coefficient W . Several different anchoring angles are imposed at the left wall. The curvature in the profiles is because the deformation is predominantly splay as $\theta \rightarrow 90^\circ$ and predominantly bend as $\theta \rightarrow 0^\circ$: the two elastic constants are different (After Ref. 58).

the elastic-theory boundary conditions coincide with the hard potential walls. There is no reason, in principle, why this should be so, although we would expect the difference in positions to be a molecular length rather than a mesoscopic length. This difference turns

out to be detectable when one studies fluctuations of the director $\mathbf{n}(\mathbf{r})$ in slabs of varying thickness between identical walls⁶⁰.

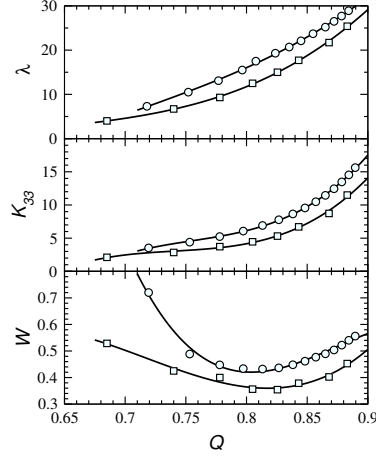


Figure 12. Extrapolation length λ , bend elastic constant K_{33} and anchoring strength W , as functions of order parameter Q . We show the results of Onsager theory (circles) and Monte Carlo simulations (squares) (After Ref. 61).

For this model, the anchoring coefficient varies in an interesting way with density or, equivalently, nematic order parameter. From the available experimental data on thermotropic liquid crystals, one can say that the extrapolation length $\lambda = K_{33}/W$ is inversely proportional to the squared value of the bulk order parameter $\lambda \propto Q^{-2}$. Taking into account that the elastic constant K_{33} is typically proportional to Q^2 , this gives⁶² for the anchoring parameter $W \propto Q^4$. However, our simulations and Onsager theory do not conform to this model. The dependences of λ , K_{33} , and W on the order parameter Q are shown in Fig 12. Rather than decreasing, the extrapolation length λ *increases* with Q , albeit not as strongly as K_{33} . The net result is that $W = K_{33}/\lambda$ has a non-monotonic dependence on Q , first decreasing, then increasing, although the variation is not dramatic. It turns out that this behaviour is connected with the variation of the density profile in the immediate vicinity of the surface, which is often neglected in theoretical treatments. Further details may be found in Ref. 61.

7 Nematic-Isotropic Interface

Now we consider what happens when the nematic and isotropic phases are at coexistence, either with or without the confining effect of walls. The N-I interface is one of the simplest fluid-fluid interfaces and poses some interesting basic questions. The simulation of this interface requires great care: the coexisting densities are typically very similar, and the surface tension will be relatively small, so fluctuations may be important. Our interest here lies in the determination of the properties of the nematic-isotropic interface, including the

surface tension, while taking care to allow all the fluctuations associated with thermal equilibrium. We have studied⁶³ a thick film of nematic phase adsorbed on a solid wall with the same kind of boundary conditions discussed earlier; the system is illustrated schematically in Fig. 13.

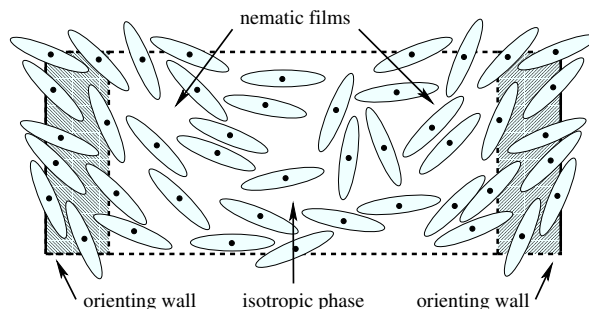


Figure 13. Schematic picture of simulation to investigate the isotropic-nematic interface, with director orientation controlled by anchoring at the confining walls. At the hard walls, strong, but short-ranged, orienting fields are applied to fix the nematic director in a chosen direction. The two wall anchoring conditions may be chosen independently, provided the isotropic region in the middle is sufficiently wide

7.1 Orientational Order and Density Profiles

For an isolated surface, the film thickness is determined by the deviation of the chemical potential from the bulk coexistence value $\mu - \mu_{\text{IN}}$. In a closed, confined system with adsorbed films on both walls, assuming the wall separation is large enough that the two films are essentially independent, the film thickness is determined by the overall density. The phase transitions in this type of system, such as the wetting behaviour, and the nematic-isotropic equivalent of ‘capillary condensation’, are of great interest^{64,65}; but here the focus of attention will be the structure of the nematic film and the interface between the two fluid phases.

We have, once more, studied ellipsoids of molecular elongation $A/B = 15$. Figure 14 gives an example of the order parameter and density profiles obtained by simulation, and compared with Onsager theory predictions, for the case when the orienting field at the hard wall aligns the director in the plane of the wall, i.e. at 90° to the interface normal. The agreement between theory and simulation is good, bearing in mind that Onsager theory is known to overestimate the bulk coexistence densities; indeed, very detailed comparisons of the profile shapes may be made⁶³.

There is some interest in studying even longer molecules⁶⁶. Firstly, it has been predicted using Onsager theory for infinitely long rods⁶⁷ that the number density $\rho(z)$ varies in a significantly non-monotonic way near the I-N interface, for the lowest free energy case of in-plane alignment. There is no evidence of this for $A/B = 15$, in either simulation or theory (see Fig. 14), which raises the possibility that the effect is only seen for higher elongations. Secondly, again for this alignment, the orientational ordering near the interface should be *biaxial*: symmetry is broken in two directions, namely normal to the

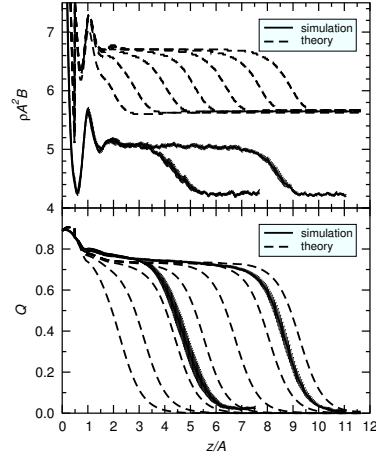


Figure 14. Profiles of number density in reduced units $\rho A^2 B = NA^2 B/V$, and order parameter Q (largest eigenvalue of the order tensor $\mathbf{Q}$), for a nematic film anchored at 90° to the normal at the left wall (in-plane alignment), in equilibrium with the isotropic phase. Simulation results are given for two different film thicknesses (solid line with error bars) and Onsager theory predictions for seven different film thicknesses (dashed lines) (After Ref. 63).

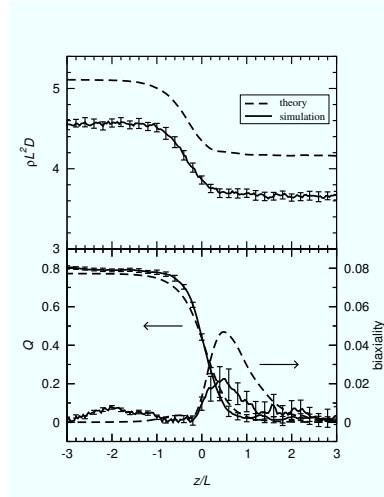


Figure 15. Profiles of reduced number density $\rho L^2 D$, orientational order parameter Q , and orientational biaxiality (on a $\times 10$ expanded scale), in the vicinity of the I-N interface, for $L/D = 50$ soft spherocylinders, determined by simulation (solid lines with error bars) and from Onsager theory (dashed lines) (After Ref. 66).

interface, and in the plane of the interface along the director. However, this biaxiality is extremely small for the $A/B = 15$ elongation described earlier. In Fig. 15 are the relevant profiles for a soft spherocylinder model with length/width = $L/D = 50$, studied by molecular dynamics, using a free-standing nematic film surrounded by isotropic phase in fully periodic boundaries. It can be seen that the density profile changes monotonically,

both in simulation and theory. However, the biaxiality (defined as the difference between the middle and lowest eigenvalues of the order tensor $\mathbf{Q}$) is detectable in the simulation, and is overestimated by the Onsager theory by about a factor of 2.

7.2 Surface Tension

One of the most important characteristics of an interface is the surface tension or surface free energy γ . For the I-N interface, the variation of this quantity with director angle gives, essentially, the anchoring potential at the interface. It will dictate the shape of droplets of one phase suspended in the other. Determining this quantity for an equilibrium I-N interface is a significant challenge for computer simulation. For a planar interface normal to the z -direction, a microscopic expression for surface tension γ is:

$$\gamma = \int_{-\infty}^{\infty} dz [P_{\parallel} - P_{\perp}(z)] \equiv \int_{-\infty}^{\infty} dz \pi(z).$$

Here, $P_{\parallel} = P_{zz} = P$ is the normal component of the pressure tensor, which is independent of z throughout the system; $P_{\perp}(z) = P_{xx}(z) = P_{yy}(z)$ is the transverse component of the pressure tensor, and this varies with z near interfaces. Far from the interface, $P_{\perp}(z) = P_{\parallel} = P$. In the above formulae, the microscopic definition of $P_{\perp}(z)$ is not unique, although the value of the surface tension does not depend on the choice^{68–70} and in the following it turns out that the differences in the profiles are quite small⁷¹.

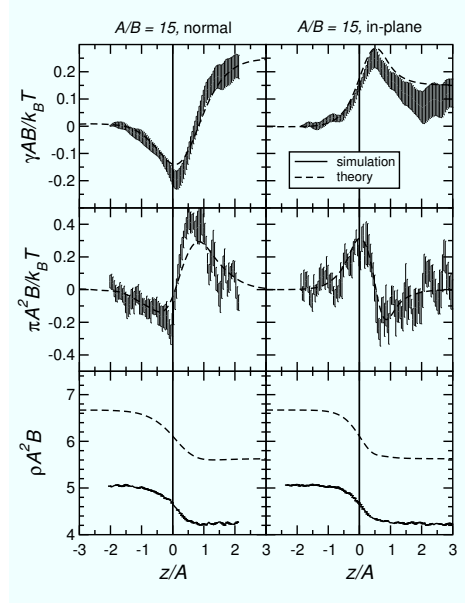


Figure 16. Profiles of number density ρ , pressure tensor difference $\pi(z)$, and running integral $\int \pi(z) dz$ (which gives the surface tension γ), all in appropriate reduced units. Director alignment is normal to the interface or in the plane of the interface, as indicated. $A/B = 15$ hard ellipsoids. Simulation results are represented by error bars, calculated from 8 independent simulations, and the predictions of Onsager theory are represented by dashed lines (After Ref 72).

The pressure tensor profiles for the I-N interface of hard and soft ellipsoids with $A/B = 15$ are measurable in simulation, taking care to eliminate systematic errors, finite size effects, and statistical fluctuations⁷². The results are shown in Fig. 16. Several interesting things emerge. The simulation results are quite noisy, but they are confirmed in detail by the Onsager theory; the surface tension of the in-plane orientation is indeed found to be lower than for normal alignment; and most dramatically, the curves for the two director orientations are qualitatively different in character. For in-plane alignment, the tension appears on the nematic side of the interface, and there is a small compressive region on the isotropic side, reducing the eventual value of γ somewhat. The net effect is that the surface of tension, defined by the first moment of the pressure tensor profile, is about one molecule length away from the interface on the nematic side. For normal alignment, the tension appears on the *isotropic* side, and there is some compression on the nematic side; the surface of tension is a significant distance, about two molecule lengths, from the interface on the isotropic side.

7.3 Capillary Waves

The Onsager theory *neglects* capillary wave fluctuations of the interface; these are also suppressed in computer simulations with insufficient transverse box dimensions. However, they may be detected in much larger-scale simulations⁷³. Suppose that we split the simulation box, of overall dimensions $L_\perp \times L_\perp \times L_z$ into blocks (columns) of transverse dimension $\ell_\perp \times \ell_\perp$ and length L_z ; for each block we measure the deviation $h(x, y)$ of the interfacial position from its average. The capillary wave Hamiltonian is given by the quadratic form:

$$\mathcal{H}_{\text{CW}} = \frac{\gamma}{2} \int \int dx dy \left[\left(\frac{\partial h}{\partial x} \right)^2 + \left(\frac{\partial h}{\partial y} \right)^2 \right] = \frac{\gamma}{2} \sum_{\mathbf{q}_\perp} q_\perp^2 |h(\mathbf{q}_\perp)|^2$$

with transverse wave-vector $\mathbf{q}_\perp = (q_x, q_y)$. Hence $\langle |h(\mathbf{q}_\perp)|^2 \rangle = k_B T / \gamma q_\perp^2$ and the local mean-squared deviation of the interface may be obtained:

$$s^2 = \langle h^2(x, y) \rangle = \frac{1}{4\pi^2} \int d\mathbf{q}_\perp \langle |h(\mathbf{q}_\perp)|^2 \rangle = \frac{k_B T}{2\pi\gamma} \ln \left(\frac{q_{\text{max}}}{q_{\text{min}}} \right)$$

with $q_{\text{min}} = 2\pi/L_\perp$, $q_{\text{max}} = 2\pi/\ell_\perp$. We expect a Gaussian distribution

$$P(h) = (2\pi s^2)^{-1/2} \exp(-h^2/2s^2).$$

Moreover the intrinsic thickness of the interface, t_0 , will be broadened, within a block of transverse dimension $\ell_\perp$, by capillary waves as follows

$$t^2 = t_0^2 + \frac{\pi}{2} \hat{s}^2 = t_0^2 + \frac{k_B T}{4\gamma} \ln \left(\frac{\hat{q}_{\text{max}}}{\hat{q}_{\text{min}}} \right)$$

where $\hat{q}_{\text{min}} = 2\pi/\ell_\perp$ and $\hat{q}_{\text{max}} \sim 2\pi/a_0$, a_0 = a microscopic length.

We studied a system of $N = 115200$ soft ellipsoids with $A/B = 15$, in a periodic box with dimensions $L_x = L_y = L_\perp \approx 150B$ and $L_z \approx 300B$. A nematic-isotropic film system was prepared and allowed to stabilize with the director assuming the preferred planar orientation: 1.2×10^6 MD steps were allowed for equilibration, and $\sim 2.0 \times 10^6$ MD

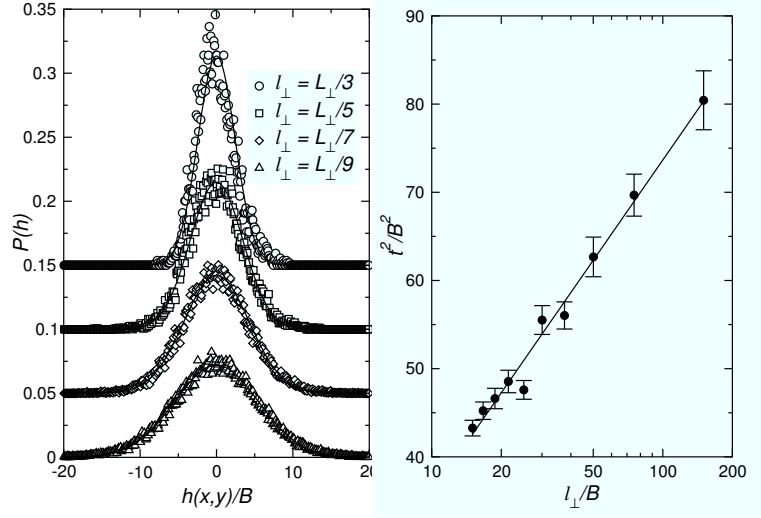


Figure 17. **Left:** Probability distribution for interface height $P(h)$ averaged over blocks of various transverse dimensions (indicated). Curves are displaced vertically for clarity. **Right:** Mean squared interfacial thickness t^2 as a function of transverse dimensions of blocks (log scale) (After Ref. 73).

steps to collect data. The local interfacial position distribution $P(h)$ is shown in Fig. 17: it is fitted well by the Gaussian distribution. The squared interfacial thickness, t^2 , is plotted as a function of block size $\ell_{\perp}$, in Fig. 17. Particularly at large $\ell_{\perp}$, the dependence of t^2 on $\ell_{\perp}$ is logarithmic, which is consistent with the capillary wave picture. The interfacial tension estimated from this logarithmic dependence is in reasonable agreement with the values obtained from pressure tensor measurements.

We also study the difference in the capillary wave spectrum as a function of $q_{\perp}$ with $q_{\perp}$ in the plane of the interface, parallel and perpendicular to the director. Typical results are shown in Fig. 18. The low- $q_{\perp}$ behaviour is roughly consistent with the values of γ deduced from the interface width, and from the earlier pressure tensor calculation. The capillary wave spectrum is anisotropic at higher values of $q_{\perp}$. Fluctuations parallel to the director are less easy than those perpendicular to the director. The curvature in these figures indicates a *negative* bending rigidity.

8 Colloid Particles in Liquid Crystals

Colloidal particles or droplets suspended in a liquid crystal are of special interest⁷⁴. Elastic deformations of the director field around the colloidal particles produce additional long-range forces between them^{75–77}. These interactions can be of dipolar or quadrupolar type depending on the symmetry of the director field around the particles⁷⁸, and this in turn is extremely sensitive to the boundary conditions on the particle surface and the size of the particles⁷⁹. The result may be supermolecular structures^{74,80}, cellular structures^{81,82}, and even a soft solid⁸³ can be observed. Colloidal dispersions in liquid crystals also have a wide variety of potential applications⁸⁴.

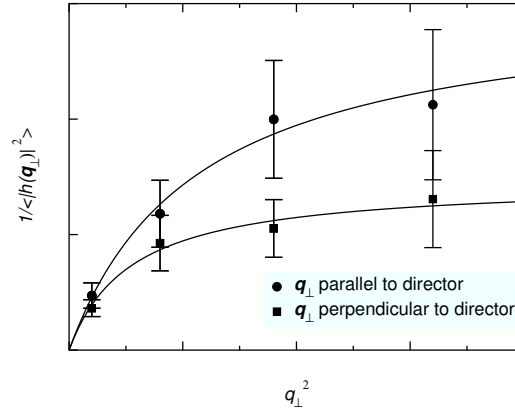


Figure 18. Capillary wave spectrum as function of wavenumber, where the wave vector is directed respectively parallel and perpendicular to the director in the plane of the interface. The lines are to guide the eye (After Ref. 73).

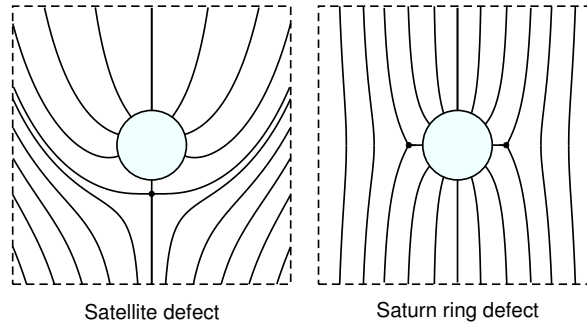


Figure 19. Sketches of director field streamlines around a colloid particle, showing the satellite defect and Saturn ring defect

A key feature here is the presence of *topological defects*^{78,85}. Strong homeotropic anchoring of the director, normal to the particle surface, induces a radial hedgehog defect with topological charge $+1$. If the director field is uniform far from the particle, i.e. the total charge of the whole system is zero, topological considerations imply that an additional defect must be created in the medium to compensate the radial hedgehog.

There are several types of defect which can arise in this case. Two are illustrated in Fig. 19. The first is a point defect: a hyperbolic hedgehog with a topological charge -1 , called a dipolar or *satellite* defect. The second is a quadrupolar or *Saturn-ring* defect, i.e. a $-1/2$ strength disclination ring that encircles the spherical particle. Theoretical and numerical work based on elastic theory^{86,85} suggest that the dipole configuration is stable for the micron-sized droplets which are usually realised experimentally; the Saturn-ring configuration should appear if the droplet size is reduced or an external field is applied^{87,88},

and, when present, it is always predicted to be most stable in the equatorial plane normal to the director. The director distortion around the ring defect vanishes with distance in a manner consistent with the *quadrupolar* nature of the defect: far from the particle, at distance r in a direction θ relative to the director, the director deviation angle β has asymptotic behaviour $\beta \sim (R/r)^3 \sin 2\theta$ where R is the particle radius. The director distortion around the satellite defect extends much further than that of the ring defect. This is in accord with its symmetry: far from the particle, the director angle vanishes as $\beta \sim (R/r)^2 \sin \theta$, i.e. it is like a dipolar term in a multipole expansion. To minimize elastic interactions between images in computer simulations, rather larger periodic boxes are required in the satellite defect case.

Earlier investigations have simulated the Saturn ring type of defect for this kind of system⁸⁹, using 2048 Gay-Berne ellipsoids. We have recently looked at larger systems⁹⁰, investigating the structure for a range of colloid radii R , using soft ellipsoids with elongation $A/B = 3$ for the liquid crystalline solvent. The interaction of each molecule with the droplet was given by a shifted Lennard-Jones repulsion potential between the centers. This results in homeotropic anchoring of the liquid crystal molecules, normal to the particle surface. The systems consisted of 8,000-1,000,000 mesogens.

8.1 Saturn Ring Defect

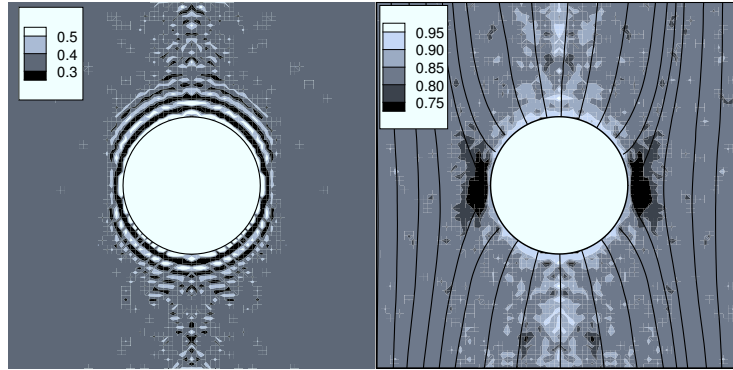


Figure 20. Density map (left) and order parameter map with director streamlines (right) in the spatial region around the colloid particle, for the Saturn ring defect. Droplet radius $R/B = 15$ (After Ref. 90).

For all studied radii in the range $R/B = 3-15$, the ring defect is relatively easy to prepare: it appears immediately after equilibration of the nematic phase of the system starting from the isotropic state, or on enlarging the droplet from a small initial radius in the nematic solvent. It seems likely, therefore, that this type of the defect is energetically more favourable for the chosen droplet sizes.

Typical maps of density and order parameter are shown in Fig. 20. Strong density oscillations can be seen *except* in the region of the defect ring, which is identifiable by greatly reduced orientational ordering. Superimposed on the order parameter map are streamlines of the director field (note that these are not reliably defined in the defect regions

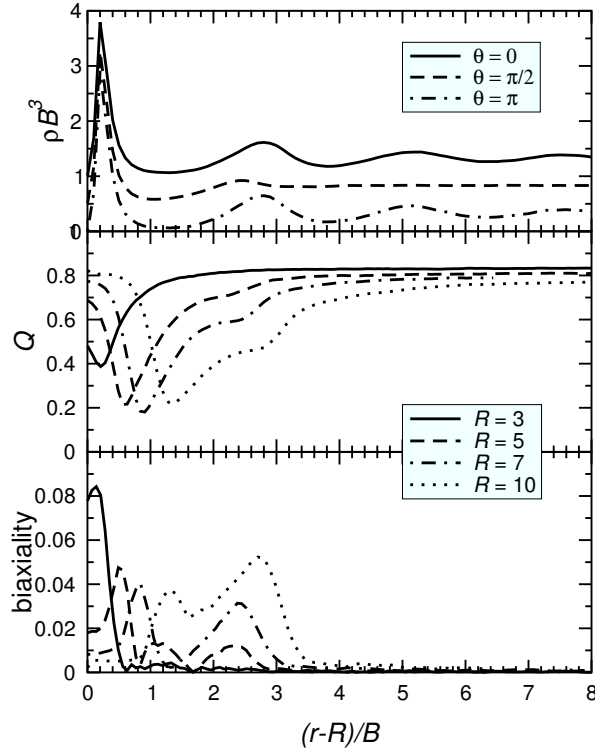


Figure 21. Above: Density profiles for the ring defect. Droplet radius $R/B = 15$, $N = 64000$ particles. The following directions are shown: $\theta = 0, \pi$ (avoiding the defect) and $\theta = \pi/2$ (crossing the disclination ring). Successive curves are offset vertically by 0.5 for clarity. Below: Order parameter and biaxiality profiles for the ring defect along the direction $\theta = \pi/2$. Different curves correspond to the different droplet radii (After Ref. 90).

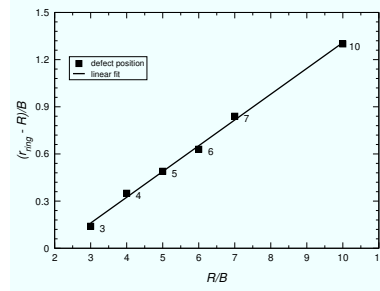


Figure 22. Distance of the core region of the ring defect from the droplet center, as a function of reduced droplet radius R/B . We also show the linear fit to the simulation results, $r_{\text{ring}} = 1.164R - 0.33B$ (After Ref. 90).

themselves, where the order parameter is very low). More detailed information may be obtained by calculating profiles in radial slices. Typical density profiles for the ring defect

with $R/B = 15$ are shown in Fig. 21 for directions $\theta = 0, \pi$ relative to the director (not intersecting the defect) and $\theta = \pi/2$ (crossing the disclination ring). The profiles which avoid the disclination have an oscillating structure near the particle surface which is typical for a *liquid-wall* interface. The profiles which cross the disclination ring do not have oscillations. The difference may be due to the partial melting of the liquid crystal in the disclination core region. This melting damps the influence of the droplet surface on the interface region.

The order parameter and biaxiality profiles for $\theta = \pi/2$ are shown in Fig. 21. The shape of these profiles in general reflects the typical structure of the core: the centre of the core has lower order than the bulk and the core region extends over a few molecular lengths. Both order parameter and biaxiality profiles indicate that the structure of the core is not simple.

Using the minima of the order profiles, we extracted the distance from the core of the disclination ring to the particle surface. The dependence of this distance on the particle radius R is shown in Fig. 22. It is interesting to compare with the phenomenological theory predicting linear dependence of the ring defect radius on the droplet radius: $r_{\text{ring}} \approx 1.25R$ from minimization of the elastic free energy using a trial function⁷⁹, $r_{\text{ring}} \approx 1.13R$ using a simulated annealing method⁸⁶, or $r_{\text{ring}} \approx 1.26R$ ⁸⁵. Our MD simulation results give $r_{\text{ring}} - R = -0.33B + (0.164 \pm 0.004)R$ which is in good agreement with the phenomenological theory, especially if one bears in mind the complex structure of the defect core.

8.2 Satellite Defect

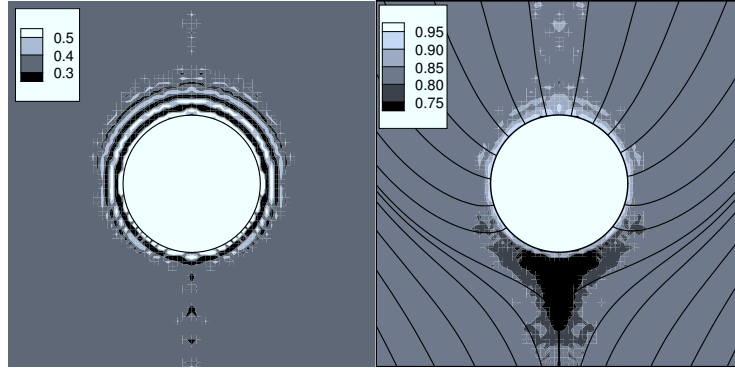


Figure 23. Density maps and order parameter maps with director streamlines for the satellite defect. Droplet radius $R/B = 15$ (After Ref. 90).

To study the satellite defect, one needs very large systems. We used one million particles and droplet radius $R = 15B$. Below this radius, an initial configuration corresponding to the satellite defect structure would transform into the Saturn ring. (For $R = 10B$, this transformation occurred slowly enough to follow; see below).

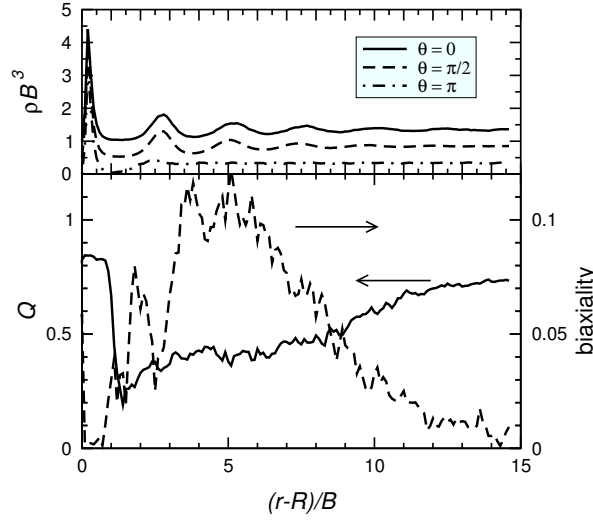


Figure 24. Above: Density profiles for the satellite defect. Droplet radius $R/B = 15$. The following directions are shown: $\theta = 0, \pi/2$ (avoiding the defect) and $\theta = \pi$ (crossing the defect). Successive curves are offset vertically by 0.5 for clarity. Below: Order parameter and biaxiality profiles for the satellite defect along the direction $\theta = \pi$ (After Ref. 90).

Typical maps of density and order parameter are shown in Fig. 23. The density profiles for the satellite defect are shown in Fig. 24. One can see that the density profile at $\theta = \pi$ (across the defect core) has less prominent oscillating structure than the other two: this is again presumably due to the partial disordering of the mesophase in the core region. The order parameter and biaxiality profiles in the $\theta = \pi$ direction are shown in Fig. 24. The defect is rather broad, and difficult to locate from these graphs. However, examination of the director field allows one to place the centre at a distance $r_{\text{sat}} \approx 1.4R$. The value predicted by the elastic theory is about $1.22R$ for the simulated annealing calculations⁸⁶; $1.17R$ ⁷⁴ or $1.46R$ ⁹¹ for the free energy minimization using a trial function.

8.3 Off-Center Ring

Simulation results show that both the satellite and ring defects are at least metastable for $R/B = 15$: once the particular defect is realized in the system, it is stable over the timescale that is accessible to our simulations. However, the satellite defect is not stable for the smaller droplets. Indeed, we observed a rapid transition of the satellite defect to the ring defect for $R/B < 10$. Equilibrating the initial configuration with the satellite defect in the cell with droplet radius $R/B = 10$, we observed that it evolves into an *off-centered* ring defect. The ring then moved slowly, evolving towards an *equatorial* Saturn ring configuration of the usual kind. Doing long runs (up to a million timesteps) we conclude that the off-centred configuration is an intermediate state between the satellite and the Saturn ring defect. The z -coordinate of the ring, relative to the droplet center, as a function of the number of the timesteps, is shown in Fig. 25. The evolution dynamics is quite slow. Note

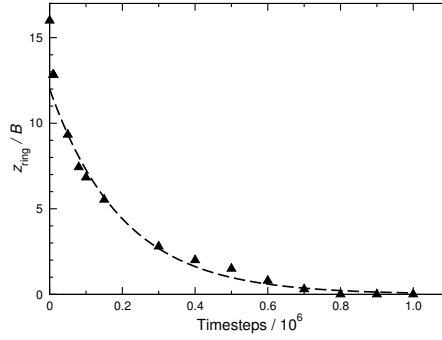


Figure 25. Distance of centre of ring from centre of colloid, as a function of time, for $R/B = 10$ (After Ref. 90).

that elastic theory also predicts the off-center ring configuration to be unstable⁸⁵, with the transition from the dipole configuration to the Saturn ring configuration occurring via this intermediate state.

Recently these studies have been extended to measure the solvent-mediated force between two spherical colloids in a liquid crystal⁹². It has been possible to resolve the component of the force perpendicular to the inter-particle vector (which would vanish in an isotropic solvent) and show that it is in qualitative agreement with the predictions of elastic theory. However the component along the inter-particle vector is dominated by solvation structure and depletion effects.

8.4 Non-Spherical Colloidal Particle

For a rod-like elongated particle with length L and transverse size $D \ll L$, and both L and D much greater than the dimensions of the molecules of the liquid crystal, one can have a disclination line of strength -1 , a pair of disclination lines of strength $-\frac{1}{2}$, as well as the ‘escaped radial’ structure, in which the director bends over to become perpendicular to the particle surface⁹³. However, from the energetic point of view, the situation is different from the case of the spherical particle. For the elongated particle, both defects are disclination *lines*. The elastic energy per unit length associated with a disclination of strength m is $\pi K m^2 \ln(R/r_0)$, where R is the size of the sample and r_0 is a lower cutoff radius (the core size)⁹⁴. This means that the free energy of a pair of $-\frac{1}{2}$ disclinations is always smaller than that of a single -1 disclination. Therefore, one can expect that the pair of $-\frac{1}{2}$ disclinations will always be a stable configuration, although in principle the -1 defect can still form a metastable state.

Phenomenological (continuum) theory⁹³ indicates that, depending on the type and strength of anchoring, the equilibrium orientation of the particle s may be either parallel or perpendicular to the liquid crystal director n , depending on the ratio of the particle radius to the extrapolation length of the nematic liquid crystal. Qualitatively, the free energy should vary as

$$\mathcal{F} = \mathcal{F}_\perp + (\mathcal{F}_\parallel - \mathcal{F}_\perp) \cos^2 \theta, \quad (5)$$

where $\cos \theta = \mathbf{n} \cdot \mathbf{s}$ and $\mathcal{F}_{\parallel}$ and $\mathcal{F}_{\perp}$ are constants. This gives a $\sin 2\theta$ dependence for the torque and predicts that the director response has a maximum at $\theta = \pi/4$ and is absent at $\theta = 0, \pi/2$. However, it is clear that this description is oversimplified: the defect structure changes while the particle rotates and the nematic ordering evolves in a complicated way.

Molecular dynamics simulations of $N = 64000$ repulsive, soft, ellipsoids of elongation 3 were carried out. For all studied diameters of the rod ($D/B = 5\text{--}20$) the -1 strength disclination appears immediately after expanding the colloid particle in the nematic state. However, during the equilibration, it splits into two $-\frac{1}{2}$ disclination lines which then move slowly towards the equatorial plane. We were not able to observe the ‘escaped radial’ configuration, probably due to the small size of the colloidal particle.

To measure the torque on a rod of finite length, we performed MD simulations in a box with periodic boundary conditions, applying a global constraint for the director along the z axis. An independent measurement was performed in slab geometry, with the walls provided strong homeotropic (along the z axis) anchoring of the director, and with the rod located in the centre of the box. The rod diameter was $D = 5B$, the rod length $L = 10B$. The torque on the rod as a function of tilt angle is presented in Fig. 26. A positive torque tends to align the rod perpendicular to the director. The results presented indicate that the dependence of the torque on the rod tilt angle is far from the $\sin 2\theta$, proposed in⁹³; indeed, the torque is not equal to zero for $\theta = 0^\circ$, i.e. there is some symmetry breaking and the orientation of the rod along the director is not even metastable.

For better understanding, a slice in the yz plane is shown in Fig. 26 for different tilt angles of the rod. This shows that the director distribution around the rod is not axially symmetric, and hence the torque is nonzero for $\theta = 0^\circ$. As the rod rotates, the director field becomes less and less frustrated, and finally we have a stable orientation of the rod perpendicular to the director, $\theta = 90^\circ$. In principle, a configuration with axial symmetry is also possible when the rod is along the z axis. However, we were not able to observe it in our simulations even when disordered, isotropic configurations containing the colloidal particle were compressed to the ordered, nematic state.

9 Conclusions

This chapter has attempted to cover several aspects of liquid crystal behaviour that are amenable to study by simulation. It is now possible to tackle features of the structure which have length scales significantly larger than molecular sizes: defects, interfacial fluctuations, distortions due to suspended colloidal particles or droplets. Little mention has been made, due to limitations of space, of many more aspects which are also of great interest: smectic (layered) phases, for instance, or reorientational dynamics and fluid flow, which are important in understanding the operation of liquid crystal devices. These also provide rich fields of study, both by molecular simulation and other modelling techniques, which may extend the accessible length and timescales towards the mesoscopic and macroscopic regimes. Many challenges remain in these areas. Finally, the predictive power of simulations remains limited: it is still quite a challenge to use atomistically correct potentials to predict, reliably, liquid crystalline phase behaviour and properties. Nonetheless, the progress that has been made using simple models is quite encouraging, and the steady increase in available computer power may bring these goals within reach in a few years.

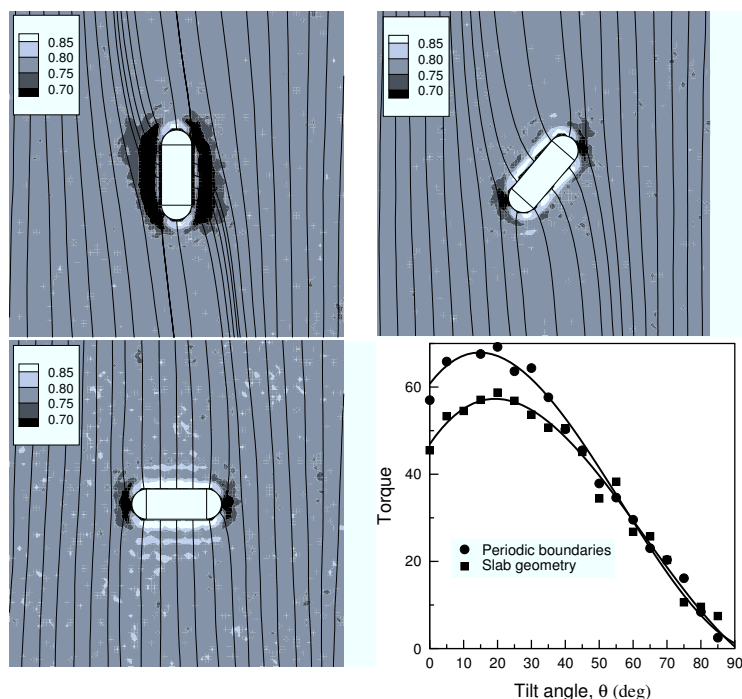


Figure 26. Director streamlines and order parameter maps for different tilt angles. A side view along the x axis is shown (the rod is tilted in the yz plane). We also plot the torque on the rod vs. rod tilt angle for two systems: one with periodic boundaries and the other in slab geometry. The lines are to guide the eye. For more details see Ref. 95

Acknowledgements

It is a pleasure to acknowledge the work of past and present members of my research group, too many to list individually. My work in liquid crystals began with an enjoyable collaboration with Daan Frenkel, and subsequently several members of his group, and was further stimulated by visits to Claudio Zannoni and a period of study leave in Mainz, hosted by Kurt Kremer and Kurt Binder. The support of the Alexander von Humboldt foundation, the British Council ARC programme, the Leverhulme Trust, and the UK Engineering and Physical Sciences Research Council, is gratefully acknowledged.

References

1. M. A. Glaser, R. Malzbender, N. A. Clark, and D. M. Walba. Atomic detail simulation studies of tilted smectics. *J. Phys. Cond. Mat.*, 6:A261–A268, 1994.
2. E. Garcia, M. A. Glaser, N. A. Clark, and D. M. Walba. HFF: a force field for liquid crystal molecules. *J. Molec. Struct. - THEOCHEM*, 464:39–48, 1999.
3. S. J. Clark, C. J. Adam, G. J. Ackland, J. White, and J. Crain. Properties of liquid

- crystal molecules from first principles computer simulation. *Liq. Cryst.*, 22:469–475, 1997.
4. S. J. Clark, C. J. Adam, D. J. Cleaver, and J. Crain. Conformational energy landscapes of liquid crystal molecules. *Liq. Cryst.*, 22:477–482, 1997.
 5. D. L. Cheung, S. J. Clark, and M. R. Wilson. Parametrization and validation of a force field for liquid-crystal forming molecules. *Phys. Rev. E*, 65:051709, 2002.
 6. J. G. Gay and B. J. Berne. Modification of the overlap potential to mimic a linear site-site potential. *J. Chem. Phys.*, 74:3316–3319, 1981.
 7. E. de Miguel, L. F. Rull, M. K. Chalam, and K. E. Gubbins. Liquid-vapor coexistence of the Gay-Berne fluid by Gibbs ensemble simulation. *Molec. Phys.*, 71:1223–1231, 1990.
 8. R. Berardi, A. P. J. Emerson, and C. Zannoni. Monte Carlo investigations of a Gay-Berne liquid crystal. *J. Chem. Soc. Faraday Trans.*, 89:4069–4078, 1993.
 9. E. de Miguel, E. Martín del Río, J. T. Brown, and Michael P. Allen. Effect of the attractive interactions on the phase behavior of the Gay-Berne liquid crystal model. *J. Chem. Phys.*, 105:4234–4249, 1996.
 10. J. T. Brown, Michael P. Allen, E. Martín del Río, and E. de Miguel. Effects of elongation on the phase behavior of the Gay-Berne fluid. *Phys. Rev. E*, 57:6685–6699, 1998.
 11. R. Berardi, C. Fava, and C. Zannoni. A generalized Gay-Berne intermolecular potential for biaxial particles. *Chem. Phys. Lett.*, 236:462–468, 1995.
 12. P. J. Camp and Michael P. Allen. Phase diagram of the hard biaxial ellipsoid fluid. *J. Chem. Phys.*, 106:6681–6688, 1997.
 13. P. J. Camp, Michael P. Allen, and A. J. Masters. Theory and computer simulation of bent-core molecules. *J. Chem. Phys.*, 111:9871–9881, 1999.
 14. J. S. van Duijneveldt and Michael P. Allen. Computer simulation study of a flexible-rigid-flexible model for liquid crystals. *Molec. Phys.*, 92:855–870, 1997.
 15. Alexey V. Lyulin, Muataz S. Al-Barwani, Michael P. Allen, Mark R. Wilson, Igor Neelov, and Nicholas K. Allsopp. Molecular dynamics simulation of main chain liquid crystalline polymers. *Macromolecules*, 31:4626–4634, 1998.
 16. J. S. van Duijneveldt, A. Gilvillegas, G. Jackson, and Michael P. Allen. 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.*, 112:9092–9104, 2000.
 17. R. Berardi, M. Fehervari, and C. Zannoni. A Monte Carlo simulation study of associated liquid crystals. *Molec. Phys.*, 97:1173–1184, 1999.
 18. D. Frenkel. Introduction to Monte Carlo methods. In *Computational Soft Matter: From Synthetic Polymers to Proteins*, 2004. this volume.
 19. M. P. Allen. Introduction to molecular dynamics simulations. In *Computational Soft Matter: From Synthetic Polymers to Proteins*, 2004. this volume.
 20. Michael P. Allen and Dominic J. Tildesley. *Computer simulation of liquids*. Clarendon Press, Oxford, paperback, 385pp edition, 1989.
 21. D. Frenkel and B. Smit. *Understanding molecular simulation : from algorithms to applications*. Academic Press, San Diego, 2nd edition, 2002.
 22. C. Oseen. Theory of liquid crystals. *Trans. Faraday Soc.*, 29:883–899, 1933.
 23. F. C. Frank. On the theory of liquid crystals. *Discuss. Faraday Soc.*, 25:19–28, 1958.

24. G. P. Crawford and S. Žumer. Saddle-splay elasticity in nematic liquid-crystals. *Int. J. Mod. Phys. B*, 9:2469–2514, 1995.
25. G. P. Crawford and S. Žumer, editor. *Liquid crystals in complex geometries formed by polymer and porous networks*, London, 1996. Taylor and Francis.
26. H. Yokoyama. Density-functional theory of surfacelike elasticity of nematic liquid crystals. *Phys. Rev. E*, 55:2938–2957, 1997.
27. P. G. de Gennes and J. Prost. *The Physics of Liquid Crystals*. Clarendon Press, Oxford, second, paperback edition, 1995.
28. L. Onsager. The effects of shape on the interaction of colloidal particles. *Ann. N. Y. Acad. Sci.*, 51:627, 1949.
29. R. Evans. *Adv. Phys.*, 28:143, 1979.
30. R. Evans. Microscopic theories of simple fluids and their interfaces. In J. Charvolin, J. F. Joanny, and J. Zinn-Justin, editors, *Liquids at Interfaces*, pages 1–98, Amsterdam, 1989. Elsevier Science Publishers B.V. Les Houches, Session XLVIII 1988.
31. R. Evans. Density functionals in the theory of nonuniform fluids. In D. Henderson, editor, *Fundamentals of Inhomogeneous Fluids*, chapter 3, pages 85–175. Dekker, New York, 1992.
32. V. V. Ginzburg, M. A. Glaser, and N. A. Clark. Liquid crystal phase diagram of the Gay-Berne fluid by density functional theory. *Liq. Cryst.*, 23:227–234, 1997.
33. J. D. Parsons. *Phys. Rev. A*, 19:1225, 1979.
34. S.-D. Lee. A numerical investigation of nematic ordering based on a simple hard-rod model. *J. Chem. Phys.*, 87:4972–4974, 1987.
35. S.-D. Lee. The Onsager-type theory for nematic ordering of finite-length hard ellipsoids. *J. Chem. Phys.*, 89:7036–7037, 1989.
36. D. Frenkel, B. M. Mulder, and J. P. McTague. Phase diagram of a system of hard ellipsoids. *Phys. Rev. Lett.*, 52:287–290, 1984.
37. D. Frenkel and B. M. Mulder. The hard ellipsoid-of-revolution fluid I. Monte Carlo simulations. *Molec. Phys.*, 55:1171–1192, 1985.
38. R. Evans. Comment. *Molec. Phys.*, 100:199–200, 2002.
39. Philip J. Camp, Carl P. Mason, Michael P. Allen, Anjali A. Khare, and David A. Kofke. The isotropic-nematic transition in uniaxial hard ellipsoid fluids: coexistence data and the approach to the Onsager limit. *J. Chem. Phys.*, 105:2837–2849, 1996.
40. A. Z. Panagiotopoulos. Direct determination of phase coexistence properties of fluids by Monte Carlo simulation in a new ensemble. *Molec. Phys.*, 61:813–826, 1987.
41. A. Z. Panagiotopoulos. Gibbs ensemble techniques. In M. Baus, L. F. Rull, and J.-P. Ryckaert, editors, *Observation, prediction and simulation of phase transitions in complex fluids*, volume 460 of *NATO ASI Series C*, pages 463–501, Dordrecht, 1995. Kluwer Academic Publishers. Proceedings of the NATO Advanced Study Institute on ‘Observation, prediction and simulation of phase transitions in complex fluids’, Varenna, Italy, July 25–August 5, 1994.
42. I. Nezbeda and J. Kolafa. A new version of the insertion particle method for determining the chemical potential by Monte Carlo simulation. *Molec. Simul.*, 5:391–403, 1991.
43. P. Attard. Simulation of the chemical potential and the cavity free energy of dense hard-sphere fluids. *J. Chem. Phys.*, 98:2225–2231, 1993.
44. D. A. Kofke. Gibbs-Duhem integration: a new method for direct evaluation of phase

- coexistence by molecular simulation. *Molec. Phys.*, 78:1331–1336, 1993.
45. Philip J. Camp, Michael P. Allen, Peter G. Bolhuis, and Daan Frenkel. Demixing in hard ellipsoid rod-plate mixtures. *J. Chem. Phys.*, 106:9270–9275, 1997.
 46. E. Martín del Río. *Estudio de las propiedades interfaciales de cristales líquidos nemáticos*. PhD thesis, Seville University, 1996.
 47. Elvira Martín del Río and Enrique de Miguel. Computer simulation study of the free surfaces of a liquid crystal model. *Phys. Rev. E*, 55:2916–2924, 1997.
 48. A. P. J. Emerson, S. Faetti, and C. Zannoni. Monte Carlo simulation of the nematic-vapour interface for a Gay-Berne liquid crystal. *Chem. Phys. Lett.*, 271:241–246, 1997.
 49. Michael P. Allen. Diffusion coefficient increases with density in hard ellipsoid liquid crystals. *Phys. Rev. Lett.*, 65:2881–2884, 1990.
 50. Siegfried Hess, Daan Frenkel, and Michael P. Allen. On the anisotropy of diffusion in nematic liquid crystals: test of a modified affine transformation model via molecular dynamics. *Molec. Phys.*, 74:765–774, 1991.
 51. Z. Zhang, O. G. Mouritsen, and M. J. Zuckermann. Weak first-order orientational transition in the Lebwohl-Lasher model for liquid crystals. *Phys. Rev. Lett.*, 69:2803–2806, 1992.
 52. Z. Zhang, M. J. Zuckermann, and O. G. Mouritsen. Phase transition and director fluctuations in the 3-dimensional Lebwohl-Lasher model of liquid crystals. *Molec. Phys.*, 80:1195–1221, 1993.
 53. Michael P. Allen and Mark A. Warren. Simulation of structure and dynamics near the isotropic-nematic transition. *Phys. Rev. Lett.*, 78:1291–1294, 1997.
 54. C. G. Gray and K. E. Gubbins. *Theory of molecular fluids. 1. Fundamentals*. Clarendon Press, Oxford, 1984.
 55. M. R. Wilson, Michael P. Allen, M. A. Warren, A. Sauron, and W. Smith. Replicated data and domain decomposition molecular dynamics techniques for simulation of anisotropic potentials. *J. Comput. Chem.*, 18:478–488, 1997.
 56. Michael P. Allen, Mark A. Warren, Mark. R. Wilson, Alain Sauron, and William Smith. Molecular dynamics calculation of elastic constants in Gay-Berne nematic liquid crystals. *J. Chem. Phys.*, 105:2850–2858, 1996.
 57. E. de Miguel, L. F. Rull, M. K. Chalam, and K. E. Gubbins. Liquid crystal phase diagram of the Gay-Berne fluid. *Molec. Phys.*, 74:405–424, 1991.
 58. Michael P. Allen. Molecular simulation and theory of liquid crystal surface anchoring. *Molec. Phys.*, 96:1391–1397, 1999.
 59. Michael P. Allen. Modelling liquid crystal structure, phase behaviour and large-scale phenomena. In C. Zannoni and P. Pasini, editors, *Advances in the Computer Simulations of Liquid Crystals*, volume 545 of *NATO ASI Series C*, pages 73–97. Kluwer Academic Publishers, Dordrecht, 2000. Proceedings of the NATO Advanced Study Institute ‘Advances in the Computer Simulations of Liquid Crystals’, Erice, Italy, June 11–21, 1998.
 60. Denis Andrienko, Guido Germano, and Michael P. Allen. Liquid crystal director fluctuations and surface anchoring by molecular simulation. *Phys. Rev. E*, 62:6688–6693, 2000.
 61. Denis Andrienko and Michael P. Allen. Theory and simulation of the nematic zenithal anchoring coefficient. *Phys. Rev. E*, 65:021704/1–11, 2002.

62. A. Mertelj and M. Čopič. Dynamic light scattering as a probe of orientational dynamics in confined liquid crystals. *Phys. Rev. E*, 61:1622–1628, 2000.
63. Michael P. Allen. Molecular simulation and theory of the isotropic-nematic interface. *J. Chem. Phys.*, 112:5447–5453, 2000.
64. R. van Roij, M. Dijkstra, and R. Evans. Orientational wetting and capillary nematization of hard-rod fluids. *Europhys. Lett.*, 49:350–356, 2000.
65. R. van Roij, M. Dijkstra, and R. Evans. Interfaces, wetting, and capillary nematization of a hard-rod fluid: theory for the Zwanzig model. *J. Chem. Phys.*, 113:7689–7701, 2000.
66. Muataz S. Al-Barwani and Michael P. Allen. Isotropic-nematic interface of soft spherocylinders. *Phys. Rev. E*, 62:6706–6710, 2000.
67. D. L. Koch and O. G. Harlen. Interfacial tension at the boundary between nematic and isotropic phases of a hard rod solution. *Macromolecules*, 32:219–226, 1999.
68. P. Schofield and J. R. Henderson. Statistical mechanics of inhomogeneous fluids. *Proc. Roy. Soc. Lond. A*, 379:231–246, 1982.
69. A. Harasima. Molecular theory of surface tension. *Adv. Chem. Phys.*, 1:203–237, 1958.
70. J. H. Irving and J. G. Kirkwood. The statistical mechanical theory of transport processes IV The equations of hydrodynamics. *J. Chem. Phys.*, 18:817–829, 1950.
71. Michael P. Allen. Pressure tensor profiles at the isotropic-nematic interface. *Chem. Phys. Lett.*, 331:513–518, 2000.
72. Andrew J. McDonald, Michael P. Allen, and Friederike Schmid. Surface tension of the isotropic-nematic interface. *Phys. Rev. E*, 63:010701(R)/1–4, 2000.
73. Nobuhiko Akino, Friederike Schmid, and Michael P. Allen. Molecular dynamics study of the nematic-isotropic interface. *Phys. Rev. E*, 63:041706/1–6, 2001.
74. P. Poulin, H. Stark, T. C. Lubensky, and D. A. Weitz. Novel colloidal interactions in anisotropic fluids. *Science*, 275:1770–1773, 1997.
75. B. I. Lev and P. M. Tomchuk. Interaction of foreign macrodroplets in a nematic liquid crystal and induced supermolecular structures. *Phys. Rev. E*, 59:591–602, 1999.
76. A. Borštnik, H. Stark, and S. Žumer. Interaction of spherical particles dispersed in a liquid crystal above the nematic-isotropic phase transition. *Phys. Rev. E*, 60:4210–4218, 1999.
77. A. Borštnik, H. Stark, and S. Žumer. Temperature-induced flocculation of colloidal particles immersed into the isotropic phase of a nematic liquid crystal. *Phys. Rev. E*, 61:2831–2839, 2000.
78. T. C. Lubensky, D. Petey, N. Currier, and H. Stark. Topological defects and interactions in nematic emulsions. *Phys. Rev. E*, 57:610–625, 1998.
79. O. V. Kuksenok, R. W. Ruhwandl, S. V. Shiyankovskii, and E. M. Terentjev. Director structure around a colloid particle suspended in a nematic liquid crystal. *Phys. Rev. E*, 54:5198–5203, 1996.
80. P. Poulin and D. A. Weitz. Inverted and multiple nematic emulsions. *Phys. Rev. E*, 57:626–637, 1998.
81. V. J. Anderson, E. M. Terentjev, S. P. Meeker, J. Crain, and W. C. K. Poon. Cellular solid behaviour of liquid crystal colloids: 1. Phase separation and morphology. *Euro. Phys. J. E*, 4:11–20, 2001.
82. V. J. Anderson and E. M. Terentjev. Cellular solid behaviour of liquid crystal colloids:

2. Mechanical properties. *Euro. Phys. J. E*, 4:21–28, 2001.
83. S. P. Meeker, W. C. K. Poon, J. Crain, and E. M. Terentjev. Colloid-liquid-crystal composites: An unusual soft solid. *Phys. Rev. E*, 61:R6083–R6086, 2000.
84. W. B. Russel, D. A. Saville, and W. R. Schowalter. *Colloidal dispersions*. Cambridge University Press, Cambridge, 1989.
85. H. Stark. Director field configurations around a spherical particle in a nematic liquid crystal. *Euro. Phys. J. B*, 10:311–321, 1999.
86. R. W. Ruhwandl and E. M. Terentjev. Monte Carlo simulation of topological defects in the nematic liquid crystal matrix around a spherical colloid particle. *Phys. Rev. E*, 56:5561–5565, 1997.
87. Y. Gu and N. L. Abbott. Observation of saturn-ring defects around solid microspheres in nematic liquid crystals. *Phys. Rev. Lett.*, 85:4719, 2000.
88. J. C. Loudet and P. Poulin. Application of an electric field to colloidal particles suspended in a liquid-crystal solvent. *Phys. Rev. Lett.*, 87:165503, 2001.
89. J. L. Billeter and R. A. Pelcovits. Defect configurations and dynamical behavior in a Gay-Berne nematic emulsion. *Phys. Rev. E*, 62:711–717, 2000.
90. Denis Andrienko, Guido Germano, and Michael P. Allen. Computer simulation of topological defects around a colloidal particle or droplet dispersed in a nematic host. *Phys. Rev. E*, 63:041701/1–8, 2001.
91. S. V. Shiyankovskii and O. V. Kuksenok. Structural transitions in nematic filled with colloid particles. *Mol. Cryst. Liq. Cryst.*, 321:489–500, 1998.
92. M. S. Al-Barwani, G. S. Sutcliffe, and M. P. Allen. Force between colloids in a nematic solvent. submitted.
93. S. V. Burylov and Y. L. Raikher. Orientation of a solid particle embedded in a monodomain nematic liquid crystal. *Phys. Rev. E*, 50:358–367, 1994.
94. M. J. Stephen and J. P. Straley. Physics of liquid crystals. *Rev. Mod. Phys.*, 46:617–704, 1974.
95. Denis Andrienko, Michael P. Allen, Gregor Skačej, and Slobodan Žumer. Defect structures and torque on an elongated colloidal particle immersed in a liquid crystal host. *Phys. Rev. E*, 65:041702/1–7, 2002.

Self-Organization of Biomolecular Systems: Simulating the Folding and Aggregation of Peptides, Proteins and Lipids.

Alan E. Mark and Siewert-Jan Marrink

Department of Biophysical Chemistry, University of Groningen,
Nijenborgh 4, 9747 AG Groningen, The Netherlands
E-mail: mark@chem.rug.nl

Computer simulations of proteins, lipids and nucleic acids at equilibrium have become largely routine. The basic methodology is well established and has been extensively verified.¹ The challenge for the future is to use such approaches to understand composition and concentration depend phenomena such as those which give rise to spontaneous self-organization in (bio)molecular systems and which by their very nature cannot be readily probed by experiment. Examples of such processes include the folding and aggregation of peptides and proteins,^{2,3} the aggregation of lipids and surfactants into micelles^{4,5} bilayers⁶ and vesicles,⁷ as well as the assembly of mixed systems into functional complexes, all fundamental to living cells.

Peptide folding: It is not yet possible to fold complete proteins using atomic models. Nevertheless, dramatic process is being made regarding the folding and aggregation of peptides. For example, in certain cases it is now possible, to simulate the folding of small (10-15 a.a.) peptides under reversible conditions from an arbitrary initial configuration with experimental precision.^{2,3} Such work has lead to a reassessment of the nature of the unfolded state and point the way forward to the folding of larger systems.⁸⁻¹⁰

Lipid aggregation: Cell membranes are the archetypal example of a self-organized system which in addition to various types of lipids can contain peptide, proteins and high concentrations of molecules such as cholesterol. Previously it was believed that it was only possible to such collective behavior in such systems using simplified or coarse-grained models. We have recently shown however that it is possible to simulate the self-assembly for a random mixture of lipids (or surfactants) in water into their correct phase in atomic detail. This has proved to be an extremely powerful approach to investigate processes such as the pore formation in bilayers, domain formation in mixed bilayers, and the mechanism of vesicle formation and fusion.^{7,11} By performing simulations of the same phenomena using a specifically parameterized course-grained model as well as in atomic detail has also been possible to investigate for what aspects modeling in atomic detail is truly needed.

Assembly of functional complexes: The process of assembly of a membrane-water interface can also provide a powerful driving force to initiate the folding and self-assembly of peptides and proteins. For example the protein hydrophobin SC3 is largely unstructured in solution but rapidly forms extended β -sheet at hydrophobic/hydrophilic interfaces.¹² In the same manner multiple copies of pore forming peptides can be included in simulations leading to spontaneously formed bilayers. This is shedding light on exactly how such

peptides manage to insert, assemble and function within membranes.¹³

Examples illustrating the extent to which simulations both at an atomic level and at a more coarse grained level can be used to understand cooperative phenomena such as those mentioned above will be presented. In particular the lecture focus on to what extent the results we obtain from such simulations are reasonable¹⁴ and how basic traps when simulating biomolecular systems can be avoided.

References

1. W. F. van Gunsteren and A. E. Mark *Validation of molecular dynamics simulation*. J. Chem. Phys. **108**, 6109–6116 (1998).
2. X. Daura, B. Jaun, D. Seebach, W. F. van Gunsteren, and A. E. Mark *Reversible peptide folding in solution by molecular dynamics simulation*. J. Mol. Biol. **280**, 925–932 (1998).
3. X. Daura, K. Gademann, B. Jaun, D. Seebach, W. F. van Gunsteren, and A. E. Mark *Peptide folding: When simulation meets experiment*. Angew. Chemie Intl. Ed. **38**, 236–240 (1999).
4. S. J. Marrink, D. P. Tieleman, and A. E. Mark, *Molecular dynamics simulation of the kinetics of spontaneous micelle formation*, J. Phys. Chem. **104**, 12165–12173 (2000).
5. S. J. Marrink and A. E. Mark, *Molecular dynamics simulations of mixed micelles modelling human bile* Biochem. **41**, 5375–5382 (2002).
6. S. J. Marrink, E. Lindahl, O. Edholm, and A. E. Mark, *Simulation of the spontaneous aggregation of phospholipids into bilayers* J. Am. Chem. Soc. **123**, 12383–12391 (2001).
7. S. J. Marrink and A. E. Mark, *Molecular dynamics simulation of the formation, structure, and dynamics of small phospholipid vesicles* J. Am. Chem. Soc. **125**, 15233–15242 (2003).
8. H. Fan and A. E. Mark, *Refinement of homology based protein structures by molecular dynamics by molecular dynamics simulation techniques* Protein Science **13**, 211–220 (2004).
9. X. Daura, W. F. van Gunsteren, and A. E. Mark *Folding-unfolding thermodynamics of a beta-heptapeptide from equilibrium simulations*. Proteins: Struct. Funct. Genet. **34**, 269–280 (1999).
10. G. Colombo, D. Roccatano, and A. E. Mark *Folding and stability of the three-stranded β -sheet peptide betanova: Insights from molecular dynamics simulations*. Proteins: Struct. Funct. Genet. **46**, 380–392 (2002).
11. S. J. Marrink and A. E. Mark, *The mechanism of vesicle fusion as revealed by molecular dynamics simulations* J. Am. Chem. Soc. **125**, 11144–11145 (2003).
12. R. Zangi, M. L. de Vocht, G. T. Robillard, and A. E. Mark *Molecular dynamics study of the folding of hydrophobin SC3 at a hydrophilic/hydrophobic interface*. Biophysical J. **83**, 112–124 (2002).
13. G. Colombo, S. J. Marrink, and A. E. Mark, *Simulation of MscL gating in a bilayer under stress*. Biophys J. **84**, 2331–2337 (2003).
14. P. Soto and A. E. Mark, *The effect of the neglect of electronic polarization in peptide folding simulations*. J. Phys. Chem. B. **106**, 12830–12833 (2002).

Amphiphiles at Interfaces: Simulation of Structure and Phase Behavior

Friederike Schmid , Dominik Duchs , Olaf Lenz , and Claire Loison

¹ Fakultät für Physik, Universität Bielefeld, Postfach 100131, 33501 Bielefeld, Germany
E-mail: schmid@physik.uni-bielefeld.de

² Centre Européen de Calcul Atomique et Moléculaire, ENS Lyon,
46, Allée d'Italie, 69007 Lyon, France

Computer simulations of coarse-grained molecular models for amphiphilic systems can provide insight into the structure of amphiphiles at interfaces. They can help to identify the factors that determine the phase behavior, and they can bridge between atomic descriptions and phenomenological field theories.

Here we focus on model systems for amphiphilic membranes. After a brief general introduction, we present selected simulation results on monolayers, bilayers, and bilayer stacks. First, we discuss internal phase transitions in membranes and show that idealized models reproduce the generic phase behavior. Then we consider membrane fluctuations and membrane defects. The simulation data is compared with mesoscopic theories, and effective phenomenological parameters can be extracted.

1 Introduction

Amphiphilic molecules are characterized by the feature that they contain both water loving (hydrophilic) and water “hating” (hydrophobic) structural units. Familiar examples are alcohols or lipids (see Fig. 1). Amphiphiles are important for many applications in technology and nature. They are very effective at helping to dissolve different substances in water, which makes them very useful, e.g., as detergents, or as coating materials to stabilize colloidal systems. Furthermore, they form a rich variety of structures at higher concentrations. In order to shield their hydrophobic parts from the water, the amphiphiles self-assemble into spherical or cylindrical micelles or bilayers¹ (Fig. 2). These structures may then order on an even higher level and form superstructures (Fig. 3)^{2,3}.

Particular interesting from an application point of view are those phases where the material is filled with bilayers. Such bilayer interfaces can serve as barriers against the diffusion of particles, and help to divide the space into compartments. Indeed, lipid bilayers are the structural basis of all biological membranes, which in turn play a central role for the function of all cells and cell organelles⁴.

From an experimental point of view, studying the properties of biomembranes on a molecular scale *in situ* is not an easy task. Therefore, several model membrane systems have been developed: (i) monolayers at the air-water interface (Langmuir monolayers), (ii) stacks of bilayers, (iii) single planar bilayers, and (iiii) giant vesicles. The simplest and oldest approach is to spread lipid molecules on a water surface⁵. Due to the amphiphilic nature of lipids, they form oriented monolayers at air/water interfaces. These are traditionally placed in a Langmuir trough with a movable barrier on one side, which allows to control the surface area (see Fig. 4). The monolayers are then studied in the microscope, with γ -ray scattering under conditions of grazing incidence, or simply by measuring the

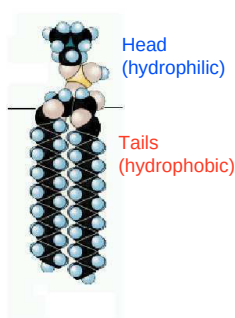


Figure 1. Structure of a single amphiphilic molecule (DPPC).

pressure-area isotherms. Since bilayers consist of two weakly coupled monolayers, many important bilayer properties can be studied already in monolayer systems.

Another experimental approach to studying lipid bilayer membranes has been to examine lamellar stacks of bilayers, i. e., lamellar phases of lipids⁶⁻⁸. Highly aligned stacks have been prepared experimentally and then examined by γ -ray diffraction^{7,8}. Such systems have also been useful to study the interactions of lipid membranes with other molecules, e. g., lipid-polymer interactions⁹, lipid-protein interactions^{10,11}, lipid-DNA interactions^{12,13}.

Furthermore, there also exist techniques by which planar bilayers or giant bilayer vesicles can be generated. γ -ray studies of such systems are difficult, due to the small size of the sample. However, they can be studied by other techniques, such as phase contrast microscopy or electron microscopy, light scattering, or transport measurements^{4,14}. It should be noted that isolated membranes are not stable in a thermodynamical sense. They are metastable structures, bound to break eventually. This, however, does not restrict their importance as model systems for biological membranes, which are themselves nonequilibrium structures.

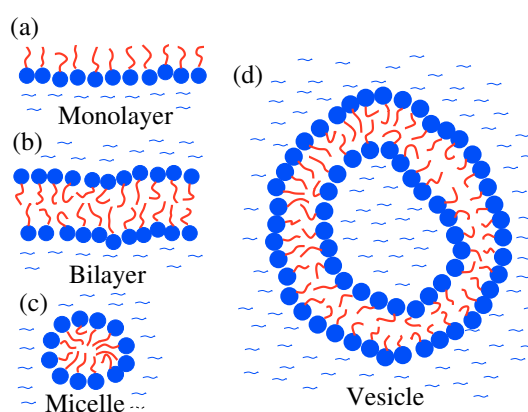


Figure 2. Self-assembled structures of amphiphiles in water.

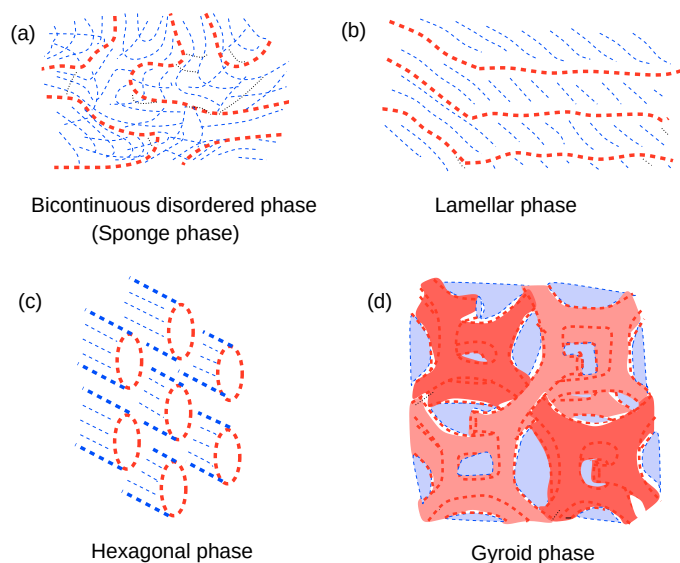


Figure 3. Selected structured phases in amphiphilic systems. The dashed lines represent bilayers. The different shadings in (d) distinguish between two different interwoven networks in the gyroid phase.

When attempting to describe lipid membranes theoretically, one faces a problem that is common for soft materials: The membrane properties result from an interplay between physical and chemical phenomena that live on a hierarchy of length scales, ranging from the atomic to the mesoscopic scale (micrometers), and likewise on a hierarchy of time scales. On the scale of atoms and small molecules (water, ions), one has the forces which keep the membrane together in the first place: the hydrophobic effect and the interaction between water and hydrophilic head groups, which is still not yet fully understood. On the molecular scale, one has the interplay between local chain conformations, membrane structure, and membrane viscosity and fluidity. We will come back to that aspect further below. One also has the electrostatic interactions between lipids. For bilayers under physiological conditions, they are shielded to some extent by the ions in the surrounding water. Nevertheless, they still influence the effective size of the head groups, thus stabilizing or destabilizing the planar lamellar structure. In monolayers, electrostatic interactions are not shielded and generate mesoscopic patterns^{15,16}. On a supramolecular length scale, one has

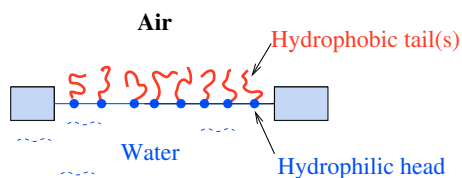


Figure 4. Schematic picture of a Langmuir trough.

the fluctuations of membrane positions. These are often described in terms of phenomenological parameters such as the bending stiffness and (if applicable) the surface tension^{2,17}. Finally, on the mesoscopic length scale, one has phase separation and domain formation. This becomes particularly important in *mixed* systems. For example, phase separation of different lipids within a membrane may trigger budding processes^{18–20}. Certain membrane lipids and proteins aggregate into “rafts”, which are believed to serve a biological purpose²¹. Polymers may induce membrane fusion²².

Computer simulations which cover all these aspects are not possible, and will not be possible for a long time. When studying membranes by computer simulation, the first and main task is therefore to *choose a problem* and then *identify an appropriate model*. Different models operate on different length and time scales. Ideally, it should be possible to establish connections between the different models (“multiscale modeling”). The problem of bridging between time and length scales is one of the great challenges in today’s computational science. We are still far from having reached that goal. However, there exist several models which can be used to study membranes on particular length scales, and computer simulations of these models can contribute to an improved understanding of physical phenomena on that scale.

In fact, several approaches are introduced in this school. Ole Mouritsen presents lattice models which allow to study the mesoscopic organization of complex biomembranes. Alan Mark takes the opposite perspective and studies membranes on an atomistic level. Here, we will assume an intermediate point of view and discuss molecular coarse grained models. Such models can be used to study the structure of a system systematically for a range of parameters. For example, they reproduce self-assembly and different lipid-water phases, as well as internal phase transitions in lipid membranes. Numerical studies provide insight into mechanism which contribute to the phase behavior. Moreover, molecular coarse grained models can be used as starting points for simulations that bridge between the molecular and the mesoscopic level. An example of such a study will be presented in Sec. 3.3.

The paper is organized as follows. We will begin with a brief discussion of typical models that are used to study membranes. In Sec. 3, we present some applications of coarse grained molecular models to study monolayers, bilayers, and bilayer stacks. We conclude in Sec. 4.

2 Models and Levels of Coarse Graining

As discussed in the introduction, it is not possible to define a single simulation model that is suited to study all relevant aspects about amphiphile membranes in one single computer simulation. Therefore, a hierarchy of models has been developed and used to study membranes from different point of views.

The models at the lowest level of the hierarchy are the atomistic models. Even these are not truly *ab-initio* in the sense that the molecules are not treated in full quantum chemical detail. Atomistic molecular dynamics simulations of amphiphilic systems use semiempirical potentials such as the GROMOS force field (see Alan Mark’s contribution), which are constructed by fitting the parameters to quantum chemical calculations *and* to experimental data. Thus they focus on the realistic description of specific substances. The time and length scales accessible to such a simulation are limited, but they increase very rapidly due

to the improving performance of the computer hardware, and to the development of good algorithms. Nowadays, atomistic simulations can deal with thousands of amphiphiles, or reach time scales up to microseconds.

The next level of coarse graining is that of idealized molecular models. Here, different molecules are still treated as individual particles, but their structure is very much simplified. Atomic and molecular details are largely disregarded. Only the features that are essential for the behavior of the molecules are kept. In the case of amphiphiles, one characteristic attribute is clearly the dual character of the molecules, with their separate water-loving and water-hating parts. Another quantity that seems to be important for the self-organization of the amphiphiles is the conformational entropy of the molecules. Many coarse-grained amphiphile models concentrate precisely on these two aspects of the molecules, and represent amphiphiles by chains of simple “water”-loving or “water”-repelling units. Coarse grained models are particularly suited to study *generic* properties of amphiphiles. They can be regarded as simple, minimal systems that provide general insight into the mechanisms that drive the self-assembly and the phase behavior of amphiphiles.

To optimize computational efficiency, many coarse grained models have been formulated on a lattice. A particular popular lattice model has been introduced about twenty years ago by R. G. Larson *et al.*²³. Water molecules () occupy single sites of a cubic lattice, and amphiphile molecules are modeled by chains of “tail” monomers and “head” monomers , which are taken to be identical to -particles. Only particles that are neighbors on the lattice interact with each other. The lattice is entirely filled by , , or particles. The interaction energy is thus determined by a single interaction parameter, which describes the relative repulsion between and or . The model can be simulated very efficiently by Monte Carlo methods. It exhibits an amazingly rich phase behavior²⁴. Even a gyroid structure can be observed under certain circumstances²⁵.

Despite the success of lattice models, they still have the obvious flaw of imposing an *a priori* anisotropy on space. Therefore, off-lattice models are attracting growing interest. To our best knowledge, the first amphiphile model of this kind was introduced by B. Smit *et al.*²⁶ in 1990. The general idea is similar to that of R. G. Larson. The amphiphiles are represented by chains made of very simple - or -units, which are in this case spherical beads. The “water” molecules are represented by free beads. Beads interact via simple short-range pairwise potentials, often truncated Lennard-Jones potentials. The parameters of the potentials are chosen such that pairs and pairs effectively repel each other. Such a model reproduces self-assembly, micelle and membrane formation.

Many similar amphiphile models have been defined and applied to study various problems. Some examples from our own work will be presented in the next section (Sec. 3).

The models discussed so far still treat amphiphiles as flexible chains. One might argue that the chain flexibility is not absolutely essential for the character of an amphiphilic system, and that a truly “minimal” model should ignore it. Indeed, “molecular” models that restrict themselves to the very basic ingredients have been efficient tools to study certain aspects of amphiphilic self-organization. For example, spin models which include just the orientation of amphiphiles and the repulsion between one molecule end and water have reproduced topological characteristics of amphiphilic phase behavior². A particularly successful class of lattice models has been designed specifically to model lipid membranes^{27–29}, and has been used to study various complex biomembranes at equilibrium and even nonequilibrium. This approach will be presented in Ole Mouritsen’s lecture.

Finally, the highest level of coarse graining is reached with the phenomenological models. These drop the notion of single particles entirely and describe amphiphilic systems by continuous fields, with a free energy functional that is governed by a few mesoscopic material parameters. For example, Ginzburg-Landau models² introduce a free energy functional, which depends on local amphiphile and water concentrations. In contrast, random interface models^{17,30} concentrate on the amphiphilic sheets, which are parametrized and characterized in terms of mechanical elasticity parameters such as the bending moduli. Mesoscopic models are good starting points for analytical approaches. Thus computer simulations of such models may also serve to test or to complement an analytical theory.

After this brief overview over different types of models for amphiphilic systems, we will now focus on coarse grained molecular models. In the next section, we will show how computer simulations of such models can help to understand the structure and the phase behavior of amphiphilic monolayers and bilayers.

3 Applications to Amphiphilic Layers

This chapter will address two different aspects of the structure of amphiphilic layers: Internal phase transitions, and fluctuations and defects in membrane stacks. It is by no means a complete overview over all simulation studies that have dealt with these and related issues. Instead, it mostly focusses on some of our own work, which is hopefully suited to illustrate the potential and the merits of molecular coarse-grained simulations.

Lamellar stacks of lipids in water assume several different structures. At high temperatures and low pressures, they are usually in the “fluid” phase, which is characterized by a low degree of conformational order and a high mobility of lipids within the membranes. Upon lowering the temperature, one encounters a first order transition – the “main transition” – to a state with higher conformational order, and lower mobility: a “gel” state. In the gel region, there exist different phases. For example, the chains may display collective tilt with respect to the surface normal, they may show local hexatic order, the lamellae may even exhibit asymmetric wavy undulations (ripple phase). Most biomembranes in living organisms are maintained in the fluid state. Nevertheless, the main transition has presumably some relevance for biological systems, as it occurs at temperatures close to the body temperature for some common bilayer lipids (e.g., 41 °C in DPPC). It will be considered in the subsections 3.1 and 3.2.

The internal ordering of the lipids is one important structural property of a membrane. Another is the spectrum of shape fluctuations, and the frequency and structure of membrane defects. Membrane defects determine critically the permeability of membranes. Moreover, the formation of point defects is believed to play a crucial role as a first step in the process of membrane fusion³¹. Membrane fluctuations and membrane defects have been described with phenomenological approaches^{32–35}, and these theories were used to interpret experimental results. In computer simulations, the local structure can be investigated in much more detail than in experiments. Therefore comparisons of phenomenological theories with molecular simulations are clearly of interest. An example of such a comparison will be presented in the subsection 3.3.

3.1 Monolayers

We begin with discussing the phase behavior in monolayers. It is closely related to that of bilayers, which is one of the reasons why Langmuir monolayers are considered to be useful model membrane systems. In particular, the main transition has a prominent monolayer equivalent – a first-order phase transition between a “liquid expanded” state and a denser “liquid condensed” state. As in membranes, the condensed state exists in several modifications, which differ, among other, by tilt order and positional order of the head groups, or by backbone order of the chains.

Experimentally, monolayer phase diagrams are often determined as a function of the temperature and the area per molecule, or alternatively, the spreading pressure of the molecules on the surface. The phase diagrams for different amphiphilic molecules are very similar. As an example, Fig. 5 shows a generic phase diagram of fatty acid monolayers^{5, 36–38}.

In order to describe such systems on a coarse-grained level, we model the amphiphiles as chains of “tail” beads with diameter σ , attached to a slightly larger “head” bead of diameter σ_h . The water surface is represented by a planar surface at $z = 0$, which repels the tail beads, and attracts the head bead.

More specifically, the model is defined as follows: The beads in the chains are connected by nonlinear springs with the potential

$$V(r) = \begin{cases} -\frac{\epsilon}{r} & \left[1 - \left(\frac{r}{\ell} \right)^2 \right] & \text{for } |r| \leq \ell \\ \infty & \text{for } |r| > \ell \end{cases} \quad (1)$$

where ℓ is the length of the spring. This potential is constructed such that it is nearly harmonic in the center, at $r \approx 0$, but diverges at $r = \pm \ell$. Thus the distance between neighbor beads cannot take arbitrarily large values, which ensures that chains cannot cross each other. A potential with a logarithmic cutoff as in (1) is often called FENE-potential

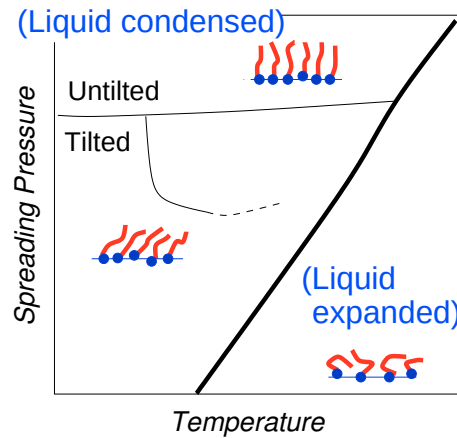


Figure 5. Generic phase diagram for fatty acid monolayers. (after Ref. 5). LE is the liquid expanded phase. In the L_2 phase and the Ov phase, the chains are tilted towards next-nearest and nearest neighbors, respectively. In the LS phase, they are on average untilted. Further phases are found at lower temperatures (not shown).

(Finite Extendible Nonlinear Elastic potential). In addition to this spring potential, chains are given a bending stiffness by virtue of a stiffness potential

$$\cdot (1 - \cos \theta) \quad (2)$$

which acts on the angle θ between subsequent springs and favors $\theta = 0$ (straight chains). The parameter k_b can then be tuned to control the conformational freedom of the chains and study the role of the chain entropy for the phase behavior.

Two beads i and j ($i \neq j$ or $j \neq i$) that are not direct neighbors in the same chain interact with a truncated and shifted Lennard-Jones potential,

$$V_{ij}(r) = \begin{cases} \left(\left(\frac{r_0}{r} \right)^{12} - 2 \left(\frac{r_0}{r} \right)^6 \right) & \text{for } r \leq r_c \\ 0 & \text{for } r > r_c \end{cases} \quad (3)$$

where $r_0 = 1.12$ nm, the cutoff is $r_c = 2.5$ nm for the tail beads, and $r_c = 1.5$ nm for the other interactions, and the shifting parameter r_0 is chosen such that $V_{ij}(r)$ is continuous at r_c . It is easy to see that with this choice of the cutoff parameters, the interactions between tail beads are attractive, and all other interactions are purely repulsive.

Finally, we have to specify the interactions of beads with the surface. We have explored two types of surface potentials: First we used a model introduced by Haas *et al*³⁹, where the surface imposes rigid constraints: The head beads are confined to stay in the plane at $z = 0$, and the tail beads are not allowed into the half-space $z < 0$ ^{40,41}. This model has the advantage of being very simple, however, it is unphysical in the sense that real water surfaces are not sharp on an atomic scale. Indeed, better agreement with the experimental phase diagram can be reached with softer surface potentials⁴². Here, head beads are subject to the potential

$$V_{ij}(z) = \begin{cases} 0 & \text{for } z \geq 0 \\ -\frac{1}{2} \left[1 - \left(\frac{z}{0.5} \right)^6 \right] & \text{for } -0.5 \leq z < 0 \end{cases} \quad (4)$$

and tail beads to

$$V_{ij}(z) = \begin{cases} -\frac{1}{2} \left[1 - \left(\frac{z}{-0.5} \right)^6 \right] & \text{for } -0.5 \leq z < 0 \\ 0 & \text{for } z \geq 0 \end{cases} \quad (5)$$

with $\sigma = 1$ nm and $\epsilon = 10$ kJ/mol. It turns out that the exact form of the surface potential is not important, almost identical results are obtained with $\epsilon = 2$ kJ/mol.

The other model parameters were $\sigma = 0.7$ nm, $\epsilon = 0.2$ kJ/mol, and $\epsilon = 100$ kJ/mol. The stiffness parameter k_b was varied between 4.5 and 100 kJ/mol. The head size was chosen 1.1 or 1.2 nm, and the chain length was 7 in most simulations.

This model was studied by Monte Carlo simulations at constant spreading pressure Π : We considered N chains with heads grafted in a planar parallelogram of variable size and shape, characterized by two side lengths a and b and one angle α . The boundary conditions were periodic in the x plane, and free in the y direction. The simulation algorithm included three elementary Monte Carlo moves:

- Single particle displacements

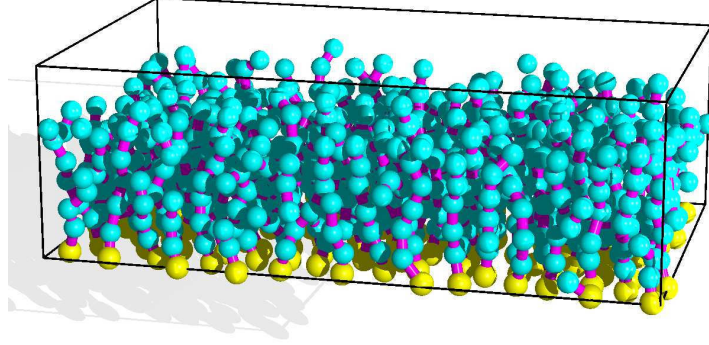


Figure 6. Snapshots of a monolayer in a system with $\sigma_h = 11\sigma_t$, $k_A = 1 \text{ } \epsilon$ and rigid constraints at the surface. ($T = 4 \text{ } \epsilon/k_B$ and $P = 5 \text{ } \epsilon/\sigma_t^2$) (Courtesy of C. Stadler).

- Variations of L_x or L_y , *i. e.*, the whole configuration is stretched or squeezed in one direction. Care must be taken that this move satisfies detailed balance. For example, a move of the form $(x, y) \rightarrow (x + \delta x, y + \delta y)$, where δx is uniformly distributed, cannot be applied without introducing a correction term in the probability for accepting the move. In contrast, the move $(x, y) \rightarrow (x + \delta x, y)$ can be applied without correction. In our simulations, we varied L_x in an additive way, $L_x \rightarrow L_x + \delta L_x$ (rejecting moves that made negative).
- Shearing the box, *i. e.*, variation of θ while keeping the total area constant.

The moves were accepted or rejected according to the Metropolis prescription^{44,45,47} with the effective Hamiltonian

$$\mathcal{H} = U - P A \quad (6)$$

where U is the (internal) energy, and A is the area of the system. To speed up the simulations, we used cell lists and Verlet lists^{46,47}. Fig. 6 shows an example of a configuration snapshot.

To analyze the properties of the monolayer, one needs to define and monitor appropriate observables. One quantity of interest is the area per molecule. Fig. 7 shows typical temperature-area isobars. One clearly discerns a discontinuous jump, which moves to higher temperatures with increasing pressure.

The nature of this phase transition can be studied in more detail by inspecting the radial pair correlation functions and the hexagonal order parameter of two-dimensional melting,

$$\left\langle \left| \frac{1}{6} \sum_i \sum_{j=1}^6 e^{i 6 \theta_{ij}} \right| \right\rangle \quad (7)$$

Here the sum $\sum_i$ runs over all heads of the system, the sum $\sum_{j=1}^6$ over the six nearest neighbors of i , and θ_{ij} is the angle between the vector connecting the two heads and an arbitrary reference axis. The data for $\langle \left| \frac{1}{6} \sum_i \sum_{j=1}^6 e^{i 6 \theta_{ij}} \right| \rangle$ which correspond to Fig. 7 are shown in Fig. 8. The hexagonal order parameter is nonzero in the low temperature phase, and jumps to a value

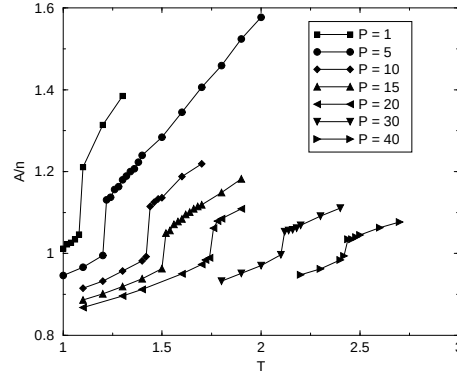


Figure 7. Area per molecule (in units σ_t^2) vs. temperature (in units ϵ/k_B) in a system with $k_A = 4.7\epsilon$, $\sigma_h = 1.1\sigma_t$ and soft surface potentials. The spreading pressures are given in units of ϵ/σ_t^2 . From Ref. 43.

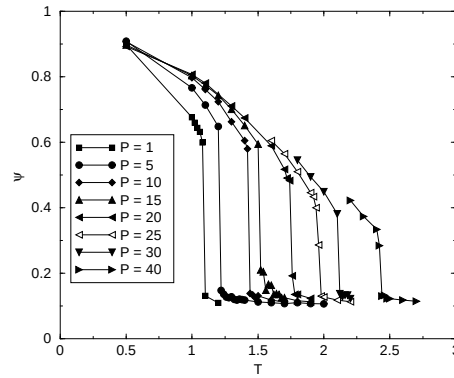


Figure 8. Order parameter S_6 vs. temperature (units ϵ/k_B) for the same system as Fig. 7. From Ref. 43.

close to zero at the transition. This indicates that the high temperature phase is liquid and the low temperature phase is hexatic or (quasi)crystalline.

A closer inspection of the ordered low-temperature region reveals that it can be subdivided in at least two regimes: At low pressures, the chains are collectively tilted in one direction, and at high pressure, they are on average untilted. The azimuthal symmetry breaking due to collective tilt order can be characterized by the order parameter

$$\sqrt{\langle x^2 \rangle} \quad (8)$$

where x and y denote the x and y components, respectively, of the head-to-end vector of a chain, averaged over all chains in *one* configuration, and $\langle \cdot \rangle$, the statistical average over all configurations. The values of $\sqrt{\langle x^2 \rangle}$ as a function of the pressure on two different isotherms are shown in Fig. 9.

The data for different temperatures and spreading pressures can be summarized in a phase diagram. An example for the system with soft potentials is shown in Fig. 10. This

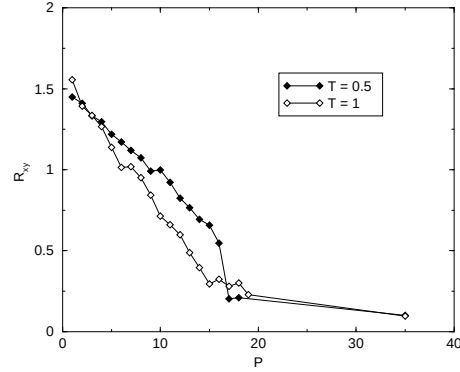


Figure 9. Tilt parameter R_{xy} vs. pressure (in units ϵ/σ_t^2) at two temperatures T (in units ϵ/k_B) for the same system as Fig. 7. From Ref. 43.

phase diagram has great similarity with the high temperature part of the experimental phase diagram in Fig. 5. At high temperature, we observe a disordered liquid expanded phase (LE). At low temperature two types of condensed phases are present – an untilted high-pressure phase (LS) and a tilted low-pressure phase (L/Ov). The temperature of the order-disorder transition increases with pressure, as in the experiment, and the pressure of the tilting transition between LS and L/Ov is almost independent of the temperature, again as in the experiment.

Thus we have established a minimal model which reproduces the main features of Langmuir monolayers in the vicinity of the main transition. At lower temperatures, the experimental phase diagram of fatty acids is much more complex than that of our model.

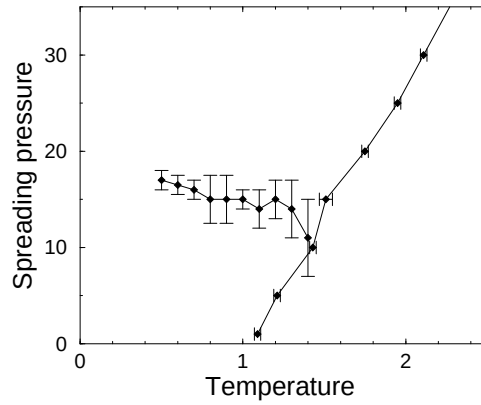


Figure 10. Phase diagram from Monte Carlo simulations of the monolayer model with $k_A = 4.7\epsilon$, $\sigma_h = 1.1\sigma_t$ and soft surface potentials. soft surface potentials in the plane of spreading pressure (units $\epsilon/k_B\sigma_t^2$) vs. temperature (units ϵ/k_B). The chain stiffness is $k_A = 4.7\epsilon$. LE denotes disordered phase, LS untilted ordered phase, and L₂/Ov ordered phase with tilt towards nearest or next-nearest neighbors (undetermined). From Refs. 42, 43

This is not surprising. To reproduce phenomena like backbone ordering, the chains have to be modeled in much more detail than has been done here. On the other hand, the details of the low temperature portion of the experimental phase diagram depend strongly on the particular choice of the amphiphile. Our model was never designed to describe specific properties of fatty acids. It was designed to reproduce the main transition in amphiphile monolayers. This is achieved in a satisfactory way.

We can learn more about the main transition by playing with the model parameters and studying how this influences the phase behavior. For example, increasing the chain stiffness shifts the order/disorder transition to higher temperatures and the tilting transition to higher pressures⁴⁰. The form of the surface potential (rigid or soft) is not important for the order/disorder transition, but it does influence the slope of the line of tilting transitions in pressure-area space⁴². Likewise, increasing the head group size by 10 % does not influence the order/disorder transition very much, but may affect the tilted phases quite dramatically⁴¹. All these observations can be summarized by the statement that the order/disorder transition is basically driven by the chains, whereas the tilting transitions result from an interplay between chains and head groups.

3.2 Bilayers

As mentioned earlier, lipid bilayers exhibit internal phase transitions which are very similar to those observed in monolayers. In other respect, however, they are fundamentally different. They do not form by adsorption on a pre-defined surface, instead, the lipids self-assemble spontaneously to a structure with planar geometry. As a result, bilayers are not strictly planar, but may curve around and undulate. In order to study membranes, one needs to define a model which reproduces self-assembly.

How can bilayer self-assembly be modeled? The simplest approach is to force amphiphiles into sheets by tethering the head groups to two dimensional surfaces⁴⁸⁻⁵¹. "Self-assembly" is then enforced by external constraints, with the obvious consequence that the head groups lose much of their translational degrees of freedom. In contrast, real self-assembled structures are held together by an interplay of amphiphile-amphiphile and amphiphile-solvent interactions. Coarse-grained models that reproduce spontaneous self-assembly must account for the presence of solvent one way or another. However, the solvent should be modeled with as little detail as possible, since the focus is still on the bilayers.

One possibility is to use the Smit model²⁶ mentioned in the introduction. The amphiphiles are modeled by chains of beads, and the solvent by beads of the same size. This model indeed reproduces bilayer self-assembly⁵² and has been applied successfully by Goetz *et al.* to study shape fluctuations of model bilayers⁵³.

Unfortunately, even the simple Smit model still has the drawback that a substantial amount of computer time in simulations is spent on the uninteresting solvent. It is therefore desirable to have a model which does not include the solvent explicitly. This idea is not particularly eccentric, such models are commonly used in simulations of polymers in solvent. The effect of the solvent is then incorporated in the effective interactions between monomers. However, it is not *a priori* clear that effective (pair) interactions will be able to bring about something as complex as membrane self-assembly.

Indeed, it was only very recently that O. Farago⁵⁴ established a solvent-free molecular model for membranes. Amphiphiles are represented by rigid linear trimers, made of

beads which interact through truncated Lennard-Jones interactions with carefully chosen interaction parameters. O. Farago showed that single bilayers remain (meta)stable in his model, that they exhibit an order-disorder phase transition as a function of the molecular area density, and that they even sustain pore formation. However, he also mentions that a lengthy “trial and error” process of fine tuning the Lennard-Jones parameters was necessary to achieve this impressive result.

In our own work⁵⁵, we propose a different, more robust approach, the “phantom solvent” model. Solvent particles are treated explicitly, but they interact *only* with the amphiphiles, not with one another. The amphiphiles perceive the solvent particles as soft beads of radius σ_s . The bulk of the solvent region is simply an ideal gas. The amphiphiles are modeled exactly in the same way as in our previous monolayer model, except that the surface potentials (4) and (5) are of course eliminated and replaced by periodic boundary conditions. The model can be implemented in a straightforward way and studied by Monte Carlo simulations. We found that it produced stable bilayers at the first go, without any parameter tuning. Two examples of snapshots are shown in Fig. 11.

The phantom solvent model has several advantages:

- It is computationally very efficient. The computational time spent on the phantom solvent is only a few percent of that spent on the amphiphiles, for any reasonable number of solvent beads. This is because only few pair interaction potentials have to be evaluated per solvent bead.
- The solvent has no local liquid structure. This is good, because we are not interested in the interplay between solvent and bilayer structure. If we intended to study this aspect, we would have to model the solvent (water!) in much more detail. Moreover, a structured solvent would introduce unwanted effective interactions between the bilayer and its periodic images in the normal direction.

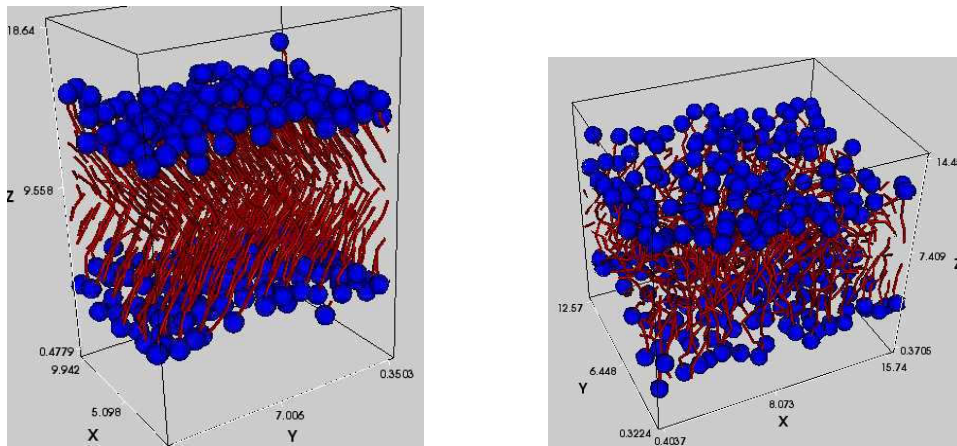


Figure 11. Snapshots of bilayers embedded in phantom solvent at pressure $P = 1 \epsilon/\sigma_t^3$ in the ordered state (left, temperature $T = 1.0 \epsilon/k_B$) and in the fluid state (right, $T = 1.5 \epsilon/k_B$). Solvent particles are not shown. ($\sigma_h = 1.1\sigma_t$, $k_A = 4.7\epsilon$)

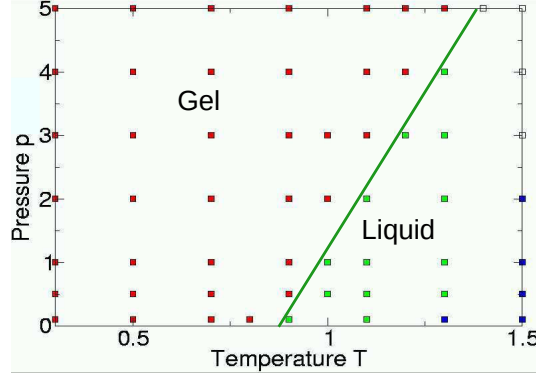


Figure 12. Preliminary phase diagram of bilayers in the phantom solvent model in the plane pressure (units ϵ/σ_t^3) vs. temperature (units ϵ/k_B). The model parameters are the same as in Fig. 10. Red points denote ordered membrane, and green points fluid membrane. At high temperatures (blue points), the membrane disintegrates.

- The phantom solvent has a simple physical interpretation. It probes the total free volume that is available to the solvent on the length scale σ_t . The self-assembly in our model is driven by the attractive interactions between the tails, and by the entropic effect that the solvents have more space when the amphiphiles are clustered together.

At the moment, the model still shares the handicap of all solvent-free models (and lattice models), that hydrodynamical interactions between the membranes and the solvent cannot be treated very easily. It might be possible to formulate dynamical equations for the phantom solvent which ensure that it behaves like a liquid and not like a gas. Otherwise, those who intend to study hydrodynamic effects might have to equip the solvent particles with a weak integrable potential, as is done in dissipative particle dynamics simulations. As long as we calculate static properties with Monte Carlo simulations, this is however not a problem.

After these general remarks, we turn back to the discussion of our particular amphiphile model. As mentioned before, it produced stable bilayers over a wide parameter range. The configuration snapshots of Fig. 11 demonstrate that the bilayers exhibit a low-temperature ordered phase and a high-temperature disordered phase. At even higher temperatures, they disintegrate. The phase transitions can be monitored, e.g., by inspecting the total nematic order parameter of amphiphiles or the area per lipid as a function of the pressure and temperature. A preliminary phase diagram is shown in Fig. 12. We note the similarity to the monolayer phase diagram of Fig. 10. This corroborates the assertion that monolayers are good model systems for membranes.

3.3 Bilayer Stacks

Having discussed internal phase transitions in monolayers and bilayers, the logical next step would be to consider internal phase transitions in lamellar stacks. They exist, of course, but they are presumably not very different from those in single bilayers. Thus we will shift focus and concentrate on other aspects of membranes in this section: shape fluctuations and defects^{56–58}.

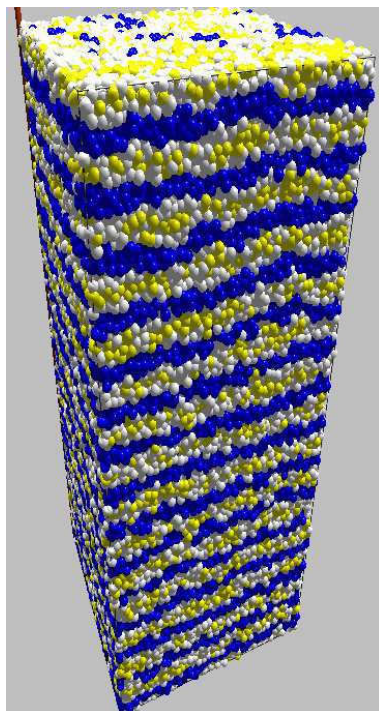


Figure 13. Snapshot of a bilayer stack with 30720 amphiphiles and 30720 solvent beads. The “hydrophobic” tail beads are blue, the “hydrophilic” head beads are white, and the solvent beads are yellow.

We consider a binary mixture of amphiphiles and solvent beads. The model shall not be described in full detail here. It was originally introduced by Soddemann *et al*^{59,60} and has some similarity with the Smit model. The elementary units are spheres with a hard core radius r_c , which may have two types: “hydrophilic” or “hydrophobic”. Beads of the same type attract each other at distances less than $1.5 r_c$. “Amphiphiles” are tetramers made of two hydrophilic and two hydrophobic beads, and “solvent” particles are single hydrophilic beads.

The pressure and the temperature are chosen such that the system is in a fluid lamellar phase. More specifically, we study configurations with 5 or 15 stacked bilayers, which are swollen with 20 volume percent solvent. The simulations were done with constant pressure molecular dynamics, using a Langevin thermostat to maintain constant temperature. The constant pressure algorithm was designed such that the box dimensions parallel and perpendicular to the lamellae fluctuate independently, in order to ensure that the pressure is isotropic and the membranes have no surface tension. A configuration snapshot is shown in Fig. 13. It was produced by equilibration of an initial configuration with lamellar order, set up such that the lamellae were oriented perpendicular to the long axis (the z -axis) of the simulation box. We have checked with smaller systems that the lamellae self-assemble spontaneously if the initial configuration is disordered.

Our study aimed at analyzing shape fluctuations of the membranes and defects in the

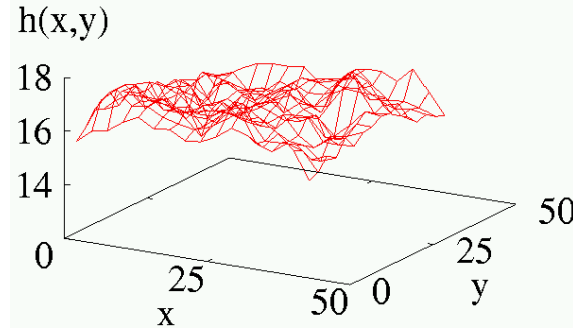


Figure 14. Typical conformation of a membrane position $h_n(x, y)$.

membranes. Therefore, the first nontrivial task was to determine the local positions of the membranes in the lamellar stack. This was done as follows:

- The simulation box was divided into $N_x \times N_y$ cells. ($N_x = N_y = 32$). Note that the size of the cells may vary between configurations because the dimensions of the box fluctuate in a constant pressure simulation.
- In each cell, the relative density of tail beads $\rho_{\text{tail}}(x, y)$ was calculated. It is defined as the ratio of the number of tail beads and the total number of beads.
- The hydrophobic space is defined as the space where the relative density of tail beads is higher than a given threshold ρ_{th} . The value of the threshold is roughly 0.7 (80 % of the maximum value of ρ_{tail} and depends on the mesh size).
- The cells that belong 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. This algorithm identifies membranes even if they have pores. At the presence of other membrane defects, additional steps have to be taken. (This happened very rarely in our system).
- For each membrane m and each position (x, y) , the two heights $h_{\text{min}}^m(x, y)$ and $h_{\text{max}}^m(x, y)$, where the density $\rho_{\text{tail}}(x, y)$ passes through the threshold ρ_{th} , are determined. The mean position is defined as the average $h_m(x, y) = (h_{\text{min}}^m + h_{\text{max}}^m) / 2$.

The algorithm identifies membranes even if they have pores. A typical membrane conformation $h_m(x, y)$ is shown in Fig. 14. The statistical distribution of $h_m(x, y)$ can be analyzed and compared with theoretical predictions.

One of the simplest mesoscopic theories for fluctuations in membrane stacks is the “discrete harmonic model”³³. It describes stacks of membranes without surface tension and assumes that the fluctuations are governed by two factors: The bending stiffness of single membranes, and the penalty for compressing the membranes, characterized by a compressibility modulus κ . The free energy is given in harmonic approximation

$$\mathcal{F}_{\text{DH}} = \sum_{\mathbf{q}} \int d\mathbf{r} \left\{ \frac{1}{2} \left(\frac{\partial h}{\partial \mathbf{r}} \right)^2 + \frac{\kappa}{2} (h - \bar{h})^2 \right\} \quad (9)$$

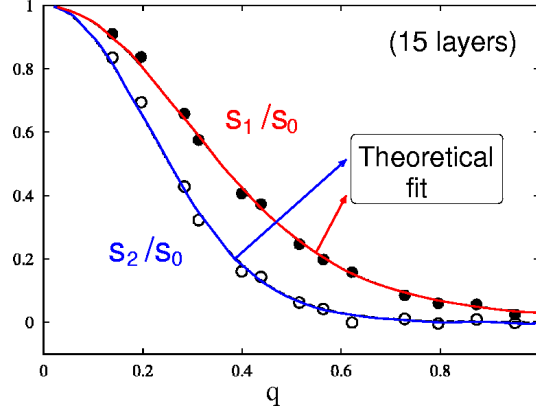


Figure 15. Ratios of transmembrane structure factors s_1/s_0 and s_2/s_0 vs. in-plane wavevector q in units of σ^{-1} . The solid lines correspond to a theoretical fit to Eq. (10) with one (common) fit parameter K_c/B .

where $\bar{d}$ is the average distance between layers. This free energy functional is simple enough that statistical averages can be calculated exactly. For example, the transmembrane structure factor which describes correlations between membrane positions in different membranes is given by⁵⁷

$$\langle s(\mathbf{q}) \rangle \langle s(\mathbf{q})^* \rangle = \langle s(\mathbf{q}) \rangle \left[1 - \frac{1}{2} + \frac{1}{2} \sqrt{1 - 4\bar{d}^2 q^2} \right] \quad (10)$$

where $s(\mathbf{q})$ is the Fourier transform of $s(x)$ in the (x) -plane and $\bar{d}$ is a dimensionless parameter. The function $s(x)$ can also be given explicitly⁵⁷. We can test the prediction (10) by simply plotting the ratio s_1/s_0 vs. q for different $\bar{d}$. The functional form of the curves should be given by the expression in the square brackets, with only one fit parameter K_c/B . Fig. 15 shows our simulation data. The agreement with the theory is very good over the whole range of q . Hence the discrete harmonic model, a mesoscopic theory, describes the membrane fluctuations in our molecular way in a satisfactory way.

Moreover, our analysis provides a value for the phenomenological parameter K_c/B . By combining it with the analyses of other quantities, it is possible to determine K_c and B separately. This establishes a bridge between the molecular and the mesoscopic description.

Next we turn to discuss the membrane defects in the lamellar stack. On principle, one can have three types of topological point defects in membranes: necks, pores, and passages. In our system, necks and passages were extremely rare, and we did not collect enough data to be able to analyze them statistically. Thus we will focus on the pores here.

As we have already mentioned, the properties of pores determine the permeability of a membrane. A number of atomistic and coarse grained simulation studies have therefore addressed pore formation^{31,61–63}, mostly in membranes under tension. In contrast, the membranes in our lamellar stack have no surface tension. As we shall see, this affects the characteristics of the pores quite dramatically.

Fig. 16 shows a snapshot of a hydrophobic layer which contains a number of pores. These pores have nucleated spontaneously. They “live” for a while, grow and shrink with-

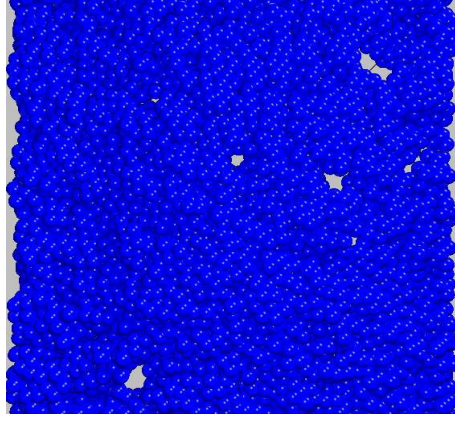


Figure 16. Snapshot of the hydrophobic beads in a single bilayer (top view).

out diffusing too much, until they finally disappear. Most pores close very quickly, but some large ones stay open for a long time.

We will first analyze the local membrane structure in the vicinity of a pore. Fig. 17 shows average composition profiles of tail beads, head beads and solvent beads as a function of the in-plane distance from the center of the pore, r , and the normal out-of-plane coordinate, z . These averages were performed with data from pores with surface areas between 4 and 16 σ^2 . The pictures demonstrate that the amphiphiles rearrange themselves at the pore edge, so that solvent beads in the pore center are exposed mainly to head beads. Such a pore is called hydrophilic. In previous simulations, both hydrophilic and hydrophobic pores have been reported, depending on the system under consideration.

Whereas the local composition profiles at the pore edge depend on the model, other structural properties of pores on larger scales are presumably generic and can again be described by simple mesoscopic theories. For example, the simplest Ansatz for the free energy of a pore with the area A and the contour length L has the form³⁴

$$E = E_c + \gamma L + \sigma A \quad (11)$$

where E_c is a core energy, γ a material parameter called line tension, and σ the surface tension. The last term accounts for the reduction of energy due to the release of surface

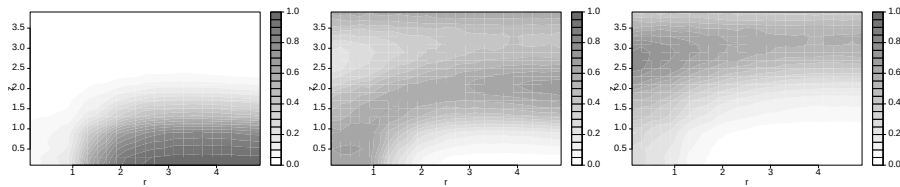


Figure 17. Density profiles of tail beads, head beads, and solvent beads as a function of the distance from the center of the pore in radial in-plane direction r and normal direction z .

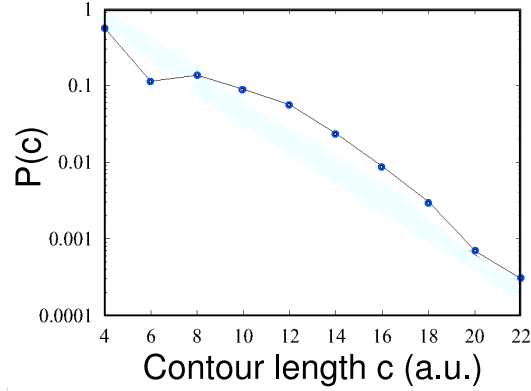


Figure 18. Distribution of pore contours in a semi-logarithmic plot.

tension in a stretched membrane. In our case, the surface tension is zero ($\gamma = 0$), and the last term vanishes. The second term describes the energy penalty at the pore rim. If this simple free energy model is correct, the pore shapes should be distributed according to a Boltzmann distribution

$$P(c) \propto \exp(-\beta \epsilon_{\text{rim}}(c)) \quad (12)$$

This can be tested in the simulation. Fig. 18 shows a histogram of contour lengths c . The bare data do not reflect the expected exponential behavior. Either the Ansatz (11) is wrong, or we have not used it correctly.

Indeed, a closer look reveals that Eq. (12) disregards an important effect: The “free energy” (11) gives only *local* free energy contributions, 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 c , and test the relation

$$P(c) \propto \Omega(c) \exp(-\beta \epsilon_{\text{rim}}(c)) \quad (13)$$

Fig. 19 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 γ . Thus we have again established a bridge between the molecular simulation model and a mesoscopic theory.

If the model (11) is correct, it 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 ring polymers. In particular, they are not round, but have a fractal structure. From polymer theory, one knows that the size R of a two dimensional self-avoiding polymer scales roughly like $R \propto c^\nu$ with the polymer length c . In our case, the “polymer length” is the contour length c . Thus the area A of a pore should scale like

$$A \propto c^{2\nu} \quad (14)$$

Fig. 20 shows that this is indeed the case.

Many other properties of the pores can be investigated, *e. g.*, pore distributions, the dynamical evolution of pores, pore life times^{56,58}. Nevertheless, we shall stop here. We

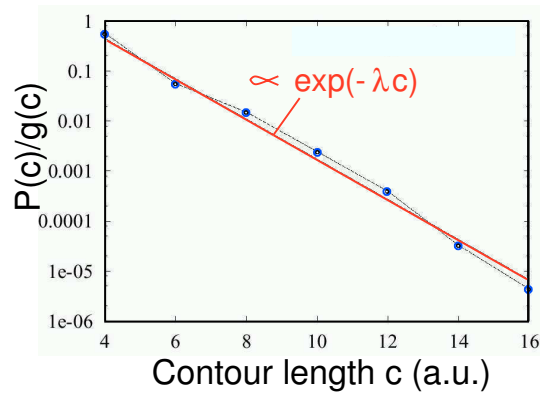


Figure 19. Distribution of pore contours, divided by the degeneracy function $g(c)$.

hope that our brief discussion has conveyed an idea how a coarse-grained molecular simulation help to test and justify mesoscopic theories, and to establish a connection between molecular and mesoscopic descriptions of amphiphilic systems.

4 Outlook

Simulations like those presented here are only a very first step towards understanding the properties of membranes. First, real biomembranes are not pure systems, but contain a mixture of many different lipids, with saturated or unsaturated chains, with charged or neutral heads etc. Second, biomembranes are filled with proteins. Typical biomembranes are not homogeneous, but compartmented into several regions with different lipid and protein composition. Furthermore, biomembranes have a complex environment, which influences

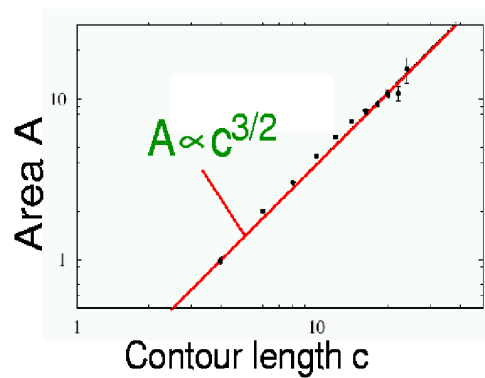


Figure 20. Pore area vs. contour length (arbitrary units).

the membrane properties. Finally, biomembranes are not equilibrium structures. They contain active proteins, and they are surrounded by an ever changing environment.

These many complications seem discouraging. We hope to have shown that a systematic approach, where the different aspects of membranes are studied one by one with appropriated idealized models, can be rewarding. Still, much remains to be done. Computer simulations of membranes and biomembranes will certainly be an active and lively research area for a long time.

Acknowledgments

We have enjoyed collaborations with C. Stadler, H. Lange, M. Mareschal, and K. Kremer. We have benefitted from fruitful discussions with F. M. Haas, R. Hilfer, K. Binder, T. Soddemann, H.-X. Guo, and R. Everaers. The simulations were carried out at the NIC Jülich, the computing center of the Max-Planck society in Garching, and the computing center of the Commissariat à l’Energie Atomique (Grenoble). We acknowledge financial support by the German Science Foundation and by the “Région Rhône-Alpes”.

References

1. J. Israelachvili, *Intermolecular and Surface Forces* (Academic Press, London, 1991).
2. G. Gompper and M. Schick, *Self-Assembling Amphiphilic Systems* Vol. 16 of *Phase Transitions and Critical Phenomena*, C. Domb and J. L. Lebowitz Eds (Academic press, London, 1994).
3. K. V. Schubert, Ber. Bunseng, Phys. Chemie **100**, 190 (1996).
4. R. B. Gennis, *Biomembranes* (Springer Verlag, New York, 1989).
5. V. M. Kaganer, H. Möhwald, and P. Dutta, Rev. Mod. Phys. **71**, 779 (1999).
6. D. A. Antelmi and P. Kekicheff, J. Phys. Chem. B **101**, 8169 (1997).
7. C. Münster, T. Salditt, M. Vogel, R. Siebrecht, and J. Peisl, Europhys. Lett. **46**, 486 (1999).
8. T. Salditt, C. Münster, U. Mennicke, C. Ollinger, and G. Fragneto, Langmuir **19**, 7703 (2003).
9. E. Freyssingeas, D. Antelmi, P. Kekicheff, P. Richetti, and A. M. Bellocq, Europ. Phys. J. B **9**, 123 (1999).
10. I. Koltover, J. O. Rädler, T. Salditt, K. J. Rothschild, and C. R. Safinya, Phys. Rev. Lett. **82**, 3184 (1999).
11. C. Münster, J. Lu, S. Schinzel, B. Bechinger, and T. Salditt, Europ. Bioph. J. **28**, 683 (2000).
12. J. O. Rädler, I. Koltover, T. Salditt, and C. R. Safinya, Science **275**, 810 (1997).
13. F. Artzner, R. Zantl, G. Rapp, and J. O. Rädler, Phys. Rev. Lett. **81**, 5015 (1998).
14. H. G. Döbereiner, Curr. Opin. Coll. Int. Phys. **5**, 256 (2000).
15. H. M. McConnell, Ann. Rev. Phys. Chem. **42**, 171 (1991).
16. C. M. Knobler and R. C. Desai, Ann. Rev. Phys. Chem. **43**, 207 (1991).
17. S. A. Safran, *Statistical Thermodynamics of Surfaces, Interfaces and Membranes* (Addison-Wesley, Reading, Massachusetts, 1994).
18. R. Lipowsky, J. Phys. II **2**, 1825 (1992).

19. H. G. Döbereiner, J. Käs, D. Noppl, and I. Sprenger, *Biophys. J.* **65**, 1396 (1993).
20. S. Yamamoto and S. Hyodo, *J. Chem. Phys.* **118**, 7937 (2003).
21. D. A. Brown and E. London, *Ann. Rev. Cell. Dev. Biol.* **14**, 111 (1998).
22. S. Safran, T. Kuhl, and J. Israelachvili, *Biophys. J.* **81**, 859 (2001).
23. R. G. Larson, L. E. Scriven, and H. T. Davis, *J. Chem. Phys.* **83**, 2411 (1985).
24. T. B. Liverpool, *Ann. Rev. Comp. Phys. IV*, D. Stauffer Edt., pp. 317-358 (World Scientific, Singapore, 1996).
25. R. G. Larson, *J. Physique II* **6**, 1441 (1996).
26. B. Smit, A. G. Schlijper, L. A. M. Rupert, N. M. van Os, *J. Phys. Chem.* **94**, 6933 (1990).
27. D. A. Pink, T. J. Green, and D. Chapman, *Biochemistry* **19**, 349 (1980).
28. A. Caillé, D. Pink, F. de Verteuil, and M. Zuckermann, *Can. J. de Physique* **58**, 581 (1980).
29. B. Dammann, H. C. Fogedby, J. H. Ipsen, C. Jeppesen, K. Jorgensen, O. G. Mouritsen, J. Risbo, M. C. Sabra, M. M. Sperotto, M. J. Zuckermann, *Handbook of Nonmedical Applications of Liposomes*, Vol. 1, p. 85, D. D. Lasic, Y. Barenholz Eds. (CRC press, 1995).
30. G. Gompper and D. M. Kroll, *J. Phys.: Cond. Matt.* **9**, 8795 (1997).
31. M. Müller, K. Katsov, and M. Schick, *J. Chem. Phys.* **116**, 2342 (2002).
32. R. Holyst, *Phys. Rev. A* **44**, 3692 (1991).
33. N. Lei, C. R. Safinya, and R. F. Bruinsma, *J. Phys. II* **5**, 1155 (1995).
34. J. Lister, *Phys. Lett.* **53A**, 193 (1975).
35. J. Shillcock and U. Seifert, *Biophys. J.* **74**, 1754 (1998).
36. A. M. Bibo and I. R. Peterson, *Adv. Mater.* **2**, 309 (1990).
37. I. R. Peterson, V. Brzezinsky, R. M. Kenn, and R. Steitz, *Langmuir* **8**, 2995 (1992).
38. G. A. Overbeck and D. Möbius, *J. Phys. Chem.* **97**, 7999 (1993).
39. F. M. Haas and R. Hilfer, *J. Chem. Phys.* **103**, 3859 (1996).
40. C. Stadler, H. Lange, and F. Schmid *Phys. Rev. E* **59**, 4248 (1999).
41. C. Stadler and F. Schmid *J. Chem. Phys.* **110**, 9697 (1999).
42. D. Düchs and F. Schmid *J. Phys.: Cond. Matt.* **13**, 4853 (2001).
43. D. Düchs, *Diplomarbeit Universität Mainz*, 1999.
44. K. Binder and D. W. Heermann, *Monte Carlo Simulation in Statistical Physics: An introduction* (Springer, Berlin, 1990).
45. D. P. Landau and K. Binder, *A Guide to Monte Carlo Simulations in Statistical Physics* (Cambridge University Press, Cambridge, 2000).
46. M. P. Allen and D. J. Tildesley, *Computer Simulations of Liquids* (Oxford University Press, Oxford, 1989).
47. D. Frenkel and B. Smit, *Understanding Molecular Simulation* (Academic Press, London, 1996).
48. D. Harries and A. Ben-Shaul, *J. Chem. Phys.* **106**, 1609 (1997).
49. A. Baumgärtner, *J. Chem. Phys.* **103**, 10669 (1995).
50. A. Baumgärtner, *Biophys. J.* **71**, 1248 (1996).
51. T. Sintès and A. Baumgärtner, *Biophys. J.* **73**, 2251 (1997).
52. R. Goetz and R. Lipowsky, *J. Chem. Phys.* **108**, 7397 (1998).
53. R. Goetz, G. Gompper, and R. Lipowsky, *Phys. Rev. Lett.* **82**, 221 (1999).
54. O. Farago, *J. Chem. Phys.* **119**, 596 (2003).

55. O. Lenz and F. Schmid, *J. Mol. Liquids*, in preparation (2003).
56. C. Loison, Dissertation, ENS Lyon and Universität Bielefeld, 2003.
57. C. Loison, M. Mareschal, K. Kremer, and F. Schmid, *J. Chem. Phys.*, in press (2003).
58. C. Loison, M. Mareschal, and F. Schmid, in preparation (2003).
59. T. Soddemann, B. Dünweg, and K. Kremer, *Europ. Phys. J. E* **6**, 409 (2001).
60. H.-X. Guo, K. Kremer, and T. Soddemann, *Phys. Rev. E* **66**, 061503 (2002).
61. M. Müller, M. Schick, *J. Chem. Phys.* **116**, 2342 (1996).
62. S. Marrink, F. Jähnig, and H. Berendsen, *Biophys. J.* **71**, 632 (1996).
63. D. Zahn, J. Brickmann, *Chem. Phys. Lipids* **352**, 41 (2002).

Random-Lattice Simulations of Biomembranes

**Ole G. Mouritsen , Myer Bloom , John H. Ipsen , Ling Miao , Morten Nielsen ,
James Polson , Jenifer Thewalt , Ilpo Vattulainen
Hong Zhu , and Martin J. Zuckermann**

¹ MEMPHYS Center for Biomembrane Physics, Physics Department,
University of Southern Denmark, DK-5230 Odense M, Denmark
E-mail: ogm@memphys.sdu.dk

² Department of Physics and Astronomy,
University of British Columbia, Vancouver, BC Canada

³ MEMPHYS, Center for Biomembrane Physics, Department of Chemistry,
University of Southern Denmark, DK-5230 Odense M, Denmark

⁴ Centre for the Physics of Materials, Department of Physics,
McGill University, Montreal, QC, Canada

⁵ Department of Physics,
University of Prince Edward Island, Charlottetown, PE, Canada

⁶ Department of Physics,
Simon Fraser University, Burnaby, BC, Canada

⁷ Laboratory of Physics and Helsinki Institute of Physics,
Helsinki University of Technology, P.O.Box 1100, FIN-02015 HUT, Finland

Lipid bilayers are models of biological membranes. The bilayers undergo thermally driven phase transitions between solid and liquid phases of different liquid-crystalline character. Cholesterol is known as a molecule that strongly modulate the phase behavior of lipid bilayers. It is surmised that the molecular evolution of sterol molecules into cholesterol was subject to a selection pressure that led to optimization of certain membrane physical properties. In this set of lecture notes we demonstrate how a combination of experiments and computer-simulation calculations can shed some light on this question by assessing the differential effects on lipid-bilayer behavior of cholesterol and one of its evolutionary predecessors, lanosterol. The simulations are based on a microscopic model for the lipid-sterol binary mixtures that consists of three basic ingredients: a representation of the translational degrees of freedom, a description of molecular conformational degrees of freedom and an expression for the potential which provides a minimal model of the interactions between the molecules. The simulations are performed using a random-lattice representation and algorithm. The methods of analysis involve umbrella sampling, histogram techniques, and finite-size scaling analysis. Results are presented for phase diagrams, order parameters, structure factors, as well as lateral diffusion. It is concluded that cholesterol in lipid membranes, in contrast to lanosterol, lead to the stabilization of a particular membrane phase, the so-called liquid-ordered phase, which may present some biological advantages over more disordered liquid phases. In particular the liquid-ordered phase sustains the formation of locally ordered domains (or 'rafts') which biologists currently believe are the platform for a number of important biological functions.

1 Introduction¹

The traditional view of biological membrane organization is that the lipid molecules form a featureless two-dimensional fluid in which proteins are kept in place by hydrophobic

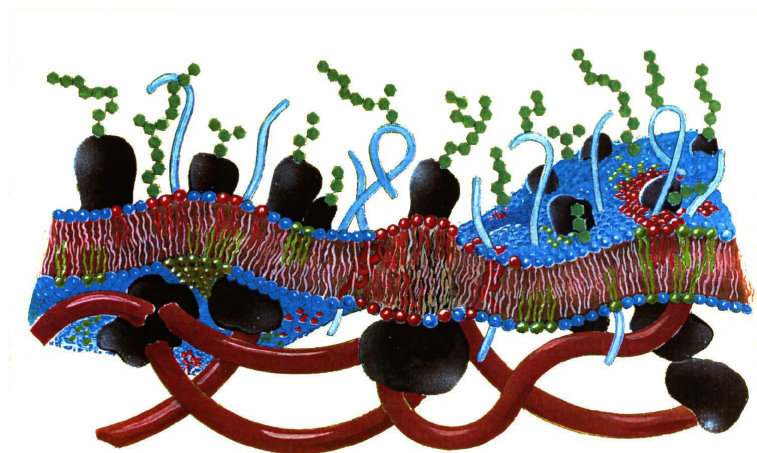


Figure 1. Schematic model of the plasma membrane of a eukaryotic cell which highlights the membrane as a composite of a central lipid bilayer sandwiched between the carbohydrate glycocalyx (which consist of polysaccharides) on the outside and the rubber-like cytoskeleton (which is a polymeric protein network) on the inside. Intercalated in the lipid bilayer are shown various integral proteins and poly-peptides. The membrane displays undulations and the lipid bilayer displays lateral heterogeneity, lipid domain formation, and thickness variations close to the integral proteins. Whereas the lipid molecules in this representation are given with some structural details, the membrane-associated proteins remain fairly featureless. In order to capture many different features in the same illustration, the different membrane components are not drawn to scale.

interactions with the lipid acyl (fatty acid) chains, hydrophilic interactions with the lipid polar heads, and perhaps anchoring interactions with the cytoskeleton². In the absence of anchoring, the proteins can diffuse laterally in the plane of the membrane. A schematic illustration of a eucaryotic cell plasma membrane is given in Fig. 1.

A large number of experimental studies of membranes and lipid bilayers have in the last decade made it increasingly clear, that membranes are highly structured fluids with a substantial degree of lateral structure³. The advent of the so-called raft hypothesis for vertebrate and fungal plasma membranes has shaped this view^{4,5}. Research into the properties of detergent resistant membrane fractions has led to the result that sorting of membrane proteins and signal transduction require co-localization of the related proteins in domains of length scale 70-200 nm, which are called rafts. The lipid compositions of these domains are far from featureless, rich in sphingolipids, cholesterol (vertebrates), or ergosterol (fungi). Moreover, the sterol is vital as its absence impedes raft formation. Finally, it has been conjectured that, in most cases, rafts form a more compact phase than that of the surrounding lipid environment, while still remaining fluid. Thus, the research on rafts suggests strongly the active roles of lipids, especially sterols, in biological functions of cells.

Since cholesterol (or ergosterol) is a key component of all plasma membranes and furthermore is central for the raft hypothesis, it is of interest to investigate, in a simple and transparent setting, the way cholesterol modulates the physical properties of lipid membranes. In particular, it is relevant to study the effect on the membrane phase equilibria and how cholesterol may induce various ordering phenomena in membranes, especially on small length scales. This calls for specific experiments on well-defined model sys-

tems together with theoretical calculations on simple molecular models. In the present set of lecture notes we shall describe a two-pronged approach to the problem by combining experiments with simulations. The experimental data is obtained mostly from nuclear magnetic resonance (NMR) spectroscopy. On the theoretical side it would be ideal to carry out Molecular Dynamics (MD) simulations on realistic atomic-scale models. Despite the advances in computer technology it is at the moment only possible to perform MD simulations on fairly small systems⁶⁻¹² (up to about 10⁶ atoms, corresponding to some hundred lipid molecules and tens of thousands of water molecules) of lipid bilayers in water, possible with a single large protein or a few smaller proteins imbedded in the membrane. The MD simulations are usually limited in time up to about 10-100 ns. In the case of lipid membranes containing sterols^{10,13,14} this implies that MD simulations are currently unable to investigate lipid membrane phase transitions and phase equilibria, whereas these techniques are capable of obtaining detailed information on the effect of local ordering due to different kinds of sterols. Therefore simpler coarse-grained models are called for that allow larger systems to be simulated. This is the strategy which forms the basis of the present set of lecture notes.

1.1 Lipid Bilayers are Model Membranes

It is generally difficult to investigate the properties of biological membranes at a molecular level because of the enormous chemical diversity of lipids and proteins involved. Many experimental and theoretical studies are therefore, performed on model membranes, which are composed of a few lipid and/or protein species. Most of such model membranes can be easily produced in-vitro as a result of the self-assembly process of their molecular components dispersed in aqueous media¹⁵.

Despite of their relative chemical simplicity the model membranes exhibit sufficient complexity to mimic some of the physical characteristics of biological membranes. For example, ternary mixtures of two lipid species with cholesterol have been used by many researches as model systems for studying raft formation⁴. In these model membranes, domains where one of the lipid species forms a compact physical phase – known as the liquid ordered (**lo**) phase¹⁶ – together with a large concentration of cholesterol molecules appear within a less compact background fluid formed by the other lipid component with very little cholesterol. The nature of the background phase is typical of single-component lipid bilayers in the liquid crystalline phase, also known as the liquid disordered (**ld**) phase¹⁶. We will give a description of the phases below.

We shall be particularly interested in investigating the thermodynamic phase behaviour of such model membranes with the view to understand the essential microscopic interactions underlying the phase behaviour. Our approach has been to carry out theoretical studies based on computer simulations of microscopic interaction models in parallel with experimental investigations. The microscopic interaction models that our group has used or developed during the past two decades have varied focuses, depending on the specific physical questions under investigation^{17,18}. Our earlier work involved the use of lattice models^{16,18,19}. Our more recent investigations, however, focused on the development and study of off-lattice models that are able to provide a genuine description of the fluid characteristics of both the **ld** and the **lo** phases²⁰. In the present lecture notes, we present our work on the application of off-lattice models to the physical properties of lipid-sterol bilayers^{21,22}.

The methodological backbone of our computer simulations is the Metropolis Monte Carlo (MC) method. This method guarantees that a simulated system will reach its state of thermodynamic equilibrium. A tutorial on the use MC methods for the simulation of lattice and off-lattice models for lipid bilayers and lateral organization has appeared recently²³ as well as a general review on lattice models used for simulation of membranes interacting with proteins, peptides, sterols, drugs, and other solutes¹⁹. The present lecture notes will furnish a description of the use of the method together with many other simulational techniques within the specific context of our modeling.

It is important to note at the outset that, given the available computational capacity, our simulations are limited in terms of the size of the sample, which is restricted to systems of the order of 10³ lipid chains for our lattice models and 10⁴ lipid chains for the off-lattice models. This enables us to study the nature of distinct phases and phase transitions using dedicated techniques, but not systems with a distribution of domains of the sizes predicted for rafts.

1.2 Pure Lipid Bilayers

Our approach has been to study model systems with the least number of molecular components first and then proceed to systems of several components^{17,18,20}. The initial systems are pure, single-component lipid bilayers which themselves exhibit several phase transitions. We will concentrate on one particular phase transition known as the main phase transition, which gives much insight into the interactions between lipid molecules. The transition takes the bilayer from the gel phase to the liquid crystalline phase as the temperature is increased and is abrupt, highly entropic. The gel phase is a two-dimensional (2d) crystalline phase in which conformationally ordered lipid chains form a structural lattice. Here conformationally ordered lipid chains are almost fully extended with very few gauche bonds towards their free ends. To describe the nature of the gel phase in more detail, we have renamed it the solid-ordered (**so**) phase. In contrast the liquid crystalline phase is a fluid phase in which conformationally disordered lipid chains can diffuse laterally in the bilayer plane. Here the conformationally disordered lipid chains contain many gauche bonds, leading to a reduced apparent length, but an increased cross-sectional area. The quantitative properties of the main phase transition depend both on the intrinsic chain length and unsaturation and on the type of polar head of the lipid molecule involved as well as on the ionic character of the aqueous medium hydrating the polar heads.

1.3 The Liquid-Ordered Phase and Sterol Evolution

The concept of liquid-ordered (**lo**) phase emerged already 15 years ago from both experimental and theoretical studies of the phase behaviour of dipalmitoyl phosphatidylcholine (DPPC)-cholesterol bilayers¹⁶. DPPC is a glycerophospholipid with a zwitterionic polar head and two saturated (16:0) acyl chains and appears as a molecular component of many biological membranes. The experimental phase diagram of DPPC-cholesterol multibilayer liposomes was first deduced by Vist and Davis²⁴ from both deuterium nuclear magnetic resonance (²H-NMR) and differential scanning calorimetry (DSC) data. The same topology was found for the phase diagram of PPetPC-cholesterol bilayers²² as shown in Fig. 2a, even though PPetPC differs in structure from DPPC in that one *trans*-double bond

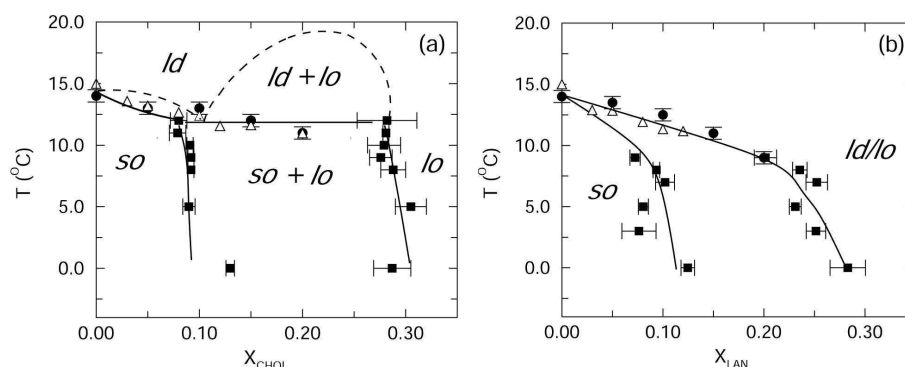


Figure 2. Experimental phase diagrams as determined by differential-scanning calorimetry (open triangles) and by NMR spectroscopy (solid squares and solid circles). (a) PPetPC-cholesterol systems; (b) PPetPC-lanosterol systems. The lines connecting the points are guides to the eyes. The phase boundaries indicated by the dashed lines in (a) are not derived from any experimental data and are shown only to illustrate the qualitative structure of the phase diagram, which is consistent with the thermodynamic phase rules.

replaces a single C–C bond in DPPC at carbon seven. Qualitatively similar phase diagrams were also found by Thewalt et. al.²⁵ for several other systems of binary mixtures containing cholesterol and lipids with PC polar heads. The most important feature of these phase diagrams is that an abrupt main phase transition is absent for cholesterol concentrations greater than 20–30 mol%, and a new phase appears in this region of the phase diagrams. This new phase is a fluid phase where the lipid chains are reasonably ordered conformationally as demonstrated by the ¹H-NMR data. Vist and Davis²⁴, who first identified this new phase named it the *l_o* phase. In the first theory for the experimental phase diagram, Ipsen et al.¹⁶ renamed this phase the liquid-ordered or *l_o* phase since this phase is a compact liquid phase with conformationally ordered chains¹⁶. This concept has gained a prominent position in raft research, because several experimental studies on rafts suggest that rafts have the physical characteristics of the *l_o* phase^{4,5}. It is worth pointing out that the only *l_o/l_d* phase separation observed for single lipid-cholesterol bilayers occurs close to the melting temperature of the pure lipid system. Thus it is reasonable to assume that mixtures of lipids and cholesterol would require at least some high- components in order to display the *l_o* phase at body temperature.

Given the existing understanding of the equilibrium phase behaviour of the single lipid-cholesterol bilayers and the underlying microscopics and inspired by the work of Bloch on sterol evolution²⁶ we have recently carried out a comparative study of lipid-cholesterol and lipid-lanosterol bilayers, both experimentally and theoretically by computer simulations. Lanosterol is the first cyclic precursor of cholesterol in the biosynthetic pathway and is regarded as a molecular fossil by Konrad Bloch who proposed that the evolutionary sequence for cholesterol is preserved in its biosynthetic pathway²⁶.

The differences in the phase behaviour of the two types of lipid-sterol systems are illustrated in the two phase diagrams in Fig. 2²². The phase diagram in Fig. 2a shows that the *l_o* phase is a well defined thermodynamic phase distinctly separate from the *l_d* phase, but that it can coexist with the *l_d* phase under appropriate temperatures and concentrations. In contrast the phase diagram in Fig. 2b shows that for PPetPC-lanosterol multi-bilayer

systems the **lo** and the **ld** phases are no longer thermodynamically distinct, i.e. that the coexistence of the two phases is absent.

Moreover, our experimental data for PPetPC-lanosterol and PPetPC-cholesterol show that cholesterol induces greater conformational order in PPetPC-lipid chains than lanosterol²². It may be reasonable to expect such a relative effect also to apply to DPPC-cholesterol and DPPC-lanosterol bilayers. Hence a possible hypothesis is that one reason for the evolutionary optimization of cholesterol is that it is able to induce the formation of rafts better than lanosterol. There is in fact some evidence that lanosterol cannot induce raft formation²⁷.

2 Off-Lattice Models^{20–22,29}

The first step in a statistical mechanical study of a model membrane system is to construct a model describing the microscopic physics of the membrane. The model should contain two basic ingredients: 1) a specific description of the microscopic states of the membrane, i.e., the relevant microscopic degrees of freedom; 2) a specific expression of the total interaction potential energy of the system as a function of the relevant degrees of freedom. In both parts, approximations or simplifications are necessary, mainly due to the limited computational capacity available and the need to underline among the almost formidably complicated microscopic details the general mechanisms that are essential to the macroscopic phenomena under investigation. The models to be described below may in fact be called ‘minimal’ in terms of the approximations and simplifications involved.

The basic molecular degrees of freedom of a lipid-bilayer membrane consists of both the positional (molecular center of mass) and internal (conformational) degrees of freedom, and both types are relevant to the phase behaviour of the membrane. The description of the positional degrees of freedom is straightforward, given by specifying the spatial coordinates of the molecules. Our description of the conformational isomerism of the lipid chains is a minimal one. Only two internal states are used to approximate the very large spectrum of conformations possible for a phospholipid chain: One state, the ‘ordered’ state, has zero internal energy and is non-degenerate, characteristic of the conformational state of a lipid chain in the gel or **so** phase. The other state, the ‘disordered’ state, is characteristic of the average conformational state of a lipid chain in the liquid-crystalline or **ld** phase. This state has a high internal energy, reflecting that energy is required for conformational excitations, and a high degeneracy, effectively representing the large number of conformational excitations of that a phospholipid chain can assume in the **ld** phase due to the presence of many gauche bonds. This idea was first proposed and used by Doniach¹⁷. Both cholesterol and lanosterol are rigid molecules in comparison with the phospholipid chains and we assume that they have no conformational degrees of freedom.

In constructing the function describing the interaction potential energy, several approximations have also been made. First, the interactions between the amphiphilic molecules (phospholipid, sterol) and the water molecules, which are predominantly responsible for stabilizing a bilayer membrane, are approximated by a surface-pressure parameter π , as first suggested by Marčelja³⁰. Secondly, the non-covalent interactions between the amphiphilic molecules may, to a first approximation, be considered as being pair-wise, of the Van der Waals type, and effective only over a short range of molecular length scale. Finally, only one monolayer is modeled based on the assumption that the two monolayers constitut-

ing a lipid bilayer are independent of each other. In essence, our model system consists of microscopic ‘particles’ – distinct lipid chains and sterol molecules – which move within a two-dimensional plane and which interact in a pairwise manner, as described below. In the rest of this text, the term ‘particle’ may refer either to a lipid chain or to a sterol molecule.

The total potential energy function that models molecular interactions in the lipid-sterol bilayer systems in our simulations is then given by

$$U = U_{\text{ph}} + U_{\text{cc}} + U_{\text{cs}} + U_{\text{ss}} \quad (1)$$

where U_{ph} is the potential energy function describing the interactions between the phospholipid molecules, whose specific form is as follows

$$U_{\text{ph}} = \sum_i \mathcal{L}_i \left[\sum_{\langle j \rangle} -\epsilon_{ij} \mathcal{L}_j \right] + \sum_i \left[-\epsilon_{ii} \{ \mathcal{L}_i \mathcal{L}_i + \mathcal{L}_i \mathcal{L}_i \} \right] \quad (2)$$

Here i is an index labeling the ‘particles’ (lipid chains and sterol molecules) in the system. Correspondingly, $\mathcal{L}_i$ and $\mathcal{L}_j$ are occupation variables which are unity when the i th particle is a lipid chain in the ordered state and the disordered state, respectively, and which are zero otherwise. ϵ_{ii} is the excitation energy of the conformationally disordered state, and $\epsilon_{ij} = \epsilon_{ji}$ and $\epsilon_{ii} = \epsilon_{ii}$ are distance-dependent and chain-conformation-dependent interactions between two lipid chains. $\langle \dots \rangle$ denotes a summation over nearest neighbors, corresponding to the approximation of short-range interactions. The energy of interaction between two chains which are both in the disordered state is approximated to be zero, which sets the reference point for energy. This potential energy function, U_{ph} , together with the configurational degeneracy of the disordered state, provides a minimal model for the main phase transition of single component bilayers.

ϵ_{ij} , ϵ_{ii} and ϵ_{ii} , as indicated by the various subscripts, represent the pairwise interactions between an ordered chain and a sterol molecule, a disordered chain and a sterol molecule, and two sterol molecules, respectively. They are written explicitly as follows:

$$\begin{aligned} \epsilon_{ij} &= \sum_{\langle j \rangle} -\epsilon_{ij} \{ \mathcal{L}_i \mathcal{L}_j + \mathcal{L}_i \mathcal{L}_j \} \\ \epsilon_{ii} &= \sum_{\langle j \rangle} -\epsilon_{ii} \{ \mathcal{L}_i \mathcal{L}_j + \mathcal{L}_i \mathcal{L}_j \} \\ \epsilon_{ii} &= \sum_{\langle j \rangle} -\epsilon_{ii} \{ \mathcal{L}_i \mathcal{L}_j \} \end{aligned} \quad (3)$$

Again $\mathcal{L}_i$ is an occupation variable which is unity when the i th particle is a sterol molecule and is zero otherwise. Clearly, $\mathcal{L}_i \mathcal{L}_j + \mathcal{L}_i \mathcal{L}_j = 1$. $\epsilon_{ij} = \epsilon_{ji}$ and $\epsilon_{ii} = \epsilon_{ii}$ are the distance-dependent, chain-conformation-dependent interaction potentials between a sterol molecule and a lipid chain. Similarly $\epsilon_{ii} = \epsilon_{ii}$ is the interaction between two sterol molecules.

All of the pair-wise interaction potentials are approximated by a sum of a hard-core repulsive potential of range r_c , a short-range square-well potential of range r_s , ϵ_s ($\epsilon_s < 0$), and a longer-range attractive square-well potential of range r_l , ϵ_l ($\epsilon_l < 0$). ϵ_s and ϵ_l

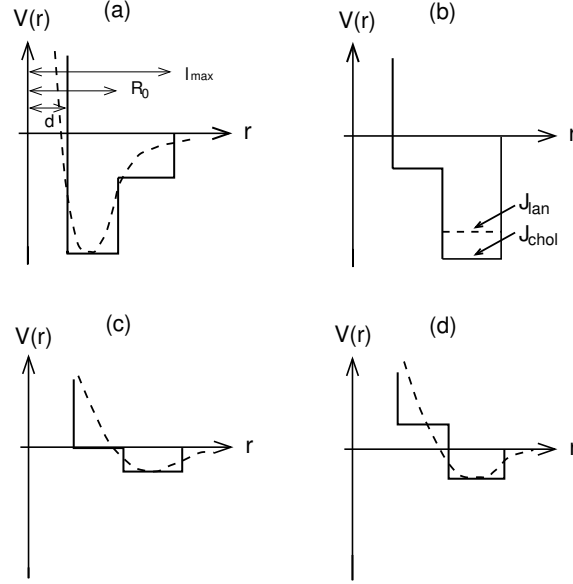


Figure 3. Schematic illustration of the model interaction potentials. (a) $V_{o-o}(r)$, where subscript 'o' refers to a lipid chain in the conformationally ordered state; (b) $V_{o-s}(r)$, where the depth of the potential corresponding to s is represented by a solid line and a parameter J_{chol} and that corresponding to o is represented by a dotted line and a parameter J_{lan} ; (c) $V_{d-s}(r)$, where subscript 'd' refers to a lipid chain in the disordered state; (d). $V_{s-s}(r)$. d is the hard-core radius, R_0 is the range of the short-range potentials and l_{max} is the effective range of the long-range potentials. The ratio of R_0/d is chosen so that the average surface area of a sterol molecule is approximately 30% larger than that of a lipid chain in the ordered state. The dashed lines illustrate the more realistic interaction potentials that the model potentials approximate.

are given by

$$V(r) = \begin{cases} -\epsilon & r \leq R_0 \\ 0 & r > R_0 \end{cases} \quad (4)$$

$$V(r) = \begin{cases} -\epsilon & r \leq R_0 \\ 0 & r > R_0 \end{cases} \quad (5)$$

where the specific values of ϵ and R_0 depend on both the molecular type and the conformational state of the interacting species. Here R_0 is a cutoff distance which will be defined when the simulation algorithm is discussed in the next subsection.

The pair-wise interaction potentials are illustrated in Fig. 3, and their construction is based on the hypothesis that cholesterol interacts with lipid molecules as follows. Cholesterol, with its streamlined, rigid hydrophobic backbone, prefers the lipid chains in its immediate neighborhood to be in the conformationally ordered (rigid) state by suppressing the formation of *gauche* bonds in those chains. At the same time cholesterol tends to disrupt laterally ordered packing of conformationally ordered chains if it is in their midst. A comparison between Fig. 3a and Fig. 3b illustrates the latter, 'crystal-breaker' mechanism, as the interactions involved imply that a cholesterol molecule dissolved in an ordered-chain

environment tends to have a larger surface area than that of a lipid chain, thereby breaking the lateral order of ordered chains. Similarly, a comparison between Fig. 3b and Fig. 3c shows that the former, ‘chain-rigidifier’ mechanism is modeled by a strong interaction between cholesterol and a neighboring, conformationally ordered lipid chain and a weaker interaction with a conformationally disordered lipid chain.

The values of all the relevant parameters in the model are chosen as follows. The unit of length scales in the model is for convenience set at the hard-core diameter, σ , of the interaction potentials and the unit of energy is defined to be a quantity ϵ . In order to convert the energy and length scales into units relevant for lipid bilayer systems, should be of the order 10^{-21} J and σ of the order 5 Å. The lateral pressure, Π , is fixed at $\Pi = 3 \times 10^{-10}$ Pa. The radius of the short-range potential, Eq. (4), is set at $r_s = 1.3\sigma$. The values of ϵ_{oo} and ϵ_{od} are chosen such that the change in surface area across the main transition is comparable to that of a pure PC (DPPC) bilayer system. The excitation energy of the disordered state of lipid chains is chosen to be $\epsilon_{ds} = 2.78\epsilon$, and the degeneracy, Ω , of the disordered state, is taken to be $\Omega = 12/78$. These values are the same as for the 10th state of the Pink Model for the main phase transition of single component PC bilayers with saturated chains³¹. Other parameters appearing in the definition of the interaction potentials are summarized in Table 1(a). They are chosen such that the theoretical phase diagram of the lipid-cholesterol model system is similar to that of the DPPC-cholesterol system.

(a)		ϵ_{oo}	ϵ_{od}	ϵ_{dd}	ϵ_{ss}	ϵ_{sd}
		0.40	-0.15	0.20	0.20	0.00
	long range	0.40	-0.15	0.20	0.20	0.00
	short range	0.45	0.40	-0.20	-0.15	-0.065

(b)		Cholesterol	Lanosterol
	ϵ_{os}	0.85	0.75
	ϵ_{os}	-0.625	-0.525

Table 1. Interaction parameters for the model potentials for the two types of lipid-sterol membranes. (a) The parameter values for the interaction potential, V_{oo} , between two lipid chains in the ordered state, the interaction potential, V_{od} , between a lipid chain in the ordered state and a lipid chain in the disordered state, the interaction potential, V_{dd} , between two lipid chains in the disordered state, the interaction potential, V_{ss} , between two sterol molecules, and the interaction potential, V_{sd} , between a lipid in the disordered state and a sterol molecule. Note that these values are identical for both the lipid-cholesterol and the lipid-lanosterol systems. (b) The parameter values for the interaction potential, V_{os} , between a lipid chain in the ordered state and a sterol molecule, where the subscript s corresponds to either chol or lan, respectively. All the parameters are given in units of J_0 (see text).

The parameters given in Table 1(b) reflect our – again, minimal – modeling of differences in lipid-cholesterol and lipid-lanosterol interactions. The model arises from a comparative analysis of the molecular structures of the two sterols which are shown in terms of space-filling models in Fig. 4. Figure 4 indicates that the major structural difference is on their respective β -faces. In detail, cholesterol has a smooth β -face, whereas the β -face of

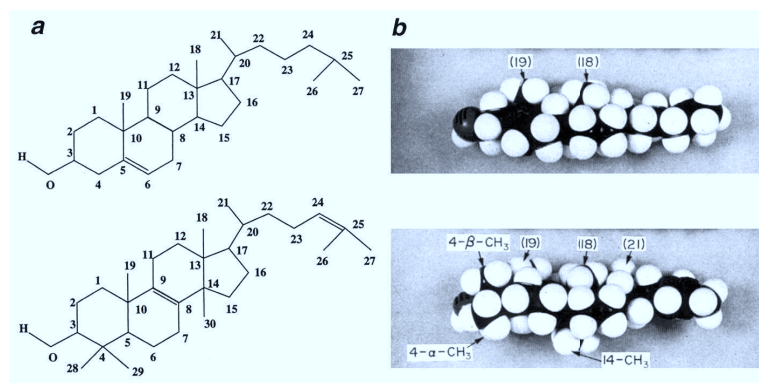


Figure 4. Molecular structures of cholesterol (top) and lanosterol (bottom). (a) Chemical structures; (b) Space-filling models. The three additional methyl groups on lanosterol are indicated in (b) as 14-CH₃, 4-α-CH₃, and 4-β-CH₃.

lanosterol has an extra methyl group, which forms a rather significant protrusion. Together with two other additional methyl groups this β -face protrusion would decrease the ability of lanosterol to order neighboring lipid chains. It is interesting to note that the β -face methyl group is the first one to be removed in the biosynthetic pathway at considerable cost in metabolic energy.

Hence, in our theoretical model, the microscopic lipid-cholesterol and lipid-lanosterol interactions differ only in the part that describes the interaction between a sterol molecule and an ordered lipid chain, χ_{SL} . The values listed in Table 1(b) correspond to a decreased strength of the cohesive lipid-lanosterol interaction relative to that of the cohesive lipid-cholesterol interaction. They were determined by fitting the theoretically calculated value of the lipid-chain order parameter, S_{C} (see later), to the acyl-chain order parameter experimentally derived from the ^1H -NMR data, over a significant range of sterol concentrations for both types of the lipid-sterol systems. The fitting was carried out at a single temperature, which was relatively high so that no complications associated with phase transitions arose. The fit is shown in Fig. 5.

3 Simulation Methods

3.1 Monte Carlo Techniques

In our computer simulations of the thermodynamic phase behaviour of the models described in the previous section, the MC method provides the backbone method. Generally speaking, the MC procedure in each simulation samples stochastically and ergodically the complete configuration space spanned by all possible microscopic states of the system and generates a Markovian sequence of microscopic states whose distribution over the configuration space approximates the equilibrium distribution. Explicitly, each simulation consists of many repeated cycles of procedures, and each cycle, labeled by an integer i , is implemented as follows:

1. An arbitrary microscopic state, Ω_i , is chosen.

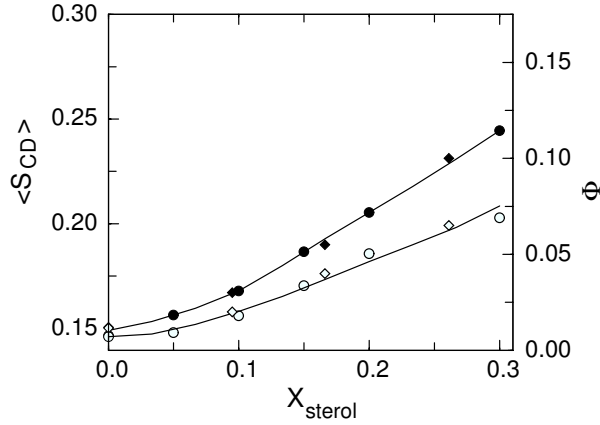


Figure 5. Fitting of the theoretical (ϕ) order parameter to the equilibrium averages of the experimental order parameter ($\langle S_{CD} \rangle$, as given in Eq. (14)) which leads to the specific value of J used for lipid-lanosterol interactions in the model. The experimental data are obtained at $T = 4^\circ\text{C}$. The theoretical data are calculated for a system containing $N = 16$ particles at a temperature of $T = 1.35 T_m$. Corresponding to cholesterol and lanosterol, the experimental data are shown by filled and open circles, respectively, and the theoretical data are given by filled and open diamonds, respectively. The lines connecting the points are guides to the eye.

2. Next a trial microscopic state, $\mathbf{r}'$, is established by a random, ergodic sampling of the complete configuration space.
3. A quantity, $\Delta E \equiv E(\mathbf{r}') - E(\mathbf{r})$, is calculated, where E depends both on the potential-energy function of the system and on the choice of thermodynamic ensemble.
4. A random number, ξ , is generated.
5. If $\xi \leq \exp(-\Delta E/k_B T)$, the trial microscopic state, $\mathbf{r}'$, is accepted as the new configuration of the system, i.e., $\mathbf{r} \leftarrow \mathbf{r}'$. Otherwise the system remains in the same state, i.e., $\mathbf{r} \leftarrow \mathbf{r}$.
6. Quantities of interest for ensemble averaging are calculated and recorded.
7. A new trial microstate, $\mathbf{r}'$, is chosen again randomly and ergodically, and the cycle is started by going back to step 3.

It is implied in the MC procedure that a specific choice of thermodynamic ensemble has been made, which affects both the definition of the complete configuration space and the definition of E . In addition to choosing temperature T and surface pressure π as thermodynamic control variables, we are still left with two different choices regarding the number of ‘lipid-chain’ particles, N_L , and the number of ‘sterol’ particles, N_S : (a) fixing both N_L and N_S ; and (b) fixing the total number of particles, $N = N_L + N_S$, and letting the chemical composition of the system fluctuate under the control of an effective chemical potential, μ . We take $\mu = \mu_S - \mu_L$ to represent the difference between the chemical potentials of the lipid and the sterol particles. We have used either (a) or (b) in our simulations,

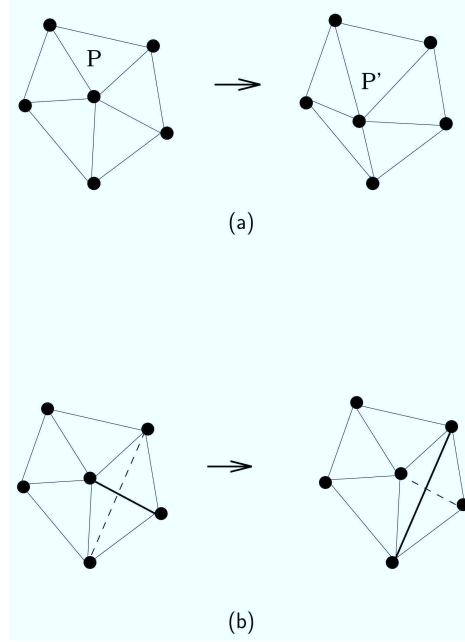


Figure 6. (a) Particle Displacement. The particle at position P is moved to position P'. (b) Link Flip. A tether (shown as a thick line) is replaced by another tether along the unoccupied diagonal (once the first tether is removed) provided that the new tether length does not exceed $l_{\max}$.

3.1.2 Particle Displacement

This move is illustrated in Fig. 6a. A particle is chosen at random and its center is subject to a random displacement (x, y) where

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2\alpha - 1 \\ 2\beta - 1 \end{pmatrix} \quad (8)$$

and α and β are random numbers, $0 \leq \alpha, \beta \leq 1$. The value of α is adjusted during the simulations so that approximately 25 % of trials are accepted. Consistent with the hard-core repulsion between two interacting particles, moves which would result in an overlap of the hard cores of the particle are always rejected. Another constraint is that the length of every tether is not allowed to exceed a maximum value $l_{\max}$.

3.1.3 Link Flip

The second move is referred to as 'link flip,' which alters the local connectivity in the tether network. It is shown schematically in Fig. 6b. In each configuration of the tether network, each tether forms one diagonal of a quadrilateral formed by the two adjacent triangles. In the 'link flip,' a tether is chosen at random and then removed; a new tether is placed along the other diagonal of the quadrilateral, provided that its length does not exceed $l_{\max}$.

3.1.4 Change of System Size

Consistent with our choice of thermodynamic ensemble, where the surface pressure is a control parameter, fluctuations in the total area of the system must be sampled. In our simulations this is achieved via a third move: a random uniform expansion or contraction of the whole system. In this step, a random change in the size of the system is generated as

$$(2 - 1) \quad (9)$$

where α is a random number, $0 \leq \alpha \leq 1$; and the coordinates of all particles in the system are rescaled accordingly. If the distance between any two particles after the rescaling is smaller than the hard-core diameter the change is always rejected. The maximum possible size change, ΔA , is adjusted during the simulation to give an acceptance ratio of about 50 %. The logarithmic term in Eq.(6) or Eq.(7) will necessarily be nonzero following this types of moves. A detailed discussion of this types of sampling moves are given in the book by Frenkel and Smith²⁸, pp. 116-122.

3.1.5 Dealing with the Conformational and Compositional Degrees of Freedom

In simulations in the canonical ensemble, sampling the part of the configuration space that is associated with the conformational degrees of freedom of the lipid chains is straightforward. A lipid chain is chosen at random and its conformational state is always changed in the trial state. For example, if the chain is in the ordered state, it will be in the disordered state in the trial state and vice versa.

In the case of the semi-grand canonical ensemble, a different updating procedure is used, where the conformational degrees of freedom and the compositional degrees of freedom are handled on the same footing. Explicitly, each particle in the system is assigned one of three possible ‘internal’ states: the conformationally ordered and disordered states of a lipid chain, or a third, ‘sterol’ state. Then a random number is used to choose among the three internal states now available.

In order to ensure detailed balance in the simulations, or the symmetry of the Markov chain generated by the algorithm, all of the different moves discussed above are combined in a random manner. In other words, the simulation algorithm does not impose any preferred sequential order in the sampling procedures that update different types of degrees of freedom²⁸. Monte Carlo step (MCS), a time unit for the simulations, is defined to be the time needed to perform on average one complete pass through all the different types of moves.

3.2 Umbrella Sampling²⁸

Up to this point, we have given a detailed recipe for performing an MC simulation for a system described by the potential-energy function of Eq. (1). We now discuss a method for determining phase boundaries specific to our study of the phase behavior of lipid-sterol bilayer membranes. This method is a version of umbrella sampling and it is particularly useful for determining phase transitions whose kinetics are strongly hysteretic. The transitions between the **so** and the **ld** phases in our systems are precisely such transitions. The high free-energy barriers separating the two coexisting phases make it very difficult to achieve an adequate sampling of both phases.

This difficulty was overcome by the use of a simulation technique, which exploits the idea of developing an ‘artificial’ potential-energy function that yields a considerably diminished free-energy barrier. The simulation carried out for the artificial potential-energy function is able to yield an adequate sampling of the two coexisting phases. The equilibrium distribution functions (defined below) for the original potential-energy function can then be established from simulations based on the modified potential-energy function through a simple re-weighting relation³².

The upshot of our simulations performed within the semi-grand canonical ensemble was to provide accurate numerical data, from which an equilibrium distribution function, $\mathcal{P}(\epsilon, x)$, could be derived for each given set of values of ϵ and x . With ϵ representing the internal interaction energy per particle and x the sterol concentration, $\mathcal{P}(\epsilon, x)$ is proportional to the probability of finding the system in states characterized commonly by ϵ and x , and has the property, that if ϵ^* and x^* correspond to a point on a phase-coexistence boundary, then it will exhibit two degenerate local minima, representing the two coexisting phases. Thus, by analyzing the change in the ‘landscape’ of $\mathcal{P}(\epsilon, x)$ as values of ϵ and x are tuned, the phase boundary can be determined.

The umbrella sampling method was employed to this end and its implementation is described as follows. At values of ϵ and x estimated to be close to true coexistence conditions, a single but very long simulation is performed for a system of relatively small size N to yield an approximation of the equilibrium distribution function, $\mathcal{P}(\epsilon, x)$. When summed over x , $\mathcal{P}(\epsilon)$ yields the equilibrium distribution function $\mathcal{P}(\epsilon)$. A spectral free-energy function, defined by $\mathcal{F}(\epsilon) = -\ln \mathcal{P}(\epsilon)$, displays a free-energy barrier between the two local minima. An iterative cycle is then set up:

1. Based on $\mathcal{F}(\epsilon)$ an extrapolation based on the size dependence of the free energy barrier is used to approximate the barrier of a system of a larger size, N' , and in turn, the following function:

$$\mathcal{F}'(\epsilon) = \mathcal{F}(\epsilon) - \frac{N'}{N} \mathcal{F}(\epsilon) \quad (10)$$

when ϵ lies in the barrier region. $\mathcal{F}'(\epsilon)$, known as the shape function, defines the modified potential energy function as $\mathcal{P}'(\epsilon, x)$.

2. $\mathcal{P}'(\epsilon, x)$ is used in a second simulation of a system of size N' . From this simulation, a modified probability distribution function, $\mathcal{P}'(\epsilon, x)$, is obtained. The spectral free-energy function corresponding to the modified potential energy function should not show a significant barrier.
3. The required distribution function, $\mathcal{P}(\epsilon, x)$, is easily reconstructed from $\mathcal{P}'(\epsilon, x)$ ³².
4. Based on $\mathcal{P}'(\epsilon, x)$, a method known as the Ferrenberg-Swendsen re-weighting technique³³ is applied in order to obtain a better estimate of the coexistence condition, ϵ^* and x^* and an improved approximation of $\mathcal{F}(\epsilon^*, x^*)$ at coexistence from $\mathcal{F}'(\epsilon^*, x^*)$.

5. If desired, another iteration is started from step (1) with $\mathcal{F}'(\beta^*, \beta^*, x)$, either to obtain an improved statistical sampling of $\mathcal{P}'(\beta^*, \beta^*, x)$ for the same system size or to simulate a larger system.

This method enables us to simulate systems of relatively large sizes. Repeated applications of the above procedures for systems with systematically varying sizes can yield a series of size-dependent $\mathcal{P}(\beta^*, \beta^*, x)$. A finite-size analysis of the series of data based on the method of Lee and Kosterlitz³⁴ can provide information on the specific nature of the corresponding transition.

Nevertheless, the method is still quite time consuming. In the iterations, the system size can only be increased in very small steps; the statistics required to obtain the initial estimate of the spectral free-energy function $\mathcal{F}$ as described in step (1) is already considerable, being typically of the order of $50 \cdot 10^3$ MCS per particle for a system of size $N = 10^4$.

3.3 Calculation of Physical Quantities

Given that a simulation based on the MC procedure described above does in fact generate an ensemble of microscopic states that are distributed in the configuration space according to the equilibrium distribution, the macroscopic thermal average of a microscopic physical quantity, $\mathcal{O}$, can be approximated by its average over the ensemble of states, $\{\Omega\}$:

$$\langle \mathcal{O} \rangle = \frac{1}{|\Omega|} \sum_{\Omega} \mathcal{O}(\Omega) \quad (11)$$

where $|\Omega|$ is the total number of microscopic states in the ensemble. It is important to note that the precise value of $|\Omega|$ required for Eq. (11) to be a good approximation for the true thermodynamic averages depends on the nature and size of the system as well as the specific physical observable under investigation.

In general, the macroscopic thermodynamic quantities that are of interest for the systems that we have studied are response functions such as specific heat and isothermal area compressibility as well as a macroscopic conformational order parameter characterizing the conformational ordering of the lipid chains. The response functions can be calculated directly from the simulation data by using the fluctuation-dissipation theorems:

$$C_V = \frac{1}{k_B} \left(\langle \mathcal{E}^2 \rangle - \langle \mathcal{E} \rangle^2 \right) \quad (12)$$

$$\chi_T = \frac{1}{k_B} \left(\langle \mathcal{A}^2 \rangle - \langle \mathcal{A} \rangle^2 \right) \quad (13)$$

where $\mathcal{E}$ is the potential-energy function defined in Eq. (1), C_V is the molecular specific heat at constant volume, and χ_T is the isothermal area compressibility. The average conformational order parameter of the lipid chains is calculated as

$$\frac{1}{2} \left(\left\langle \frac{\sum_i (\mathcal{L}_i - \mathcal{L}_i^0)}{\sum_i (\mathcal{L}_i + \mathcal{L}_i^0)} \right\rangle + 1 \right) \quad (14)$$

In addition to the physical quantities mentioned above, the so-called structure factors, which provide information on the lateral organization of the constituent molecules in the

system, have also been calculated, based on the simulations performed within the canonical ensemble. Two of them are listed here, for which numerical data will be provided later:

$$\begin{aligned} \rho(\mathbf{q}) &= \frac{1}{N} \{ \langle \rho(\mathbf{q}) \rho(-\mathbf{q}) \rangle - \langle \rho(\mathbf{q}) \rangle \langle \rho(-\mathbf{q}) \rangle \} \\ \rho_s(\mathbf{q}) &= \frac{1}{N_s} \{ \langle \rho_s(\mathbf{q}) \rho_s(-\mathbf{q}) \rangle - \langle \rho_s(\mathbf{q}) \rangle \langle \rho_s(-\mathbf{q}) \rangle \} \end{aligned} \quad (15)$$

Here $\rho(\mathbf{q})$ is the Fourier transform of the total density, $\rho(\mathbf{r}) \equiv \sum_i \delta(\mathbf{r} - \mathbf{r}_i)$ and $\rho_s(\mathbf{q})$ the Fourier transforms of the partial density of the sterol molecules, $\rho_s(\mathbf{r}) \equiv \sum_i \delta(\mathbf{r} - \mathbf{r}_i) \mathcal{L}_s$.

Finally, some simulations performed in the canonical ensemble were also used to derive the tracer diffusion coefficient D . According to the Einstein relation, the diffusion constant can be expressed as

$$D = \lim_{t \rightarrow \infty} \frac{1}{2t} \langle |\mathbf{r}(t) - \mathbf{r}(0)|^2 \rangle \quad (16)$$

where 2 is the number of spatial dimensions of membranes, t is time, and the factor in the angular brackets is the mean-square displacement of the diffusing particle over time interval t . In practice, the simulation data $\langle |\mathbf{r}(t) - \mathbf{r}(0)|^2 \rangle$ was plotted as a function of t and the plot displayed linearity for sufficiently large t . D was then determined from the slope of the linear part of the plot. Since the simulations were carried out using Monte Carlo methods, the time scales are meaningful only in a relative sense, as opposed to the absolute physical time scales in the system.

A practical remark on the simulations may be useful. In each simulation run, the system was equilibrated over a period of 200,000 MCS, and the various physical quantities were averaged over a period of 5-20·10⁵ MCS.

4 Results

In this section we present a selection of the results of the Monte Carlo simulations, in terms of the equilibrium phase diagrams, collective conformational ordering of the lipid chains and structural characterizations of the lipid-sterol bilayer membranes. All the results in this section were obtained using the potential energy functions and the simulation methods presented in the previous section.

4.1 Phase Diagrams

4.1.1 Pure Lipid Bilayers

The phase diagram obtained from our simulation study for model lipid bilayers in the absence of cholesterol²⁰ is shown in Fig. 7. The phase diagram was calculated using the potential of Fig. 3a and is represented in terms of temperature and the parameter, ϵ , which is a measure of the relative strength of the two square-well attractions. The point of key importance in this phase diagram is the appearance of two distinct regimes, separated by a triple point, of different types of macroscopic interplay between the translational and conformational degrees of freedom. All the phase lines represent first-order phase

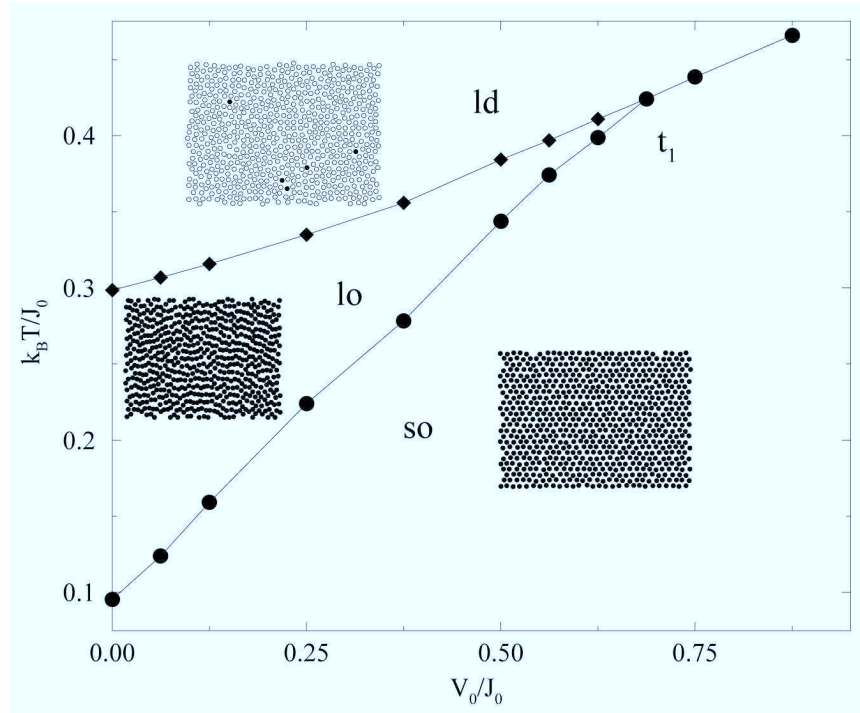


Figure 7. Phase diagram for the model of the pure lipid bilayer system. All three phase boundaries are first-order phase boundaries. The insets show snapshots of typical micro-configurations for the three different phases labeled **so** (solid-ordered), **ld** (liquid-disordered), and **lo** (liquid-ordered). Chains in the disordered state are plotted as (o) and chains in the ordered chain state as (•). The three snapshots are not given to scale. t_1 is the triple point described in the text.

transitions and a triple point is a phase point where two first-order phase lines join to form a single first-order phase line. The two degrees of freedom are uncoupled for values of V_0/J_0 smaller than the triple-point value, where two distinct ordering transitions from the **so** to the **lo** phases and from the **lo** to the **ld** phase take place successively. The first of these transitions is a solid-fluid transition where the conformational degrees of freedom do not change whereas the second transition is between two fluids where the conformation degrees of freedom change. However, the two degrees of freedom are macroscopically coupled for values of V_0/J_0 greater than the triple-point value and the intermediate **lo** phase vanishes. This is exactly the case of the main phase transition in single component lipid bilayers where one passes from directly from an **so** phase to an **ld** phase. We conclude, therefore, that although there is no fundamental reason why the intermediate **lo** phase cannot appear as a physical phase of a single-component lipid bilayer, the physical reality is such that the main phase transitions of single component lipid bilayers take place always above but close to the triple point. For our simulations for lipid-sterol systems, the values of the parameters in the model were then set such that the intermediate phase did not appear as a distinct thermodynamic phase of the corresponding system of pure lipid bilayers.

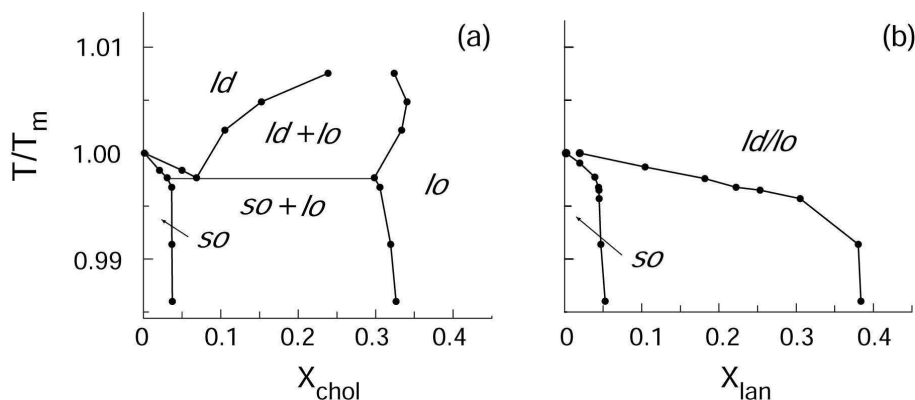


Figure 8. Theoretical phase diagrams determined from the Monte Carlo simulations of the microscopic model of interactions for the lipid-sterol membranes. (a) Lipid-cholesterol membranes; (b) lipid-lanosterol membranes.

4.1.2 Lipid-Sterol Bilayers

The simulated equilibrium phase diagrams for the lipid-cholesterol and the lipid-lanosterol membranes^{21,22} are shown in Fig. 8. They were obtained from simulations in the semi-grand canonical ensemble and are presented in terms of the sterol concentration x (where $x = x_{\text{chol}}$ or x_{lan}) and a reduced temperature, T/T_m , where T_m is the main-transition temperature for the system of the pure lipid membrane. In terms of microscopic interaction parameters, the lipid-cholesterol systems are distinguished from the lipid-lanosterol systems by a rather modest (about 10 per cent) increase in the strength of the microscopic interactions between a sterol molecule and a lipid chain in its conformationally ordered state (see Table 1). This was illustrated in Fig. 3b.

It is clear from Fig. 8 that this small modification in the lipid-sterol interaction strength leads to considerable differences in the overall topologies of the two phase diagrams. In the case of lipid-cholesterol systems (Fig. 8a), the most significant characteristic of the phase diagram is a stable region of coexistence between the *ld* and *lo* phases. Associated with this coexistence are necessarily a stable critical point and the existence of a three-phase line. The critical point is found to be located close to $T/T_m \approx 1.0075$, $x \approx 0.28$ for the interaction parameters used in the simulations. The temperature of the three-phase coexistence is estimated to be $T/T_m \approx 0.9977$, and the concentrations of cholesterol in the three coexisting phases are $x_{\text{so}} \approx 0.030$, $x_{\text{ld}} \approx 0.068$, and $x_{\text{lo}} \approx 0.28$, respectively. By contrast, Fig. 8b shows that no *ld-lo* coexistence can be identified in the phase diagrams for the lipid-lanosterol systems. Correspondingly, only a metastable critical point exists and there is no three-phase line.

The systematic development of the stable *ld-lo* coexistence as lanosterol ‘evolves’ to cholesterol is a macroscopic signature of an increase in the capacity of the sterols to stabilize the *lo* phase. Another consistent signature is that with the sterol ‘evolution,’ the *lo* phase boundary of the low-temperature *so-lo* coexistence moves towards lower sterol concentrations, indicating a broadening of the region of stability of the *lo* phase in lipid-cholesterol bilayer membranes. A comparison between Fig. 2 and Fig. 8 shows a close correspondence between experimental and simulation results for both phase diagrams.

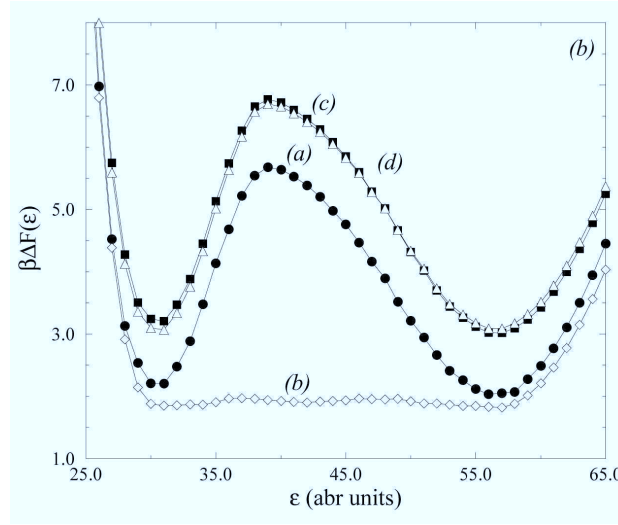


Figure 9. The different spectral free energies calculated using the modified potential-energy function. (a) The initial estimate of the shape function, $f(\epsilon)$; (b) The spectral free-energy function for the modified potential-energy function; (c) The spectral free-energy function for the original system and (d) (open triangles) the spectral free-energy function at phase coexistence. $\beta = 1/k_B T$.

4.2 Umbrella Sampling and Finite-Size Analysis

It may be instructive at this point to demonstrate, with specific reference to the phase diagrams, how the umbrella sampling technique and finite-size analysis were used to determine the phase diagrams^{21,29}.

First, Fig. 9 shows the results from the different steps in the iteration procedure of the umbrella sampling described previously, which was applied in the simulations of the theoretical model for single-component lipid bilayers. In this particular example, the transition between the **so** and the **ld** phases was investigated. At $T = 0.6$, a temperature just slightly below the transition temperature, an initial estimate for the spectral shape function was made for a small system $N = 14$. This estimate is shown in curve (a) of the figure and has a distinct barrier between the two minima representing different phases. The shape function for a larger system, $N = 16$, was then obtained by the rescaling given in Eq. (10) and the spectral free energy for the modified potential-energy function was derived from the new simulation based on the modified energy function, as shown in curve (b). The spectral free energy corresponding to the original potential-energy function (shown in curve (c)) was reconstructed. Finally, use of the re-weighting technique yielded the spectral free energy at coexistence $T = 0.6035$, now shown in curve (d).

Next, finite-size analysis was used to deal with the simulation data obtained for the two types of lipid-sterol systems in order to determine if phase coexistence occurred or not in the thermodynamic limit or if a critical point existed as a terminal point for the **ld-lo** coexistence. Figure 10 gives examples of such a finite-size analysis for the lipid-cholesterol system (refer to the phase diagram in Fig. 8a). This figure shows the spectral free energies, $\mathcal{F}(x)$, calculated as a function of the system size N using umbrella sampling, for the **so-lo** coexistence at $T = 0.860$ (Fig. 10a), the **ld-lo** coexistence at $T = 1.0035$

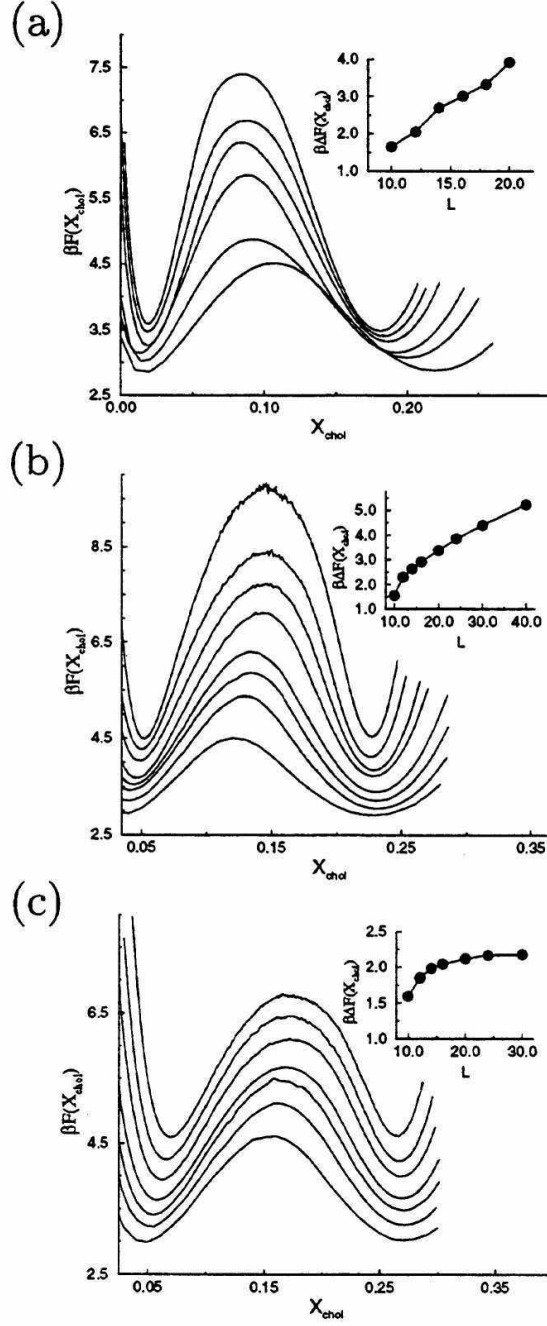


Figure 10. Finite-size scaling plots of both $\mathcal{F}_L(x_s)$ and the barrier height $\mathcal{F}_L(x_s)$ (the insets) at (a) $T = 0.6 T_m$ (**so-lo** phase coexistence); (b) $T = 1.35 T_m$ (**ld-lo** phase coexistence); (c) $T = 1.75 T_m$ (close to the critical point). The system sizes are $L=8, 10, 12, 14, 16$ and 20 . $\beta = 1/k_B T$.

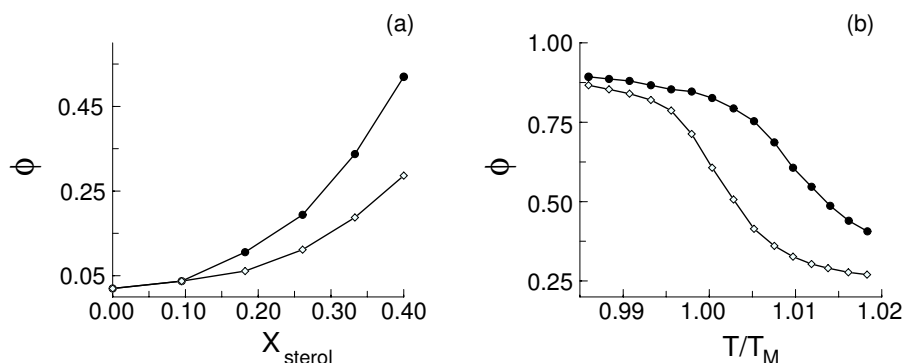


Figure 11. Theoretically calculated average of lipid-chain order parameter, ϕ , for the lipid-sterol membranes containing $N = 16$ particles. The solid circles correspond to lipid-cholesterol membranes, and the open diamonds correspond to lipid-lanosterol membranes. (a) ϕ as a function of sterol concentration at a fixed temperature, $T = 1.12 T_M$; (b) ϕ as a function of temperature for a fixed sterol concentration $x_s = 0.367$.

(Fig. 10b), and the **ld-lo** coexistence at $T = 1.0075$. Shown in the insets are the corresponding barrier heights as a function of ϕ for the three cases. In the first two cases, the barrier height increases linearly with ϕ when ϕ is sufficiently large. This finite-size behaviour is a characteristic of a first-order transition³⁴. In the third case, however, the barrier height approaches a constant value as ϕ is increased, indicating that $T = 1.0075$ is very close to the critical point terminating the **ld-lo** coexistence.

4.3 Conformational Ordering of the Lipid Molecules

In order to characterize quantitatively the differential effects of the two sterols on the physical properties of lipid-sterol bilayer membranes²², the calculated conformational order parameter, ϕ , as defined in Eq. (14), is presented in Fig. 11.

This thermal average represents the collective conformational ordering of the lipid molecules. Figure 11a shows ϕ as a function of sterol concentration for both types of the lipid-sterol systems at a fixed temperature $T = 1.012 T_M$. This temperature is above that corresponding to the critical point of the **ld-lo** coexistence region of the lipid-cholesterol system. At this temperature, neither of the two types of the lipid-sterol systems undergo any phase transitions as the sterol concentration is changed. Figure 11b gives ϕ as a function of temperature for a fixed sterol concentration $x = 0.367$. Clearly both cholesterol and lanosterol order the lipid chains, but the ordering effect of cholesterol is much stronger, qualitatively similar to what has been shown by the experimental data in Fig. 5. For example, Fig. 11a shows that cholesterol at $x = 0.40$ is able to rigidify close to 55 % of the lipid chains, whereas lanosterol at the same concentration can only rigidify roughly 30 %. This differential effect is also clearly illustrated in Fig. 12. Figure 12a and Fig. 12b are, respectively, snapshots of microscopic states of the lipid-cholesterol and the lipid-lanosterol systems in the **lo** phase.

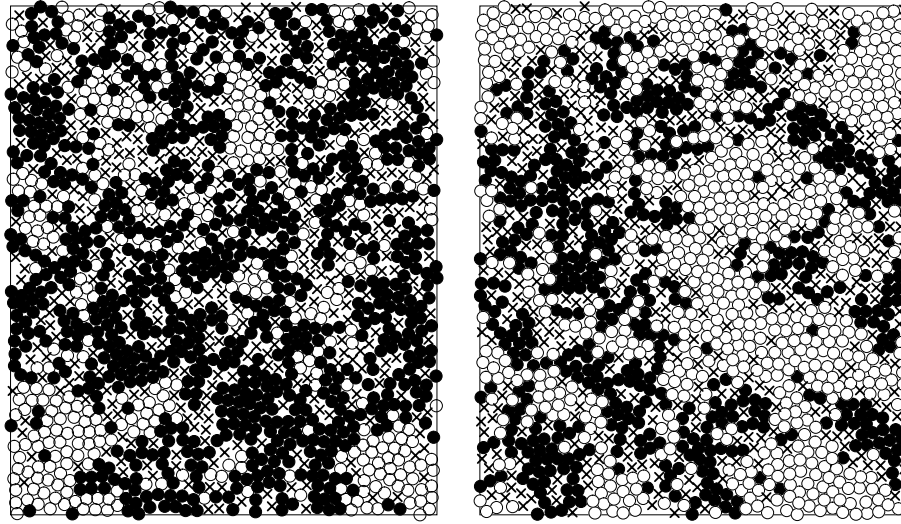


Figure 12. Snapshots of the micro-configurations for lipid-sterol systems in the **lo** phase very close to T_m when the sterol is cholesterol (the left-hand panel) and when the sterol is lanosterol (the right-hand panel). In the snapshots, a lipid chain in the ordered state is shown as (●), a lipid chain in the disordered state as (○) and a cholesterol molecule as (×).

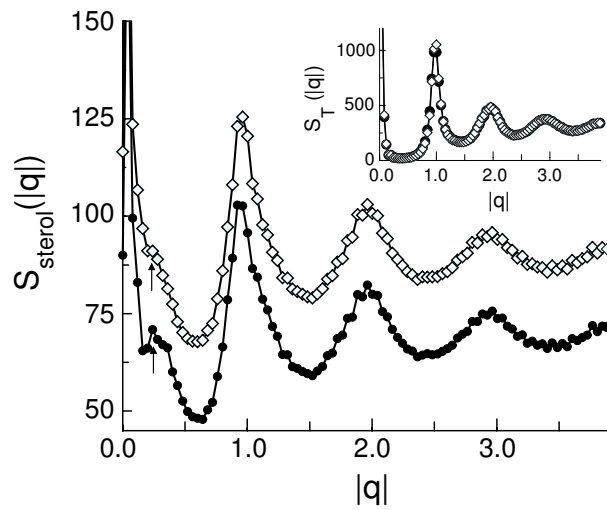


Figure 13. The partial structure factor describing the distribution of the sterol molecules in the membranes, $S_{\text{sterol}}(|q|)$, calculated at $T = 0.6 T_m$ and $x_s = 0.367$. $S_{\text{chol}}(|q|)$ is shown by the filled circles and $S_{\text{lan}}(|q|)$ by the open diamonds. For clarity, the curve for $S_{\text{lan}}(|q|)$ has been shifted along the y -axis. The values of $|q|$ are given in units of $2\pi/d$, where d is the hard-core diameter assigned to the particles. The arrows indicate an unusual structural signal in addition to the usual peaks characteristic of liquid structure. The inset shows the total structure factor, $S_T(|q|)$, for the two lipid-sterol systems.

4.4 Molecular Organization

The MC simulations of the microscopic model allows us to analyze quantitatively the lateral distribution of the molecules in the lipid-sterol systems^{21,22} in terms of the structure factors defined in Eq. (15). Figure 13 exhibits examples of the calculated structure factors calculated at $\phi = 0.806$ and $x = 0.367$, where the bilayers are in an **lo** state. The figure shows circular averages over the directions of the Fourier wave vectors, $\mathbf{q}$, of the sterol structure factor $S(\mathbf{q})$, characterizing the distributions of the sterol molecules. Shown in the inset are the circular averages of the total structure factor $S(\mathbf{q})$. The total structure factor has the typical features common to all liquid systems. The partial sterol structure factor, however, reveals more interesting structural information. Specifically, there appears a distinct, low- $|\mathbf{q}|$ peak at $|\mathbf{q}| \simeq 0.35 \cdot 2\pi/a$, where a is the hard-core radius as discussed in the section describing the theoretical model. This particular $|\mathbf{q}|$ -value corresponds to a real-space length scale that covers several molecules, if a representative value of 5 Å for a may be used. The signal almost disappears in $S(\mathbf{q})$ for the lipid-lanosterol systems.

4.5 Thread-Like Structures

Analysis of microscopic configurations such as those shown in Fig. 14 suggests that the special peak in the cholesterol structure factor is related to a micro-structure consisting of aligned ‘threads’²² of the cholesterol molecules, interspersed with ‘threads’ of lipid molecules with conformationally ordered chains as highlighted in Fig. 15. A qualitative analysis of the microscopic interactions in the model predicts that the stronger the interaction between a sterol molecule and an ordered lipid chain is, the more likely such micro-domains are to appear. This explains our observation through simulations that such micro-structures disappear almost entirely in the lipid-lanosterol systems, where the lipid-sterol interaction becomes weaker.

4.6 Lateral Diffusion

The simulation techniques used to monitor lateral diffusion in membranes were recently reviewed³⁵ and the differences between tracer diffusion and collective diffusion in relation to different types of experimental approaches were thoroughly discussed³⁶. The lipid tracer diffusion coefficient in a model lipid-cholesterol binary mixture was calculated as a function of cholesterol concentration and temperature^{37,36}. The model system that was studied was identical to that used in the calculation of the phase diagram of Fig. 8(a). The results are summarized in Fig. 16. The diffusion coefficient increases monotonically with increasing temperature and this qualitative trend is independent of cholesterol concentration. The cholesterol concentration dependence of D , however, shows more interesting behaviour. At higher temperatures, D decreases with increasing x , while at temperatures below T_c , D increases monotonically with x . At temperatures slightly greater than T_c , D first decreases, then increases slightly with increasing x , as can be seen more clearly in the insert to the figure. A qualitative interpretation of the physical origin of this behavior is provided by the free volume theory of diffusion. An increase in D with x is due to an increase in the free area per molecule with increasing x . Decreases in D arise from the fact that cholesterol also promotes conformational ordering of the lipid chains, which in turn causes the chains to interact more effectively, and

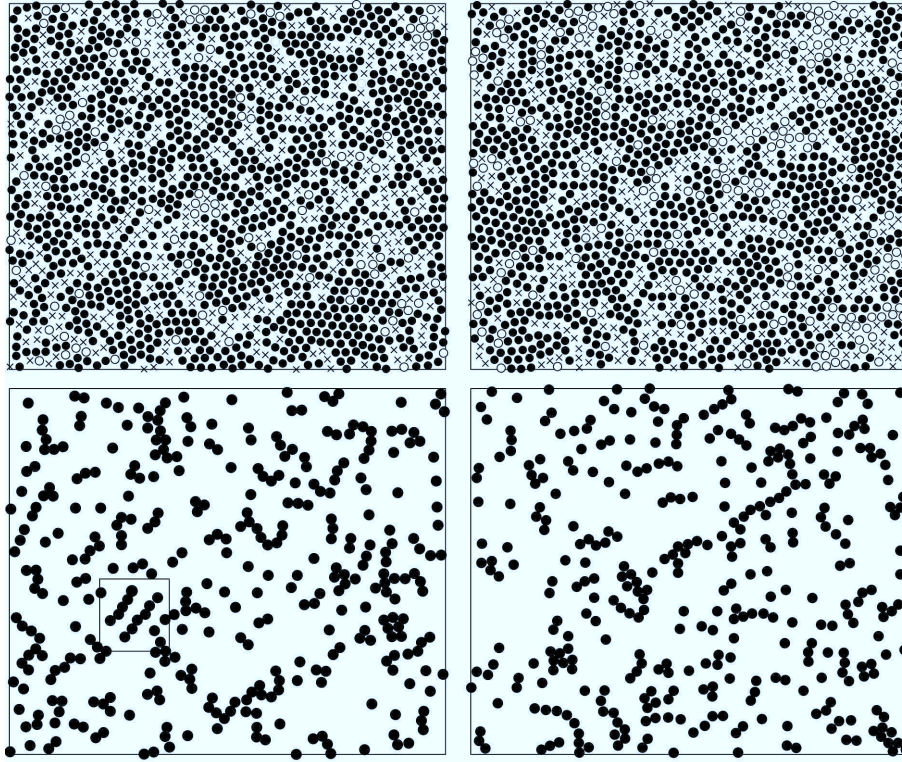


Figure 14. Snapshots of micro-configurations for lipid membrane systems containing cholesterol (left) and lanosterol (right), calculated at $T = 6 \cdot T_m$ and $x_s = .367$. The upper panel shows all particles, with lipid chains in the ordered state shown by filled circles, lipid chains in the disordered state shown by open circles, and sterol molecules shown by crosses. The lower panel shows only the corresponding sterol molecules as filled circles. The part highlighted by the box in the lower left snapshot is shown in detail in Fig. 15.

thus increase the effective activation energy for particle movement. These two effects compete with one another and give rise to the non-monotonic variation of κ with x at intermediate temperatures. The diffusion results of Fig. 16 are qualitatively consistent with those of a Fluorescence-Recovery-After-Photobleaching (FRAP) experimental study of lateral diffusion in dimyristoylphosphatidylcholine (DMPC)-cholesterol binary mixture model membranes³⁸. Thus, the minimal model designed to reproduce the phase behaviour of lipid-sterol mixtures also yields a correct description of the equilibrium dynamics.

5 The Final Words

In this set of lecture notes we gave a detailed description of our minimal, off-lattice models for lipid-cholesterol and lipid-lanosterol bilayers using random-lattice simulation techniques. The difference in behavior between these two sterols was modeled, based on their specific molecular characteristics, in terms of their differential interactions with lipid molecules, namely, that cholesterol tends to rigidify neighboring lipid chains more strongly

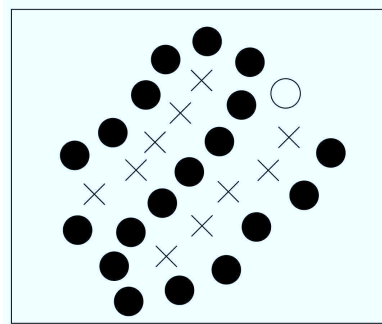


Figure 15. Enlarged version of the box highlighted in Fig. 14, showing a local 'thread-like' distribution of lipid and cholesterol molecules. Lipid chains in the ordered state are shown by filled circles, lipid chains in the disordered state are shown by open circles, and cholesterol molecules are shown by crosses.

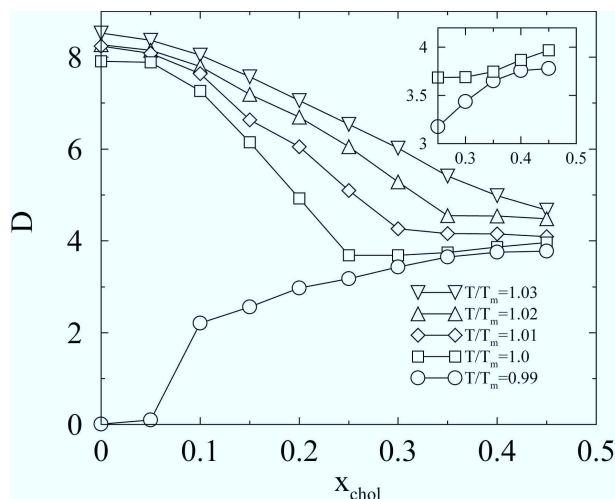


Figure 16. Diffusion coefficient D vs cholesterol concentration fraction x_s . The inset shows the data for T/T_m and 1. T_m in an expanded scale to illustrate the increase in D with x_{chol} in the **lo** phase. The diffusion results for D are given in units of $d^2/\tau \times 10^6$, where d is the hard-core diameter.

than lanosterol due to the relative smoothness of its β -face compared to that of lanosterol. We also discussed the MC method as well as the algorithm used in the MC simulations of the model. We also presented applications of various modern statistical-mechanical methods used to obtain the phase behavior of both lipid-cholesterol and lipid-lanosterol bilayers.

We were able to establish the occurrence/absence of the **lo** phase as a distinct, thermodynamically stable phase in the lipid-cholesterol and lipid-lanosterol system, respectively. We pointed out that this might have significant implications for understanding sterol evolution within the context of the structural evolution of membranes as well as the sterol requirement for the formation of rafts. In particular our results correlate with the current

hypothesis that cholesterol is an important factor in the formation of rafts and their stability and that lanosterol is unable to stabilize rafts.

We were also able to use our model to predict the behavior of the diffusion constant in lipid-cholesterol systems as a function of both temperature and cholesterol concentration. No additional parameters or parameter values were used in these simulations and the results were found to agree qualitatively with experimental results. It is significant that the minimal theoretical models have both descriptive and predictive capabilities, as this shows that the essential microscopic physical mechanisms have been included correctly in the models.

Acknowledgments

This work was supported by the Danish Natural Science Research Council, le FCAR du Quebec (Canada) via centre and team grants, and the NSERC of Canada via operating and equipment grants. MEMPHYS – Center for Biomembrane Physics is supported by the Danish National Research Foundation. Both MN and LM acknowledge support from the Danish Research Academy.

References

1. The present set of lecture notes is based on a recent review article by the present authors, M. J. Zuckermann, M. Bloom, J. H. Ipsen, L. Miao, O. G. Mouritsen, M. Nielsen, J. Polson, J. Thewalt, I. Vattulainen, and H. Zhu, *Meth. Enzymol.* (in press, 2003).
2. S. I. Singer and G. L. Nicolson, *Science* **175**, 720 (1972).
3. O. G. Mouritsen and O. S. Andersen (eds.). *In Search of a New Biomembrane Model*, *Biol. Skr. Dan. Vid. Selsk.* **49** (1998) 224.
4. K. Simons and E. Ikonen, *Nature* **387**, 569 (1997); D. Brown and E. London, *J. Biol. Chem.* **275**, 17221 (2000).
5. F. R. Maxfield, *Curr. Opin. Cell Biol.* **14**, 483 (2002).
6. U. Esmann and M. L. Berkowitz, *Biophys. J.* **76**, 2081 (1999).
7. P. B. Moore, C. F. Lopez, and M. K. Klein, *Biophys. J.* **81**, 2484 (2001).
8. H. L. Scott, *Curr. Opin. Struct. Biol.* **12**, 495 (2002).
9. L. Saiz, S. Bandyopadhyay, and M. L. Klein, *Bioscience Rep.* **22**, 151 (2002).
10. C. Hofsass, E. Lindahl, and O. Edholm, *Biophys. J.* **84**, 2192 (2003).
11. M. Patra, M. Karttunen, M. Hyvönen, E. Falck, P. Lindqvist, and I. Vattulainen, *Biophys. J.* **84**, 3636 (2003).
12. R. Böckmann, A. Hac, T. Heimburg, and H. Grubmüller, *Biophys. J.* **85**, 1647 (2002).
13. A. M. Smondyrev and M. L. Berkowitz, *Biophys. J.* **80**, 1649 (2001).
14. S. W. Chiu, E. Jakobsson, R. J. Mashl, and H. L. Scott, *Biophys. J.* **83**, 1842 (2002).
15. See for example: G. W. Feigenson and J. T. Buboltz, *Biophys. J.* **80**, 2775 (2001).
16. J. H. Ipsen, G. Karlström, O. G. Mouritsen, H. Wennerström, and M. J. Zuckermann, *Biochim. Biophys. Acta* **905**, 162 (1987); J. H. Ipsen, M. J. Zuckermann, and O. G. Mouritsen In *Cholesterol in Model Membranes* (L. X. Finegold, ed.) CRC Press, Boca Raton, Florida, 1993, pp. 223-257.

17. S. Doniach, J. Chem. Phys. **68**, 4912-4916 (1978).
18. O. G. Mouritsen, B. Dammann, H. C. Fogedby, J. H. Ipsen, C. Jeppersen, K. Jørgensen, J. Risbo, M. C. Sabra, M. M. Sperotto, and M. J. Zuckermann, Biophys. Chem. **55**, 55 (1995).
19. O. G. Mouritsen. In *Advances in the Computer Simulations of Liquid Crystals* (P. Pasini and C. Zannoni, eds.) Kluwer Academic Publ. Dordrecht, 2000, pp. 139-187.
20. M. Nielsen, L. Miao, J. H. Ipsen, O. G. Mouritsen, and M. J. Zuckermann, Phys. Rev. E **54**, 6889 (1996).
21. M. Nielsen, L. Miao, J. H. Ipsen, M. J. Zuckermann, and O. G. Mouritsen, Phys. Rev. E **59**, 5790 (1999).
22. L. Miao, M. Nielsen, J. Thewalt, J. H. Ipsen, M. Bloom, M. J. Zuckermann, and O. G. Mouritsen, Biophys. J. **82**, 1429 (2002).
23. M. C. Sabra and O. G. Mouritsen, Meth. Enzymol. **321**, 263 (2000).
24. M. Vist and J. H. Davis, Biochemistry **29**, 451 (1990).
25. J. Thewalt and M. Bloom, Biophys. J. **63**, 1176 (1992); J. Thewalt, F. M. Hanert, M. Liseisen, A. J. Farral, and M. Bloom, Pharmacology **42**, 9 (1992).
26. K. Bloch, Science **150**, 19 (1965); CRC Crit. Rev. Biochim. **14**, 47 (1983).
27. X. Xu and E. London, Biochemistry **39**, 843 (2000).
28. An excellent textbook for both MC and Molecular Dynamics is: D. Frenkel and B. Smit, *Understanding Molecular Dynamics*, 2nd Edition, Computational Science Series, Vol. 1, Academic Press, 2002.
29. M. Nielsen, Ph. D. Thesis, McGill University, Montreal, 1999.
30. S. Marcelja, Biochim. Biophys. Acta **367**, 165 (1974).
31. D. A. Pink, T. J. Green, and D. Chapman, Biochemistry **19**, 349 (1980).
32. J. Risbo, Ph. D. Thesis, Technical University of Denmark, 1997; J. Risbo, G. Besold, and O. G. Mouritsen, Comp. Mat. Sci. **15**, 311 (1999).
33. A. M. Ferrenberg and R. H. Swendsen, Phys. Rev. Lett. **61**, 2635 (1988); *ibid* **63**, 1195 (1991).
34. J. Lee and J. L. Kosterlitz, Phys. Rev. Lett. **65**, 137 (1990); Phys. Rev. B **43**, 3625 (1991).
35. I. Vattulainen and O. G. Mouritsen. In *Diffusion in Condensed Matter* (J. Kärgner, P. Heitjans, and R. Haberlandt, eds.) Springer-Verlag, Berlin (in press, 2003).
36. I. Vattulainen, J. H. Ipsen, and O. G. Mouritsen, Biophys. J. (submitted, 2003).
37. J. M. Polson, I. Vattulainen, H. Zhu, and M. J. Zuckermann, Eur. Phys. J. **5**, 485 (2001).
38. P. F. F. Alameida, W. L. C. Vaz, and T. E. Thomson, Biochemistry **31**, 6739 (1992).

Statistical Mechanics of Protein Folding: Some Outstanding Problems

A. Grosberg

Department of Physics, University of Minnesota, Minneapolis, MN 55455, U.S.A.
and
Institute of Biochemical Physics, Russian Academy of Sciences, Moscow 117977, Russia
E-mail: grosberg@physics.umn.edu

We consider broken ergodicity, memorization of random choice, and spontaneous symmetry breaking in both sequence space in the evolution and in conformation space in connection with folding. We pay major attention to the topology of conformation space, its connectivity, and the determination of both the order parameters and the (free) energy landscapes.

1 Introduction

It is useful to start the introduction with the recommendation to the reader. If one wants to study proteins, then there is a very good book to jump-start the process, it is “Lectures on Protein Physics,” by Alexei Finkelstein and late Oleg Ptitsyn¹.

1.1 Basic Terminology

Proteins are heteropolymer molecules, consisting of 20 different monomeric units, called amino acid residues, or simply amino acids. These units are connected in a chain-like topology, such that the main chain of the polymer, called polypeptide chain, is universal, does not depend on the specific types of the amino acids, while different amino acids form different side groups at every chain monomer.

Proteins are characterized by their primary, secondary, ternary and quaternary structure. The primary structure is defined as the sequence of amino acids along the polymer (polypeptide) chain. The fact that polypeptide chain is regular, sequence independent, gives rise to the formation of certain regular conformational elements, such as helices or sheets (called, since Linus Pauling, α -helices and β -sheets). These elements are usually referred to as secondary structure. However, in order to perform its biochemical function, the overwhelming majority of the proteins acquires a complex three-dimensional structure during the protein folding process. Protein folding is the process in which molecule self-organizes into a native three-dimensional structure, called tertiary structure. Finally, the quaternary structure is the arrangement of the constituting tertiary structure domains, belonging either to the same or to different polymer chains.

The difference between secondary and tertiary structure frequently seems confusing for the beginners. One aspect of the difference, which appears purely quantitative, if not vague, is that secondary structure, such as spiral-like α -helix, involves relatively short pieces of the chain, while folding of the tertiary structure involves spatial scales of the entire molecule. Another, possibly more clear aspect, is that secondary structure motives are determined by the geometry (or stereochemistry) of the main polymer chain. For instance,

in α -helix, there are hydrogen bonds between monomers with numbers n and $n+3$ along the chain (for some series of subsequent numbers n). Historically, people thought that energetics of the secondary structure elements is also due to the interactions between the main chain atoms, or, better to say, due to the competition between hydrogen bonds of the (main chain) peptide groups to each other or to surrounding water. In fact, bulky side groups, although located outside α -helix, or above and below the β -sheet, can be in proximity to each other and, therefore, can affect energy and stability of the corresponding secondary structure elements. In other words, the stability of secondary structure is strongly affected by the tertiary structure, and vice versa. Nevertheless, the distinction between secondary and tertiary structure is fairly sharp.

The protein folding process, during which the heteropolymer chain reaches its folded, native structure, is one of the central, but not yet completely understood problems of the molecular biology. This lecture gives an introduction to some of the most pressing issues of protein folding from the viewpoint of the theoretical physics (or, it maybe better to say, from the personal subjective viewpoint of the present author).

1.2 Folding and Re-Naturation

To describe the “creative” process of protein folding, it is convenient to start with the “destructive” process of protein denaturation. Denaturation is what happens when one boils an egg. In a better controlled experiment, one can take a very dilute solution of native proteins, where every protein molecule is very far and independent from any other, and thus acts independently of others (unlike in the egg). Then, when the solution is heated up, something happens. For instance, if the protein is an enzyme, it keeps working as a catalyst up until certain temperature, called denaturation point, above which no appreciable catalytic activity is observed.

The story of denaturation studies is a long and interesting one. For instance, it was found that temperature of denaturation for every protein is not very far above the temperature at which host cell normally lives and functions: proteins from bacteria living on glaciers denature at about 20° , while organisms living in the hot water of geiser have their proteins stable up until 140° or so. Another exciting observation is that denaturation is a cooperative all-or-none transition, it is a finite system counterpart of the first order phase transition². Denaturation transition is accompanied by a significant latent heat, usually about 7×10^{-4} kcal/mol. Per unit mass, it is about the same order of magnitude as the latent heat of melting of usual molecular crystals, such as naphthalene. Finally, denaturation can be caused not only by heating, but also in many cases by cooling down (“cold denaturation”), as well as by changing pH, salinity, etc, and, most easily, by adding to the solution some “denaturing agents,” most of which disrupt hydrophobic interactions (such as urea).

What happens to the molecule when protein denatures? In many cases, although not in all cases, it is globule-coil transition: native globule opens up and becomes essentially a Gaussian coil, in which volume interactions between aminoacids are of minor importance. In some other cases, native globule gets transformed into another globular form, called molten globule, which does not have so uniquely determined tertiary structure. Once again, we should refer the reader to the book¹ for further details.

In fact, when protein denatures, one should be wondering if some irreversible changes occurred, such as breaking of covalent bonds connecting protein polymer chain together.

Indeed, hard boiled egg is certainly irreversibly changed from the fresh, no chicken can be produced from it by any manipulations. Is the same true for proteins denatured in the dilute solution? It was a major discovery by C. Anfinsen about 50 years ago that gently denatured proteins can be returned back to their native states, their enzymatic activity can be restored, and, therefore, nothing irreversible happens to them during denaturation. The phenomenon of return to the native state is called re-naturation. After this was first established in classical experiments on ribonuclease by C. Anfinsen³, renaturation was observed with many proteins.

Anfinsen's result means, first of all, that all the information necessary to form the native tertiary structure is contained in the molecule itself, because renaturation is observed in a very dilute solution, where protein molecules do not interact. Therefore, all the necessary information is contained in the primary sequence. In Anfinsen's own words in his Nobel acceptance speech in 1972⁴

“The native conformation is determined by the totality of interatomic interactions and hence by the amino acid sequence, in a given environment.”

As we already said, not all of the proteins renature starting from completely open coil-type state. Furthermore, not all of the proteins renature at all. It is however most important to realize that there **exist** many proteins which do not require anything, like any external help or assistance, to fold, their folding is an example of true *self*-organization. A lot of proteins will correctly fold into their native state in very dilute water-based solutions, provided only with appropriate pH and salt concentration. There are some other proteins which do require assistance to fold; they only fold in the presence of chaperons. Note also that not all of the proteins function in the folded globular state. Some, such as, e.g., collagen function as fibrils. Others are attached to membranes. These lectures do not discuss these cases. However important biologically, all these aspects are of secondary relevance to the physics approach to the subject of folding. Since there are proteins which do fold without any help, it is clear that they should be the first subject of attention and they should be understood first.

As Anfinsen experiment suggests beyond a doubt, native state of a foldable protein is stable and thus represents a minimum of free energy. This does not mean it is necessarily the global minimum. In other words, Anfinsen's experiment leaves open the possibility that native state is just a local free energy minimum and that there is one or more even more stable states, which correspond to deeper minima. This possibility is very likely realized in some cases (e.g., prion proteins), when both correctly folded and some misfolded states are realized under different circumstances. These misfolded states are very important, because they are responsible for a number of diseases, like, e.g. some illnesses of the nervous system (“mad cow disease”). This already suggests the importance of both thermodynamic (equilibrium) and kinetic aspects of folding, because folded state may be either thermodynamically most stable (global minimum), or thermodynamically metastable (local minimum), but kinetically most easily accessible.

1.3 Quenched, but not Disordered

It is useful to emphasize from the very beginning that the physics of protein conformations, including the folding phenomenon, is closely related to the branch of physics which

is called physics of disordered (or, sometimes, “dirty”) systems, such as doped semi-conductors. The relation between proteins and disordered systems may appear bizarre and unfounded. Indeed, proteins are very pure: each protein molecule is an exact copy of every other molecule of the same protein, they have identical molecular mass, identical number of monomers in the chain and, moreover, identical sequences of links along the chain. The confusion here is entirely due to the historically accepted terminology, which for the regular systems, such as semi-conductors, concentrates on the aspect of disorder. Indeed, donor and acceptor atoms in a semi-conductor are positioned randomly. These samples are manufactured from a melt, a liquid mixture of different atoms, and when melt is quenched and frozen, the minority atoms in the host lattice are essentially fixed at the random places where they happen to be at the moment of quench.

Similar situation is known in the materials called spin glasses. Those are the dilute solutions of paramagnetic atoms in the non-magnetic matrix. Once again, if we start from a melt, and then quench it, then paramagnetic centers are caught in the random places where they happen to be at the moment of quench.

When such systems are called disordered, it is meant to emphasize that although they are solids, they have no translational or any other simple order in their structures. In protein context, this aspect of disorder is neither interesting nor productive. More important is the aspect that the structure is quenched. That means, for the spin glass example, that the spatial positions of the paramagnetic centers are fixed and, once the sample is prepared, they do not change. This is interesting, because every paramagnetic atom has obviously a spin, and exchange interactions between these spins is determined by the distance between them in space. Thus, in essence, when spin glass sample is quenched, what happens is we have a system of spins, $\{S_i\}$, with interaction between any two spins i and j , J_{ij} , where matrix J_{ij} is quenched. Thus, physics of spin glasses is the statistical mechanics of the spin system which is pretty much like the regular Ising system, except interactions matrix is not regular, but quenched.

In proteins, sequence, or primary structure is under strict genetic control. This means, every protein molecule is prepared by the protein synthesis cell machinery under DNA and RNA guidance, such that all copies of the same protein have identical sequences. When protein is prepared, its sequence does not change during folding or any other conformational transformations. In this sense, proteins are the systems with quenched sequences. In this respect, whether we call proteins ordered or disordered is more emotional than scientific issue, the only thing that is important is that they are quenched.

This said, it is important to understand where do these sequences come from. In fact, they are the product of biological evolution.

1.4 Quenching and Self-Averaging

In order to understand the above concepts, let us consider a “photo” of an Ising spin system, where the spins are oriented randomly. The randomness can be caused by two reasons: either we are above the critical temperature in the uniform Ising system with ordered interactions, or we deal with irregular system with disordered interactions.

In the first case the spin configuration changes in time, and it is a realization of the thermal noise. In the second case the spin configuration also changes in time, but does so on the background of a “prepared” interaction pattern which does not change in time. How-

1.5 Plan

In this lecture, we shall discuss basic issues in the statistical mechanics modeling of proteins, which involves both sequence evolution and conformation dynamics.

2 Sequences

2.1 Diffusion in the Sequence Space

Evolution of sequences is itself an exciting problem with a large statistical mechanics component. Needless saying, it has got particularly exciting now, with plenty of uncovered experimental (statistical) material. The general understanding of the area is impossible without the realization of the basic fact: evolution of sequences is about as old as the Universe itself. Indeed, according to the modern estimates, the age of the Universe since the Big Bang is about 15 billion years, the age of the Earth is about 4.5 billion years, and the oldest clear signs of life on the Earth are about 3.8 billion years old. So, all these numbers are fairly close. Thus, speaking about evolution of sequences, we should have in mind an event which happened once and did not have time to be repeated to explore different possibilities.

Let us look at evolution mathematically, considering it as some sort of diffusion in the sequence space. This sequence space is obviously a discrete one: if we have, say, Σ possibilities for every monomer (Σ letters in the alphabet, $\Sigma \approx 20$ for real proteins), then we only have a finite number of sequences Σ^L , where L is the maximal acceptable length. This space is big, because the number Σ^L is astronomical. In fact, this number is so large that, for instance, H. Fraunfelder⁵ suggests to call it a “biological number,” where biological numbers dwarf astronomical ones. Two ingredients are vital to imagine diffusion in such space: first, we should allow some dynamics, which means there should be some mechanism allowing sequences to change from time to time producing some other sequences; second, the system must be able to decide which of the newly produced sequences are good and which are bad, which means there should be a feedback.

A few comments are in order. In real biology, the mechanism of sequence change is incredibly complex, it involves mutations in DNA, biosynthesis of proteins based on mutated DNA, their folding, and inheritance of the mutated DNA to new generations of organisms. By the way, real mutations in DNA are not only point mutations, there are things as sophisticated as swapping of long blocks etc.

Let us simplify this as far as possible. Namely, let us imagine that the time is discrete, and pretend that at every time step, with every “click of the evolutionary clock” every sequence, s , can jump and transform into some other sequence, s' . In reality, these time steps are perhaps close to the life time of one cell before it produces the offspring.

This gives rise to a convenient mathematical scheme. Let us imagine every possible protein sequence as a dot and let us connect two dots by a line if they can be transformed into one another in one time step. We call these lines “mutations,” although they are not real mutations, real mutations occur in DNA, and we speak of the effective “mutations,” which are the jumps in the space of protein sequences. Thus, with sequences represented as nodes and connected by bonds representing “mutations,” our sequence space is now presented as a graph.

Now we return to the question of the feedback. Arriving at every sequence, or every node on the sequence space graph, our system has to be able to decide whether this sequence is good and should be multiplied, or it is bad and should be eliminated. Phenomenologically, we should imagine that there is some function f which is defined for every sequence, $f(s)$, and which measures fitness of this sequence. By analogy with energy, it is common to define $E(s)$ such that the better fitness corresponds to the lower values of $E(s)$.

In real biology, once again, this is incredibly complex. The “test” involves folding of a protein with the new sequence, its function, and whether it is appropriate for the cell and the organism. In many cases, perhaps in the majority of cases, the test is an outright failure: either the polymer does not fold, or does not function well enough, or does some other nasty things, such as aggregates with something else. Rather rare steps lead to something very good.

According to Kimura and his neutral theory of evolution, there are also great many steps which are neutral, they lead to the sequences which under current conditions are neither better nor worse than existing ones. This neutral theory suggests that diffusion over neutral regions of the sequence space creates a pool of sequences to choose from when the conditions change.

Now, we can try to formulate the mathematical model of diffusion in the sequence space. One way to do this is to use the Metropolis Monte Carlo method, considering fitness as effective energy (divided by $k_B T$). During the “evolution” represented by the Monte Carlo dynamics of the Metropolis algorithm, a biased diffusion will take place in the sequence space. The equations for the “flux” from site s to s' or back are such that the system generally moves in the direction of decreasing E , however, with some small but non-zero probability it can also move in the opposite direction, which allows overcoming the local barriers and prevents the system from falling down in a local trap. More comments about Metropolis method will be given later, when we shall use it to study diffusion in conformation space.

2.2 Spontaneous Symmetry Breaking and Memorization of The Random Choice

The Metropolis scheme is not the only possible one to describe biased diffusion in the sequence space, nor is it necessarily the best one. It is brought here as just an example. One can very easily argue about lots of important factors which can be included only in a doubtful way. Perhaps the most important is the fact that proteins interact strongly, they interact in both physical and information sense, which means that the performance, or fitness of a given sequence, depends strongly on the set of other sequences currently in the pool. However correct are all the critical remarks, there is one general conclusion which is robust, universal and fully insensitive to the details: it is that evolution did not have time to visit all sites in sequence space, the evolution had time to try only a very small fraction of the sequences.

Indeed, it is a very general property of diffusion type equations in a restricted domain to exhibit two distinct regimes, which correspond to early time and late time asymptotics. The late time asymptotics is characterized by exponential approach of the distribution to the “ground state” eigenfunction of the corresponding operator. The existence of such eigenfunction is guaranteed when the domain is restricted. This late time asymptotics

corresponds to exhaustive (or ergodic) sampling of all possibilities. This is what happens in the thermodynamic equilibrium. On the other hand, the early time asymptotics corresponds to the initial spread out of the diffusing cloud (or probability distribution), in this regime overall size and shape of the domain is not relevant. In this sense, evolutionary diffusion in the sequence space is currently in the very early stage, it is very far from exhaustively sampling all possible sequences. This most fundamental fact easily follows from the very simple estimate. Total number of sequences of the length, say, 200, is $20^200 \approx 10^{60}$. This is a “biologically large” number⁵. It is easy to establish that even if we use all the mass of the Earth (6×10^{24} g) to make proteins (mass about 3.4×10^{-22} g each), and will change all the sequences every nanosecond, even then the total number of tested sequences can reach “only” about 10^{10} , which is still nowhere near the 10^{60} required. Therefore, most of the possible protein sequences have never been “tested”: we are in the short-time limit. In particular, we cannot say that current proteins are the best possible ones, because Mother Nature did not try most of others...

The fact that all sequences could not possibly be tested indicates that the existing choice of sequences has a significant element of randomness to it. As Oleg Ptitsyn expressed it once⁶, proteins are slightly edited random copolymers. Edited - because they did go through some evolutionary selection; but edited only slightly - because we are still in the early time limit in terms of diffusion in the sequence space.

3 Conformations, Conformation Space, and the Levinthal Paradox

3.1 Conformation Space: Lattice Polymer

Computationally most tractable polymer model is the lattice one. It represents the polymer as a self avoiding walk on a (cubic) lattice. An appropriate set of local moves is usually defined to allow the dynamics realized as a succession of these moves. In the most common Stockmayer-Verdier model⁷, these moves include end flip, corner flip, and crankshaft flip. The only requirement for the set of local moves is that they have to provide for ergodicity: every conformation should be possible to transform into every other conformation. Importantly, ergodicity does not require that the path from one conformation to another be short or simple, it only requires that such path exists.

We now have to make one step further in the direction of abstraction and ask: what is the space of conformations for such lattice polymer? This space is in fact a graph^{8,9}. Each vertex of the conformational graph represents one particular conformation. Two vertices of the graph are connected by a bond if and only if the corresponding conformations can be transformed into one another by a *single* elementary move. When all conformations are included, this graph is connected, which is to say that the system is ergodic. In general, this graph, or certain regions of it, may have non-trivial fractal dimension¹⁰. In addition, certain parts of conformation graph are shown to be of a small world network type¹¹. Finally, if we restrict consideration with only compact conformations, then ergodicity may be broken¹²; controlling, say, maximal gyration radius of the polymer in real space we can observe percolation transition in the graph of conformations.

3.2 Conformation Space: Off-Lattice Polymer

In the simplest off-lattice model, the chain conformation is fully described by the set of monomer position vectors, $\mathbf{r}$. Naively, this suggests that the space of conformations for the N -mer is a $3N$ -dimensional Euclidean space. This is wrong. To see why and how this is wrong, it is enough to consider a “polymer” with two monomers ($N=2$) with fixed bond length: apart from translational freedom of the first monomer, the “space of conformations” is, of course, a sphere - curved surface with non-Euclidean internal geometry. This very simple argument can be pulled quite far to show that conformation space is very much curved for long polymers, and its internal geometry is not at all Euclidean¹³. One should not be distracted by the fact that conformation *can* be fully characterized with the Euclidean position vectors $\mathbf{r}$ - this only means that the curved space of conformations can be *embedded* in a bigger Euclidean space - just like we usually imagine a sphere in a 3 space. This embedding in no way cancels or downplays the importance of the fact that internal geometry of conformation space is not Euclidean, because of course all the dynamical trajectories of the systems are in this space, not in the bigger flat one.

Clearly, the difference between on-lattice and off-lattice cases is neither deep nor important in this context. To understand it, let us begin with neglecting the excluded volume constraints. In this case, lattice model has just 6 (on the cubic lattice in 3) possible positions for every bond. Similarly, in an off-lattice model based on rotational isomers, there are a few permissible rotational states for every bond. With 2 possible states for every one of N monomers, the graph of conformations would have been an N -dimensional hypercube; similar figure with 6 or other finite number of vertices along each of the axis has no special name, so, for the lack of a better word, let us call it a “cube”. Of course, imposition of the excluded volume constraints erases many (in fact, even the majority) of the “cube” vertices, the ones which correspond to conformations with overlapping monomers. Nevertheless, the remaining part, which is the real conformations graph, is still a part of a cube, which obviously has nothing to do with the regular lattice - discrete analog of a flat Euclidean space. Similarly, for the off-lattice model with continuous set of rotations, conformation space is compact (in the strict mathematical sense of the word) and curved.

Thus, the stage of the protein folding drama is a compact space, some part of a “cube” or its continuous analog. It is terribly curved and restricted, it has complex topological and fractal properties. How can we understand the motion of a protein in such space?

3.3 Levinthal Paradox

The most pragmatic formulation reduces protein folding problem to that of prediction: knowing the amino-acid sequence of a protein, how to determine its ternary fold? This is potentially a multi-billion dollar issue, as it involves many protein-based drugs. To achieve this goal is very difficult, the problem resists best efforts of best scientists for a few decades now. Why is it so difficult? The nature of difficulty is well explained by the celebrated Levinthal paradox¹⁴. In the most standard formulation, Levinthal paradox arises from the idea that the time required for a protein molecule to sample all of its conformations is **at least** M , where M is the number of distinct conformations, and τ is the time necessary to sample one conformation. Then, the paradox goes, unguided folding into one particular (native) state requires **at least** time of order $M\tau$ which is far too long, because M is astronomically (biologically?) large. Similar to the sequence evolution case described

above, it is easy to establish an emotionally impressive fact, that the age of the Universe is by far insufficient for one typical length (≈ 200 , $M \approx 10^4$) protein to sample all possible conformations even if we assume τ as little as 1 ns.

Of course, the above formulation relies on the discrete conformation space, such as the conformation graph of a lattice polymer. However, the idea remains the same for continuous models as well.

Levinthal paradox was a subject of numerous discussions over the years. One of them was held recently in the Journal of Biomolecular Structure & Dynamics¹⁵. The present author contribution to this discussion emphasized the role of topology of conformational space. Specifically, the time estimate of the exhaustive search through the conformations depends on how these conformations are mutually arranged, how protein can move from one conformation to another. This aspect was long underappreciated, and deserves more detailed presentation.

3.4 Connectivity of the Conformation Space and Scaling of Levinthal Time

It seems that the idea of this section is well illustrated by the following story told in the author's contribution to the discussion¹⁵.

A long time ago, well before the breakup of Soviet Union, a large biophysics meeting was held in then Soviet Republic of Georgia. The site was a rural place in the center of a famous wine producing region, the month was October, and the major event was the all-Georgian wine testing festival, advertised as a merry traditional peasant holiday. Upon arrival, biophysicists found a large open field in the valley between mountains, covered with dozens of pavilions, each representing a particular village, and each offering for free a glass of young wine. Very soon, the cloud of biophysicists seemed perfectly obeying the diffusion equation, with each individual in the cloud undergoing random walks.

Wine testing continued for an unexpectedly long time. Assuming visit to one pavilion takes time τ , and assuming there were some M pavilions, one could have naively expected that after time close to $M\tau$ the testing would be over. Such expectation proved totally wrong. One possible reason is trivial: wine testers, even if they were to complete the exhaustive testing, are unlikely to realize the completion of the task and to stop at that. But there is also another more interesting reason: the time necessary for the random walk to visit all M sites does not scale as M , it can be significantly larger than that, simply because random walk visits some sites great many times before the first visit to some other sites.

Indeed, in terms of wine testing, $M\tau$ is the time required to visit all pavilions in an orderly fashion, one pavilion after another, never returning to the already visited place. Of course, a sober person can do that, but sober model is unrealistic for the wine tester. Equally unrealistic is the model of protein chain dynamics which orderly samples all conformational states, one after another, never returning to the once visited conformation. For all other sampling strategies, the exhaustion time is larger than $M\tau$. Of course, Levinthal did not say that the time of exhaustive conformation sampling (or wine testing) was $M\tau$ - he said it was **at least** $M\tau$; in other words, he said it is $\geq M\tau$. This was sufficient for him to conclude that exhaustive sampling is impossible for realistic τ , such as 150 or 200 ns.

The opposite, and perhaps more realistic, model of protein dynamics (and also of wine

testing) would be purely random walk in the space of conformations. To begin with, suppose it is an unbiased random walk, which means there is no conformation dependent (free) energy landscape involved. How can we estimate the exhaustive sampling time for the unbiased random walk model? Consider first that wine testing pavilions arranged along a line. Then random walk of longevity τ brings us as far as about $\sqrt{\tau}$, which means we cover all M sites when $\sqrt{\tau} \simeq M$, and the time of exhaustive sampling is $\tau \simeq M^2$. Needless saying, the difference between M and M^2 is very significant.

In fact, accurate estimate of exhaustive sampling time by a random walk is not completely trivial. More sophisticated estimate for one dimensional case, which will not be derived here, reads $\tau \simeq M^2 \ln M$. For the random walk in the space of higher dimension d , the result depends on d . When d crosses over 2, the mechanism of sampling changes, because random walk tends to leave behind large unvisited regions. At $d = 2$, exhaustive sampling is only possible because the overall volume is restricted, and random walk is forced to come back.

What is d in reality is anybody's guess. Please do not forget that d here is the dimension of the abstract space of protein conformations, not the real three-dimensional space.

The result close to M^2 would be correct for the dimension d as high as M ; in this case, exhaustion time would have scaled as $M^2 \ln M$. However, this estimate is completely unrealistic, because $M^2 \ln M$ corresponds to the situation where each conformation (site in conformation space, or wine pavilion) can be equally probably reached from every other conformation in just one step. Clearly, real protein chains are nowhere near this extreme.

3.5 Levinthal Exhaustion Time and Folding Time

In fact, the exhaustive sampling time of all conformations is not the quantity of major physical interest. Indeed, we are not interested in trying all conformations, we want to know how long it takes to arrive at the particular native state. More specifically, in the language of random walks, folding time is defined as mean first passage time to the native state starting from an arbitrary open conformation. Here, "native state" usually means one particular native conformation, but, of course, this is just a simplified coarse grained view of things. In fact, native state is also a macrostate, or, in other words, it corresponds to some region, $\mathcal{N}$, in conformation space. This region most likely has a funny non-trivial shape, because being in the native state imposes very strict requirements (say, below 0.1 Å) on some of the coordinates, while some other might be restricted in a much more liberal way (say, 5 Å for some side groups in the loop regions), or not restricted at all (as, e.g., angles of the side groups in the loops). This implies that $\mathcal{N}$ should be thought of as a pancake-shaped or some fractal with large aspect ratio(s).

Furthermore, in reality there are forces and energies involved, which means random walk in conformation space occurs on top of the energy landscape. Since folding is similar to first order phase transition, there should be a transition state, which is similar to a critical nucleus. In this case, the question is how long does it take for a random walk in conformation space to bring the system into some critical region. Indeed, in a more physical language, when people speak about nucleation and growth mechanism for the kinetics of the first order phase transition¹⁶, it is assumed that critical nucleus plays the role of a "transition state," and that after having achieved this state the system completes the transition very rapidly. This can be re-phrased by saying that *before* crossing the barrier, or before the

critical nucleus is assembled, the conformational motion can be seen as a largely unbiased random walk in the space of conformations, while the way from the transition to the native state is a bias-dominated walk. Thus, the above mentioned critical region is another name for the critical nucleus - it consists of all the microstates which together constitute the transition state ensemble. Thus, our question is how long does it take for the random walk to enter, for the first time, some region starting from a random point in conformation space? To answer this, in addition to the overall conformation space geometry, one has to know more about Ω : what is the shape and fractal dimension of Ω , what is its boundary, etc. We do not know answers to these pressing questions. Only very crudely, we can imagine that Ω , just as $\mathcal{N}$, is in some sense pancake-shaped.

What we do know however is how to formulate the Levinthal estimate of folding time in this language: it is $\tau \sim |\Omega|/|\Omega_{\text{TS}}|$, where $|\Omega|$ is the entire conformation space, and $|\Omega_{\text{TS}}|$ means the number of conformations in the domain Ω_{TS} . This estimate is similar in spirit to the M estimate of exhaustion time, as discussed above, because M , roughly, is $M \sim |\Omega|$. When random walk terminology is considered, it is obvious that this estimate has nothing to do with the first approach time which we are looking for. Indeed, $|\Omega|/|\Omega_{\text{TS}}|$ is in fact an average time between two subsequent visits in Ω_{TS} by a very long trajectory assuming that these visits occur rare enough such that roughly ergodic covering of Ω is established between two visits of Ω_{TS} . Thus, in addition to the question of energetic bias towards the native state, an understanding of random walks in conformation space is crucial to the understanding of protein folding.

From our discussion of the exhaustion time, it also follows that the result for the first passage time will very strongly depend on the shape of critical region Ω_{TS} . Very roughly, if we imagine Ω as a horizontal pancake, then random walk is only important in vertical direction, and unimportant in horizontal direction. This leads to dramatic reduction of the space dimension in which exhaustive search by a random walk should be performed, thus dramatically reducing the time estimate. This fact was recently emphasized by T. McLeash¹⁷.

3.6 What is Common, and What is Different, Between a Chemical Reaction and a First-Order Phase Transition?

A chemical reaction can be represented in terms of proper reaction coordinates as a movement through a saddle point of the (free)energy surface. The reactants and the products correspond to different local minima. As the temperature increases, the reactants will be able to reach the top of the saddle with an increasing probability, such that overall reaction rate is proportional to $(e^{-\Delta G^\ddagger/kT} - e^{-\Delta G/kT})$ (there is also a factor determined by the curvature of the free energy surface in the vicinity of the saddle¹⁸). Temperature dependence of the rate constant is a very steep increasing function, and thus the reaction will proceed with a significant speed only when temperature is not too low.

Somewhat similarly looking but very importantly different probabilities describe the equilibrium properties instead of the rate, this is the probability for the molecule to be in either reactant or a product, $P_{\text{R}} = e^{-G_{\text{R}}/kT}/Z$ and $P_{\text{P}} = e^{-G_{\text{P}}/kT}/Z$. The ratio of the probabilities is proportional to $(e^{-G_{\text{R}}/kT} - e^{-G_{\text{P}}/kT})/e^{-G_{\text{TS}}/kT}$. This implies that the thermodynamic (equilibrium) probability to be in either product or reactant state, whichever is higher in energy, is very low at low temperatures, and grows to become 1/2 at high temperatures. Temperature dependence of this probability is a characteristic sigmoidal shaped curve.

First order phase transition is somewhat similar in the sense that we also have two minima of (free) energy, and they are separated by the barrier. This gives rise to the “switching” behavior somewhat similar to that of the sigmoidal temperature dependence described above. There is, however, an important difference: the width of the switching function, in case of the first order phase transition, goes to 0 as N goes to infinity. In a first-order phase transition the transition part of the function becomes more and more steep as the number of microscopic units (atoms, spins, monomers) grows, and finally becomes a step function. Indeed, as $N \rightarrow \infty$, the free energy difference between the phases (local minima of the landscape) varies as $\Delta F \sim (T_c - T)$. This is because phase transition is a cooperative phenomenon, it is driven by the interactions between particles. Thus, the free energy difference, as a function of the temperature, roughly speaking, is either zero or infinite (in thermodynamic limit), that is the probability to be in a phase switches from zero to one. That means, the sigmoidal temperature dependence is degenerated into a step function. In chemical systems, there is no cooperativity between the microscopic units, and therefore the switching is soft, it is not a step but a sigmoidal curve.

Proteins, and protein folding, belong somehow to both worlds of chemical reactions and of first order transitions. Protein folding is similar to a chemical reaction because folding is the event that involves just one protein molecule, and there is no cooperativity between protein molecules. But, protein folding is also close to a first order transition, because protein molecule contains a large number of monomers, and there is strong cooperativity between monomers in the protein molecule.

Kinetics of protein folding includes, first of all, polymer chain collapse and also self-organization of the “correct” 3D structure. It is known that some proteins first collapse and then search for the native state through the re-arrangements of the collapsed globule, while others wait longer in the open coil state, but then jump into a more or less correct globular state. Initial stages of folding, initial collapse of the coil, seems to be insensitive to the details of the heteropolymer sequence. This argument suggests that these stages should be well modeled by a homopolymer collapse.

3.7 Reaction Coordinate and Levinthal Paradox

Our discussion of the chemical reactions and first order phase transitions allows us to look at the Levinthal paradox from a new refreshing point of view.

As a starting point, it is convenient to imagine a plot showing free energy as a function of an order parameter, or reaction coordinate, x . As we discussed, it must have two minima separated by a maximum. When we are close to the first order transition, the dependence $\Delta F(x)$ evolves with changing temperature such that one minimum gets deeper, while the other one gets more shallow. At the transition point, they switch the roles of the deepest and the less deep ones. This free energy can be defined if one keeps the value of an order parameter, x , fixed, while allowing all other degrees of freedom to relax to thermodynamic equilibrium. For equilibrium statistical mechanics, this consideration is OK, and basically anything can serve as an order parameter, provided that it has different values in two different phases. In kinetics, the situation is far more delicate.

Indeed, let us consider temperature at which both minima of free energy $\Delta F(x)$ are at roughly the same depth. They are separated by a barrier. It looks like the system has to overcome THIS barrier to go from one phase to another. Then, the transition time

should be proportional to $\exp(-\Delta G^\ddagger)$, where $\Delta G^\ddagger(x) - \Delta G^\ddagger(x)$ is the barrier height. However, as free energy is an extensive quantity, the height of the barrier, $\Delta G^\ddagger$, is proportional to V and goes to infinity in thermodynamic limit. In fact, this is just another formulation of Levinthal paradox, because we arrived at the transition time which is exponential in V .

Luckily, there is nothing specific for proteins in this consideration, and we can turn for lessons to other, better understood fields of physics. Consider, for instance, gas-liquid phase transition, perhaps the simplest first order phase transition. For the equilibrium theory of this transition, the most natural order parameter is the overall averaged density of the system. As a function of density, free energy has exactly the shape that we discussed, with two minima separated by a maximum. What is the nature of the maximum? If we start from a gas and increase its density, say, to half that of liquid, then we are at disadvantage from both energy and entropy view points: energetically, our molecules are not close enough to enjoy attractive interactions, so we gained nothing; entropically, density is much higher than in the gas, so we lost a lot. Thus the maximum.

So, how does the gas condenses? This occurs through nucleation. That means, the system does not progress through a succession of overall uniform states with gradually increasing density, instead it goes through the states with dramatically non-uniform density, which is the droplets of liquid in the gas. In this sense, density is not an appropriate order parameter, or not a good reaction coordinate in chemical kinetics language.

What turns out a better choice is the radius of the nucleus. We imagine that a nucleus is a spherical shaped droplet of a new phase, and look at its radius, r . As a function of r , free energy is usually written in the form $\Delta G(r) = (4/3)\pi r^3 \Delta \mu - 4\pi r^2 \sigma$, where the first term is gain in free energy due to the formation of the new phase, its proportional to the volume of the nucleus, while the latter term is the surface energy of the contact between the droplet and the surrounding gas, it is unfavorable (positive), and proportional to the surface area. It is easy to see, that as a function of r this free energy, $\Delta G(r)$, has a characteristic shape with the maximum. It is this maximum, the theory says, that has to be overcome. Obviously, it has nothing to do with V , it does not diverge in thermodynamic limit. It corresponds to the so-called critical nucleus size. In the beginning, the system has to climb on the barrier, it must be a thermo-activated process, entirely due to fluctuations. However, as soon as the nucleus is larger than critical, it starts sliding down the potential hill.

What does this theory assume? Although this is rarely emphasized, this theory assumes that the radius r is a very slow variable. Indeed, this is the only assumption that allows to write down $\Delta G(r)$ as we did before. For instance, by using the surface tension we assume that the surface structure achieves thermodynamic equilibrium at every value of r . By the same token, we assume that the shape of nucleus relaxes much faster than its size, etc. By the way, recent simulations by Daan Frenkel and his co-workers¹⁹ suggest that all these assumptions are not so benign and not universally applicable even for the gas-liquid transition.

Returning to the general consideration, the problem is easy if there is a variable, let say x , which relaxes much slower than all other variables. Then, at the fixed x , the system has time to equilibrate all of its other degrees of freedom, and in this case our considerations become applicable kinetically. That is the problem of the choice of reaction coordinate, or kinetic order parameter. It is a difficult problem, which does not have known general solution.

3.8 Commitment as the Best Possible Reaction Coordinate

Let us assume that we can calculate for each point in the conformation space the probability that the protein is going to fold before ever touching the unfolded state. This way, we can formally define a function κ over the conformation space^{20–22}. This probability is almost unity around the native state, and is almost zero around the unfolded states. This probability is also called commitment, it measures, for every conformation, to which extent this conformation is committed to fold.

Let us first discuss this quantity for a lattice model²⁰. To understand what is κ , we assume that there are two well defined states of the system, in case of lattice toy proteins those are folded and unfolded states. For the present discussion, it is not important that folded state is usually represented by just a single compact conformation, while unfolded state is a big ensemble of coil-like conformations. What is important is that in the conformation space graph certain vortices are labeled as belonging to unfolded state, while certain others are labeled as belonging to the folded state.

Now, we pick an arbitrary conformation, or an arbitrary vertex of the conformation space graph, and define κ for this conformation in the following way. Imagine that we perform many Monte Carlo runs using the chosen conformation as an initial one. Every time, we run Monte Carlo dynamics for as long as it takes for the system to arrive for the first time into either folded or unfolded state. Then, we have to collect the statistics over many runs, and determine the probability, κ , that a run, or randomly chosen trajectory, will first arrive into the folded state before ever touching the unfolded state. This probability is

Quite similarly, commitment can be defined for an off-lattice system^{21,22}. In this case, we choose a particular configuration (or conformation), and then initiate many trajectories going out of this point by randomly choosing the full set of momenta (or velocities). Once again, we determine the probability that the trajectory first hits the folded state before ever touching the unfolded state.

What is the meaning of this quantity, whether one prefers calling it κ or commitment? If κ is large, close to unity, that means the system is very close to the folded state, its descend back to the unfolded state is unlikely. Vice versa, if κ is small, that means the system is close to the unfolded state.

We shall argue that κ , or commitment, is, in principle, THE BEST possible REACTION COORDINATE, or an order parameter. Of course, practical disadvantage of the value κ is that it is difficult to compute and it has not obvious meaning, it is not anything like nucleus radius. But as a matter of principle, there cannot be better reaction coordinate. We shall prove this fact considering kinetics of the lattice model. This is essentially the random walk of a particle on the graph.

Indeed, if there were the slowest degree of freedom in the system, then κ would be clearly a well defined monotonic function of this coordinate. As the system progresses along the selected single degree of freedom from an unfolded to the folded state, its value of κ necessarily increases. Thus, if the reaction coordinate in the traditional sense existed, it would be basically the same as commitment. However, κ can be computed by a constructive algorithm, albeit time consuming one, while reaction coordinate is never known. And, most importantly, κ , or commitment, remains well defined even for the generic system, with no special slow degree of freedom.

The concept of commitment, or τ , is currently widely used in protein folding simulations (see, for instance,²³). Nevertheless, there is a feeling that better intuition about this quantity is necessary. For instance, the works on the statistical mechanics of folding trajectories²⁴ lead to the questions like the ones about local extrema of τ . Is it possible to find a state such that all surrounding states have τ smaller (or larger) than the given one? If the answer were to be positive, that would render the concept of commitment useless at the very least. Luckily, the intuition suggests that it should be impossible. Below, we shall prove that it is impossible indeed.

3.9 Direct Current Analogy

In this section, we shall describe the physical analogy which helps intuitive understanding of the commitment, and also allows to prove their general properties, like monotonous behavior mentioned in the previous paragraph. This analogy is presented in terms of direct currents governed by the Kirchoff rules. To formulate it, let us be a little more specific about the model under study.

We shall consider Metropolis Monte Carlo dynamics - the method first described in a 1953 paper²⁵ by Nicholas Metropolis, Arianna Rosenbluth, Marshall Rosenbluth, Augusta Teller, and Edward Teller. The idea of the method was reportedly conceived by Metropolis, Rosenbluth, and Teller during a Los Alamos dinner party.

Suppose our polymer has certain available conformations, and let us call the conformations with letter $\mathcal{C}$. There is potential energy landscape, which means that there is certain energy $\epsilon(\mathcal{C})$ associated with every conformation $\mathcal{C}$. Consider now a pair of conformations, say $\mathcal{C}$ and $\mathcal{C}'$, such that they are connected on the conformation graph; in other words, they can be transformed one into another via a single Monte Carlo move. Consider the flow between $\mathcal{C}$ and $\mathcal{C}'$. Of course, this means, we should imagine a large number (an ensemble) of computers independently running Monte Carlo simulation of our system, and then we should ask what is number of copies switching from $\mathcal{C}$ to $\mathcal{C}'$ at the given time? Equivalently, we can speak of a probability to find a single Monte Carlo process jumping from $\mathcal{C}$ to $\mathcal{C}'$ at the given time. To be specific, let us assume that $\mathcal{C}$ is higher in energy than $\mathcal{C}'$: $\epsilon(\mathcal{C}) > \epsilon(\mathcal{C}')$. Then, Metropolis criteria²⁵ accepts the move from $\mathcal{C}$ to $\mathcal{C}'$ every time the move is offered, we get that the flow from $\mathcal{C}$ to $\mathcal{C}'$ is simply proportional to the occupation number of $\mathcal{C}$ at the present time, $n_{\mathcal{C}}$. As regards the opposite moves from $\mathcal{C}'$ to $\mathcal{C}$, they are accepted with probability smaller than unity, namely $\frac{\epsilon(\mathcal{C}) - \epsilon(\mathcal{C}')}{\epsilon(\mathcal{C})}$. We should also keep in mind that the connectivity of the conformation graph in different points might be different. Specifically, let us say that every conformation $\mathcal{C}$ has some $\nu_{\mathcal{C}}$ neighbors on the graph. Then, the bond leading from $\mathcal{C}$ to $\mathcal{C}'$ is offered with the probability $1/\nu_{\mathcal{C}}$. Therefore, we can summarize all this by writing the following relation for the flux of probability from $\mathcal{C}$ to $\mathcal{C}'$:

$$J_{\mathcal{C} \rightarrow \mathcal{C}'} = \frac{c}{c'} \left\{ 1 - \frac{c}{c'} \left[\frac{\epsilon(\mathcal{C}) - \epsilon(\mathcal{C}')}{\epsilon(\mathcal{C})} \right] \right\} \quad (1)$$

Similarly, the flux in the opposite direction is

$$J_{\mathcal{C}' \rightarrow \mathcal{C}} = \frac{c'}{c} \left\{ 1 - \frac{c'}{c} \left[\frac{\epsilon(\mathcal{C}') - \epsilon(\mathcal{C})}{\epsilon(\mathcal{C}')} \right] \right\} \quad (2)$$

Below, to save some writing, we shall omit the temperature factor, assuming $k_B T = 1$; in other words, we pretend that potential energy $\mathcal{E}(\mathcal{C})$ is measured in the units of $k_B T$: $\mathcal{E}(\mathcal{C}) \rightarrow \mathcal{E}(\mathcal{C})$.

It is convenient to rewrite equations (1) and (2) by introducing the following notations. For every conformation $\mathcal{C}$, we define the quantity

$$c_{\mathcal{C}} = \exp(-\mathcal{E}(\mathcal{C})) \quad (3)$$

and for every connection between $\mathcal{C}$ and $\mathcal{C}'$, we define another quantity

$$cc' = \left\{ c_{\mathcal{C}}^{-1} \mathcal{E}(\mathcal{C}) - c_{\mathcal{C}'}^{-1} \mathcal{E}(\mathcal{C}') \right\} \quad (4)$$

One nice thing about quantities $c_{\mathcal{C}}$ is that in equilibrium, when occupation numbers $c_{\mathcal{C}}$ are governed by Boltzmann distribution $c_{\mathcal{C}} \propto \exp(-\mathcal{E}(\mathcal{C}))$, the values of $c_{\mathcal{C}}$ become independent of $\mathcal{C}$, a constant all over the conformation space. Using $c_{\mathcal{C}}$ and cc' , the master equations (1) and (2) can be presented in the form

$$cc' \equiv c_{\mathcal{C} \rightarrow \mathcal{C}'} - c_{\mathcal{C}' \rightarrow \mathcal{C}} = \frac{c_{\mathcal{C}} - c_{\mathcal{C}'}}{cc'} \quad (5)$$

In this formula, it is now easy to recognize the Ohm's law for direct currents, which leads to the following physical interpretation.

We imagine that the conformational graph of the system is a network of resistors. The resistance between nodes $\mathcal{C}$ and $\mathcal{C}'$ is cc' . Governed by the Kirchoff rules, there are potentials on each node, $c_{\mathcal{C}}$, and currents between the nodes, cc' . The knowledge of these potentials and the current distribution is equivalent to the full knowledge of the dynamics of the system in question. Thus, we see that occupation numbers (or probabilities) $c_{\mathcal{C}}$ satisfy the equation which is an analog of diffusion equation:

$$-\frac{c_{\mathcal{C}}}{c} = \sum_{\mathcal{C}'} cc' \quad (6)$$

where the summation runs over all sites $\mathcal{C}'$ which are neighbors of the site $\mathcal{C}$ on the graph.

In fact, as we know from the experience with the resistor networks, the important ingredient of the problem is the externally applied voltage. In the presence of such applied voltage, direct currents may flow and remain time-independent. The easiest way here is to look at the analog of Green's function of the diffusion equation, which is equivalent to applying a δ -type voltage source at just one point, say, $\mathcal{C}$:

$$-\frac{c_{\mathcal{C}}}{c} = \sum_{\mathcal{C}'} cc' - \delta(\mathcal{C}) \quad (7)$$

It is useful now to realize the more specific meaning of occupation numbers $c_{\mathcal{C}}$ in our specific context. One way to understand them is to imagine that there are many processors simultaneously running the Monte Carlo simulation of a given protein. Then, $c_{\mathcal{C}}$ is the fraction of all processors whose simulations are at the point $\mathcal{C}$ at the time moment t . From that point of view, we should imagine the rule best adjusted to computing the τ -commitment: every time that a particular computer achieves folding, it is re-launched again starting from the point $\mathcal{C}$. In this setting, we will eventually achieve a stationary distribution with time-independent $c_{\mathcal{C}}$, but with non-zero and non-trivial currents. The equation

for such stationary distribution is derived by time integration of the diffusion equation (7):

$$0 = - \sum_{c'} \sum_c c c' \quad (\mathcal{C}) \quad (8)$$

or, in the other words,

$$\sum_{c'} \sum_c \frac{c}{\left\{ \frac{c}{c} \frac{(c)}{(c)} - \frac{c'}{c'} \frac{(c')}{(c')} \right\}} \quad (\mathcal{C}) \quad (9)$$

3.10 Direct Current Formulation of the First Return Problem and its Easy Solution

The concept of commitment, as it is formulated above, is based on the classical probabilistic “first return” problem. Indeed, commitment of the state $\mathcal{C}$ is the probability that the trajectory of the random walk on the graph, starting from $\mathcal{C}$ will arrive *for the first time* into the folded state $\mathcal{F}$ before ever hitting the unfolded state $\mathcal{U}$. Let us consider the simple classical example²⁶.

Suppose a gambler arrives at the casino with certain amount of money x and keeps gambling indefinitely unless arriving at the desperate zero money state ($x = 0$), in which case he/she is discarded. The question is this: what is the probability that the player loses money (arrives for the first time at $x = 0$) at the time t ? The standard way to solve this problem is to realize that the time-dependent probability distribution of the money at possession of the gambler satisfies the diffusion equation:

$$\frac{\partial P(x, t)}{\partial t} = D \frac{\partial^2 P(x, t)}{\partial x^2} \quad (x > 0) \quad (10)$$

where D is an appropriate diffusion coefficient (which can be absorbed into the time measurement units), and $P(x, 0)$ -functions describe the initial condition. The central idea of this approach is to realize that the concept of *first return* is embodied in the absorbing boundary condition:

$$P(0, t) = 0 \quad (11)$$

This boundary condition ensures that in the path integral solution of the diffusion equation, all trajectories are discarded which ever visit the left half-line ($x < 0$). Now, assuming $P(x, t)$ is found, the probability to arrive at $x = 0$ at time t *for the first time* is given by the flux into $x = 0$: $J(0, t) = -D \frac{\partial P(x, t)}{\partial x} \bigg|_{x=0}$. The corresponding solution is not difficult to find:

$$P(x, t) = \frac{x}{2\sqrt{Dt}} \exp\left(-\frac{x^2}{4Dt}\right) \quad (12)$$

and the classical theorem²⁶ says that $\int_0^\infty J(0, t) dt = 1$: sooner or later, gambler loses all money with probability one.

Here is now the simple solution of this problem based on the direct current analogy (see also²⁷). Consider semi-infinite line of resistors (or just a wire of uniform resistance) going along the x -axis. In terms of analogy, the absorbing boundary condition (11) means that the boundary is grounded, i.e., is kept at the zero potential (3). Suppose now we feed a direct current into the point x (of course, applying a constant voltage to the point x). Then, the whole statement of the theorem is that the current leaving the system through the

grounded terminal at $x = 0$ is equal to the current entering the system through the terminal at $x = L$ - which is trivial.

Of course, this simple solution comes from the use of stationary equation (9).

3.11 Direct Current Formulation of the Commitment

Now we are prepared to formulate the way to determine the commitment of any given state $\mathcal{C}$ using the direct current analogy. We understand that the first arrival condition can be imposed by absorbing boundary conditions, or grounding the corresponding sites. Therefore, the formulation is as follows.

Suppose all the sites $\mathcal{U}$ corresponding to the unfolded state are grounded, as well as the site $\mathcal{F}$ corresponding to the folded state. Suppose further that we feed a direct (stationary, time independent) current I into the site $\mathcal{C}$. This current flows partly to $\mathcal{U}$ and partly to $\mathcal{F}$; in obvious notations, $I_{\mathcal{U}} + I_{\mathcal{F}} = I$. Then, the commitment is nothing but the fraction of current going into the $\mathcal{F}$ terminal. According to the Ohm's law, this can also be written in terms of the corresponding resistances:

$$(\mathcal{C}) = \frac{I_{\mathcal{F}}}{I} = \frac{R_{\mathcal{CU}}}{R_{\mathcal{CU}} + R_{\mathcal{CF}}} \quad (13)$$

It is instructive to consider here a simple example, which is the diffusion in one-dimensional potential landscape $V(x)$. Of course, in this case the choice of reaction coordinate is trivial, there nothing but x . Nevertheless, we can imagine that there are two ends of the diffusion interval, which we call points $\mathcal{U}$ and $\mathcal{F}$, and we want to know what is the probability to start from some point $\mathcal{C}$ between $\mathcal{U}$ and $\mathcal{F}$ and arrive at $\mathcal{F}$ for the first time without ever touching $\mathcal{U}$. In this case, the resistor network is one-dimensional, all resistors are connected in series. Remembering the expression (4), we see momentarily that formula (13) yields

$$(\mathcal{C}) = \frac{\int_{\mathcal{C}}^{\mathcal{U}} V'(x) dx}{\int_{\mathcal{F}}^{\mathcal{U}} V'(x) dx} \quad (14)$$

It is clearly seen both in the example (14) and in the general formula (13) that the commitment is always positive (non-negative) and never exceeds unity - as it must be for the probability.

3.12 Direct Current Formulation of the Landscape

Our result (13) generates an insight into the general meaning of commitment and its properties. For instance, we can proceed in the following way. Let us ask what are the surfaces of $V(x)$, and what is their topology.

To understand it, let us imagine yet another experiment with the direct current. Let us ground this time only the folded state site $\mathcal{F}$, and let us feed the current into the unfolded state sites $\mathcal{U}$. In other words, the potentials of the terminals at the folded and unfolded states are equal to $\mathcal{F} = 0$ and $\mathcal{U} = V$, respectively. In this case, according to the Ohm's law, the potential at the arbitrary site $\mathcal{C}$ should be equal to

$$\frac{V_{\mathcal{C}}}{V} = \frac{R_{\mathcal{CF}}}{R_{\mathcal{CU}} + R_{\mathcal{CF}}} \quad 1 - \quad (15)$$

Thus, this potential essentially gives us the value of the commitment Φ_c . Of course, this realization of the direct current model corresponds very directly to the most common computer experiment in which we start from somewhere in the unfolded region and look for the first arrival into the folding state.

The result (15) indicates that the topology of every surface of constant commitment is such that it separates folded state from unfolded state; moreover, and more general, every surface Φ_c separates the regions with $\Phi < \Phi_c$ and with $\Phi > \Phi_c$. In particular, the commitment has no local maximum or local minimum. Its only minimum is the global one at the unfolded state, where it is equal to zero, and its only maximum is that at the folded state where it is equal to unity. This confirms the intuition according to which the commitment can be used as a reaction coordinate.

Now, the central question arises: what is the relevant free energy profile if we choose to use Φ_c as the reaction coordinate? It turns out possible to address this question in quite general form. In this argument, we shall disregard the discrete character of conformation space graph and pretend to work with the continuous conformations space. Moreover, we shall assume that the conformation space is flat, which means that Φ_c does not depend on $\mathcal{C}$.

Consider our latest direct current formulation, in which terminals are at the folded and the unfolded states, and equipotential surfaces are simultaneously the surfaces of constant commitment. Since the current does not flow along the equipotential surface, we can ask what is the resistance of the conformation space layer sitting between the surfaces Φ_c and $\Phi_{c'}$. Since these surfaces are equipotential, all resistors connecting them may be viewed as connected in parallel. For them, the conductivities must be summed together, which yields

$$\frac{1}{R(\Phi_c, \Phi_{c'})} = \sum_c \frac{1}{R_{cc'}} \quad (16)$$

Remembering the formula (4), it is natural to define $R(\Phi_c, \Phi_{c'})$ according to $R_{cc'} = R_{eff}(\Phi_c, \Phi_{c'})$, and then we obtain

$$R(\Phi_c, \Phi_{c'}) = \left[\sum_c \frac{1}{R_{cc'}} \right]^{-1} \quad (17)$$

Strikingly, this formula coincides with the definition of the partially equilibrium free energy as a function of order parameter, except it does not involve any assumptions regarding slow degree of freedom and the like. Instead of being the definition of the statistical mechanics partition function, formula (17) describes the parallel connection of the resistors. What this formula proves is that one can use the commitment as the reaction coordinate, with the relevant free energy profile given by the formula (17). In particular, the concept of the folding barrier is clearly formulated as the place with highest resistance.

It should be born in mind that equation (17) deserves further attention. In particular, it remains unclear how we should proceed with generalizing it for the case when Φ_c does depend on $\mathcal{C}$. This raises the question of the geometry of the real conformation space.

4 Compact Conformations: Lattice Model

The study of the geometry of compact polymers is quite tricky. The theory of random compact conformations is well developed on the mean field level (see, e.g., in the book²⁸). This is the theory of homopolymer globules, because they are entropically dominated by the most typical conformations. Major conclusion of the mean field theory is that chain segments inside the globule follow Gaussian statistics, and do not exhibit any signs of order. This conclusion is in sharp contradiction with the statements in the literature²⁹⁻³¹ that compactness of the conformation may favor elements of secondary structures, such as α -helices and β -pins.

Computationally, the problem of compact conformations is closely related to that of Hamiltonian walks on the graphs. We remind the reader that the concept of a Hamiltonian walk was introduced by Hamilton in connection with famous Euler problem of Königsberg bridges: the task was to find the Sunday promenade passing every one of the seven bridges, never returning to the already visited place. In general, Hamiltonian walk on an arbitrary graph can be defined as a walk which visits every site on the graph once and only once. If our graph is, say, $3 \times 3 \times 3$ piece of the cubic lattice in $3D$, then Hamiltonian walk on such graph is the same as maximally compact conformation of the polymer filling $3 \times 3 \times 3$ domain.

Enumeration of Hamiltonian walks on graphs is well known problem in combinatorics. Of course, the best possible statistics is achieved by exhaustive enumeration of all Hamiltonian walks. This is possible for rather short polymer chains only: for the chains with 27 monomers filling $3 \times 3 \times 3$ of the cubic lattice³², and also for 36- and 48-mers, filling $3 \times 3 \times 4$ and $3 \times 4 \times 4$ segments, respectively³³. Obviously, these chains are far too short to address statistics and fractal structure of the typical conformation.

Short of exhaustive enumeration, other methods to generate larger compact conformations have been suggested. The most straightforward Monte Carlo chain growth methods³⁴ are totally inefficient for long compact chains, because of catastrophic explosion of rejected looped conformations. Transfer matrix approach put forward by³⁵⁻³⁷ is very efficient for the chains filling an elongated domain $L \times M \times N$, where one of the dimensions, say L , may be arbitrarily large. Unfortunately, to remain within computational tractability, two other dimensions, M and N , must be small, not greater than 2 or 3. An alternative approach, suggested in³⁸, is free of this limitation. It employs combinatorial techniques of two-matching and patching of bipartite graphs. Unfortunately, we found that this method generates conformations in a heavily biased way.

In the forthcoming paper³⁹, we demonstrate significant improvements to the algorithm of Ramakrishnan *et al*³⁸. Compact conformations (Hamiltonian walks) and cycles were generated up to the size $22 \times 22 \times 22$. Their fractal and topological properties were examined in details. The major findings were as follows.

First, it was confirmed that local fractal structure of typical compact conformations is Gaussian at the scale smaller than the globule size. In other words, it satisfies Flory theorem. Second, it was found that for linear compact chain the correlation between chain ends is minimal, it is basically reduced to the excluded volume effect (both ends cannot occupy the same place). Third, it was found that chain compaction quite significantly increases the probability of forming non-trivial knots. These results should be born in mind when we consider conformations of real proteins and try to analyze whether they are random or not.

4.1 Native Conformations

So far we thought about a native state of a protein in quite abstract way. In fact, several thousands real native conformations are now known due to the X-ray diffraction analysis. Surprisingly, many of the native conformations are quite alike despite they belong to totally different proteins. This might be the result of either convergent or divergent evolution. In the former case, we imagine that certain conformations have some evolutionary advantages, and then evolutionary development brings proteins to have these advantageous conformations. Another possibility is to imagine that similarities between conformations of proteins might be due to the common ancestry of these proteins. Not going too deeply into this exciting subject, we only make a reference to the interesting recent paper⁴⁰ and the references therein.

5 Conclusion

It should be noted that the understanding of Levinthal paradox has progressed very significantly since it was first formulated¹⁴. First of all, it is found that the folding time, under the conditions of thermodynamic equilibrium between folded and unfolded states, scales as $(\tau_0)^{41,42}$, which is very significantly smaller than Levinthal time proportional to $(\tau_0)^N$. This estimate, as already said, is valid under the conditions of thermodynamic equilibrium, which means that it relies on the transition between denatured and native forms being highly cooperative, of all-or-none type. Indeed, high cooperativity is a well established experimental fact². It is also well understood that high cooperativity is the property of proteins which is due to their peculiar selected sequences. Among random sequences, vast majority would not have exhibited any signs of cooperativity, as it was first established by Shakhnovich and Gutin⁴³. This latter fact has been extensively tested using lattice models (as described, e.g., in the review article⁴⁴; see also references therein). Since everything related to the lattice models is perceived with a large dose of (healthy?) skepticism in protein community, it is important to emphasize that the fact of non-cooperative folding in the majority of sequences is well understood beyond lattice models. Actually, it was foreseen by Bryngelson and Wolynes a long time ago⁴⁵.

Speaking about the relation between sequence selection, the all-or-none cooperative mechanism of folding, as the scaling of folding time, it is interesting to mention that experimental observations do not provide any evidence on the folding (under equilibrium conditions) time dependence on the chain length. That means, the above mentioned theoretical prediction, $(\tau_0)^{41,42}$, although sufficient to rule out any paradoxes, maybe still an overestimate.

The role of sequence selection is also well understood from a different view point, namely, related to the mutation stability (see also in the review article⁴⁴). In the majority of sequences, every mutation breaks the stability of the native state with the probability very close to 100%. By contrast, the selected sequences - the same ones which exhibit highly cooperative folding-unfolding transition! - are reliable in the sense that their native state with high probability survives and remains stable even after several mutations.

While the real mechanisms of evolutionary sequence selection remain unknown, and while the computational models of sequence selection keep improving since the first suggestions^{46,47}, it is getting increasingly clear that there are many sequences which meet the sufficient criteria of reliable folding.

To summarize, in the light of all the findings of the last decade, it seems clear (to the present author, at least) that the discussions about Levinthal paradox must now move forward to the new spheres. How does the sequence selection work (or worked) in real evolution? What are the specific scenario of folding dynamics for selected sequences - how specific is the nucleation, how many and which conformations belong to the transition state, what is the reaction coordinate associated with folding; in other words, how precisely do these selected sequences slide down their folding funnels^{48,23,49,50}? What are the physical principles behind the selection of certain spatial structures, or folds and fold families⁵¹? What are the general physical principles behind the enzymatic, motor and other functions of proteins, and do they have any relation to the principles involved in folding? What are the mechanisms of aggregation, or mechanisms preventing aggregation, of proteins? There are very many works on these subjects, to make the list of them is a daunting task far beyond the framework of the present note. However many questions remain open, it seems that the Levinthal's question - how can protein sample "biologically large" number of conformations - has been answered: protein does not sample them. Most importantly, there are sufficiently many "good" sequences for the evolution to select from, where every "good" sequence is capable of folding, and does not need exhaustive conformation sampling to do so. Understanding this was a remarkable achievement of the last decade.

References

1. A. V. Finkelstein, O. B. Ptitsyn *Protein Physics*, Academic Press, 2002.
2. P.Privalov Adv. Protein. Chem. **33**, 167-241 (1997).
3. C.Anfinsen, R.Redfield, W.Choate, J.Page, W.Carroll *Studies on the Gross Structure, Cross-Linkages, and Terminal Sequences in Ribonuclease* Journal of Biological Chemistry **207**, 201-210 (1954).
4. Christian B. Anfinsen *Studies on the Principles That Govern the Folding of Protein Chains, The Nobel Lecture delivered on 11 December 1972*. Les Prix Nobel en 1972, p. 103-119, Nobel Foundation, 1973.
5. H. Fraunfelder, Colloquium talk at the University of Minnesota Physics Department, October, 2002.
6. O. B.Ptitsyn, M. V.Volkenstein Journal of Biomolecular Structure and Dynamics **4**, 137 (1986).
7. P. H. Verdier, W. H. Stockmayer Journal Chem. Phys. **36**, 227 (1962).
8. J. D. Bryngelson, J. N. Onuchic, N. D. Socci, P. G. Wolynes: Proteins **21**, 167-195 (1995).
9. M. R. Betancourt, J. N. Onuchic Journal Chem. Phys. **103**, 773-787 (1995).
10. R. Du, A. Grosberg, T. Tanaka *Random Walks in the Space of Conformations of Toy Proteins* Phys. Rev. Letters **84**, 1828-1831 (2000).
11. A. Scala, L. A. Nunes Amaral, M Barthèlèmy Europhys. Lett. **55**, 594-600 (2001).
12. R. Du, A. Grosberg, T. Tanaka, M. Rubinstein *Unexpected Scenario of Glass Transition in Polymer Globules: an Exactly Enumerable Model* Phys. Rev. Letters **84**, 2417-2420 (2000).
13. E. Shnol, preprint, Pouschino (1980).
14. C.Levinthal Journal Chim. Phys. Biol. **65**, 44-45 (1968).
15. I.Berezovsky, E.Trifonov *Loop fold structure of proteins: resolution of Levinthal's*

- paradox* J. Biomol. Struct. Dyn. **20**, 5-6 (2002); A.Finkelstein *Cunning simplicity of hierarchial folding* Journal Biomol. Struct. Dyn. **20**, 311-313 (2002); I.Berezovsky, E.Trifonov *Back to units of protein folding* Journal Biomol. Struct. Dyn. **20**, 315-316 (2002); A.Grosberg *A few disconnected notes related to Levinthal paradox* Journal Biomol. Struct. Dyn. **20**, 317-321 (2002); A.Kloczkowski, R.Jernigan *Loop folds and evolutionary conservation of folding nuclei* Journal Biomol. Struct. Dyn. **20**, 323-325 (2002); M.Rooman, Y.Dehouck, J.Kwasigroch, C.Biot *What is paradoxical about Levinthal paradox?* Journal Biomol. Struct. Dyn. **20**, 327-329 (2002); A.Fernandez, A.Belinky, M. de las Mercedes Boland *Protein folding: where is the paradox?* Journal Biomol. Struct. Dyn. **20**, 331 (2002).
16. E. Lifshitz, L. Pitaevskii: *Physical Kinetics*, reprint edition (Butterworth-Heinemann, 1981).
 17. T. C. B.McLeash *Protein folding in high-dimensional spaces: hypergutters and the role of non-native interactions* <http://arxiv.org/abs/physics/0311028>.
 18. P. Hanggi, P. Talkner, M. Borkovec *Reaction-rate theory: fifty years after Kramers* Reviews of Modern Physics **62**, 251 (1990).
 19. S. Auer, D. Frenkel *Prediction of absolute crystal-nucleation rate in hard sphere colloids* Nature **409**, 1020-1023 (2001).
 20. R. Du, V. Pande, A. Grosberg, T. Tanaka, E. Shakhnovich *On the Transition Coordinate for Protein Folding* Journal Chem. Phys. **108**, 334-350 (1998).
 21. P. Bolhuis, C. Dellago, P. Geissler, D. Chandler: *Transition path sampling: throwing ropes over mountains in the dark* Journal Phys. - Condensed Matter **12**, A147-A152 (2000).
 22. P. G. Bolhuis, D. Chandler, C. Dellago, and P. L. Geissler, *Transition path sampling: Throwing ropes over rough mountain passes, in the dark* Annu. Rev. Phys. Chem. **53**, 291-318 (2002).
 23. V. Pande, A. Grosberg, T. Tanaka, D. Rokhsar *Pathways for Protein Folding: is a New View Needed?* Current Opinion in Structural Biology **8**, 68-79 (1998).
 24. P. Eastman, N. Grönbech-Jensen, S. Doniach *Simulation of protein folding by reaction path annealing* Journal Chem. Phys. **114**, 3823 (2001).
 25. N. Metropolis, A. Rosenbluth, M. Rosenbluth, A. Teller, and E. Teller *Equation of state calculations by fast computing machines* Journal Phys. Chem. **21**, 1087-1092 (1953).
 26. W. Feller: *An introduction to probability theory and its applications*, 2nd edn. (J. Wiley & Sons, New York 1971).
 27. S. Redner: *A guide to first passage processes* (Cambridge Univ. Press, 2001).
 28. A. Yu. Grosberg, A. R. Khokhlov, *Statistical Physics of Macromolecules*, AIP Press, New York, (1994).
 29. D.Yee, H.Chan, T.Havel, K.Dill Journal Mol. Biol. **241**, 557-573 (1994).
 30. A. Maritan, C. Micheletti, A. Trovato, J. R. Banavar, Nature **406(6793)**, 287-290 (2000).
 31. J. Banavar, A. Maritan *Colloquium: Geometrical approach to protein folding: a tube picture* Reviews of Modern Physics **75**, 23-34 (2003).
 32. E. Shakhnovich, A. Gutin Journal Chem. Phys. **93**, 5967 (1990).
 33. V. S. Pande, C. Joerg, A. Yu. Grosberg, T. Tanaka, Journal Phys. A: Math. Gen. **27(18)**, 6231-6236 (1994).

34. *Markov Chains and Monte Carlo Calculations in Polymer Science, Monographs in Macromolecular Chemistry*, edited by G. G. Lowry, Marcel-Dekker, New York, (1970).
35. A. Kloczkowski, R. L. Jernigan *Journal Chem. Phys.* **109**(12), 5134-5146 (1998).
36. A. Kloczkowski, R. L. Jernigan *Journal Chem. Phys.* **109**(12), 5147-5159 (1998).
37. A. Kloczkowski, R. L. Jernigan *Macromolecules* **30**(21), 6691-6694 (1997).
38. R. Ramakrishnan, J. F. Pekny, J. M. Caruthers *Journal Chem. Phys.* **103**(17), 7592 (1995).
39. R. Lua, A. Borovinskiy, A. Grosberg *Fractal and statistical properties of large compact polymers: a computational study* *Polymer* **45**(2), 717-731 (2004).
40. N. V. Dokholyan, B. E. Shakhnovich, E. I. Shakhnovich *Expanding protein universe and its origin from the biological Big Bang* *Proc. Nat. Acad. Sci. USA* **99**, 14132-14136 (2002).
41. A. Finkelstein, A. Badretdinov *Folding & Design* **2**, 115 (1997); *Folding & Design* **3**, 67 (1998).
42. R. Du, A. Grosberg, T. Tanaka *Coexistence of Native and Denatured Phases in a Single Protein-like Molecule* *Phys. Rev. Letters* **83**, 4670-4673 (1999).
43. E. Shakhnovich, A. Gutin *Biophys. Chem.* **34**, 187 (1989).
44. V. Pande, A. Grosberg, T. Tanaka *Heteropolymer Freezing and Design: Towards Physical Models of Protein Folding* *Reviews of Modern Physics* **72**, 259-314 (2000).
45. J. Bryngelson, P. Wolynes *Proc. Nat. Acad. Sci. USA* **87**, 7524-7528 (1987).
46. E. Shakhnovich, A. Gutin *Proc. Nat. Acad. Sci. USA* **90**, 7195 (1993).
47. V. Pande, A. Grosberg, T. Tanaka *Proc. Nat. Acad. Sci. USA* **91**, 12972 (1994).
48. E. Shakhnovich *Modeling protein folding: the beauty and power of simplicity* *Folding and Design* **1**, 50-52 (1996); E. Shakhnovich *Theoretical studies of protein-folding thermodynamics and kinetics* *Current Opinion Structural Biology* **7**, 29-40 (1997).
49. J. Onuchic, N. Socci, Z. Luthey-Schulten, P. Wolynes *Protein folding funnels: the nature of the transition state ensemble* *Folding and Design* **1**, 441-450 (1996).
50. K. Dill, H. Chan, *From Levinthal to pathways to funnels*. *Nature: Structural Biology* **4**, 10-19 (1997).
51. H. Li, R. Helling, C. Tang, N. Wingreen *Science* **273**, 666 (1996).

Proteins as Molecular Machines: Force Probe Simulations

Helmut Grubmüller

Theoretical and Computational Biophysics Department,
Max-Planck-Institute for Biophysical Chemistry,
Am Fassberg 11, 37077 Göttingen, Germany
E-mail: hgrubmu@gwdg.de

Many proteins are molecular “nano-machines”, which perform their biological function via well-coordinated structural transitions. In most cases, these transitions occur on much slower time scales than those accessible to conventional molecular dynamics techniques. Force probe simulations provide a powerful means to overcome this limitation and thus to get insight into the atomistic mechanisms that underly biological functions. This chapter provides a basic introduction into this method. It further sketches a simple non-equilibrium statistical mechanics treatment that shows how to connect the results of force probe simulations with atomic force microscopy experiments. Finally, the method is applied to study mechanical energy transfer steps in F_1 ATP synthase, probably the smallest rotary motor on earth.

1 Introduction

Proteins are macromolecules, consisting of many thousands of atoms. Virtually all biomolecular tasks in the cell are carried out by proteins. Examples are the maintenance of structures (collagen), enzymatic catalysis, chemical signal transduction (immunoglobulins and hormones), DNA transcription and duplication, chemico-mechanical energy conversion (motor proteins) transport across membranes, and the degradation of dispensable or degenerated proteins. Other functions are the transport or filtering of small molecules (ion pumps, aquaporins), molecular recognition (antibodies), photosynthesis, sensing (e.g. rhodopsin), or the triggering of membrane fusion¹. Traditionally, changes in the Gibbs free energy are the central concept to follow and describe these processes. More recently, with the advent of single molecule force probe experiments and the perception of proteins as “molecular machines”, the application and detection of targeted forces open a new route — both experimentally and on the theory side — to protein function². The theoretical aspects and appropriate simulation techniques are the subject of this chapter.

2 Protein Structure and Dynamics

Contrary to most other polymers, proteins adopt under native conditions (typically aqueous solvent at 20 – 40°C and neutral *pH*) a well defined spatial structure³, i.e., the (average) position of each single atom of the protein is well-defined. This remarkable fact allows to measure all the atomic positions by X-ray crystallography or nuclear magnetic resonance (NMR) spectroscopy. The Protein Data Bank⁴ currently (2004) archives more than 20 000 such atomically resolved protein structures.

Very early experiments have suggested periodic structural elements in proteins. W.T. Astbury (a scholar of W.L. Bragg), e.g., observed characteristic patterns in X-ray scattering experiments on hairs already in the thirties. Remarkably, these patterns changed upon

stretching of the hair, and Astbury referred to the two respective structures as “ α ” and “ β ”. This was the first direct observation of a structural transition in a protein. The similarly regular helical structure of DNA discovered by Watson and Crick twenty years later led to the general belief that all biological macromolecules exhibit a simple and regular structure.

When John Kendrew determined the structure of myoglobin⁵ — a small muscular oxygen transport protein — in 1958, its extraordinary complexity and lack of regularity came, therefore, to many as a surprise. Kendrew remarked: “Perhaps the most remarkable features of the molecule are its complexity and lack of symmetry. The arrangement seems to be almost totally lacking in the kind of regularities which one instinctively anticipates.”

From a physical/chemical point of view, the irregular — but highly ordered — protein structure is in fact not easy to explain. In solution, essentially all other polymers adopt a huge number of different, usually quickly interchanging structures (“random coil”). Even polypeptides with random amino acid sequence (as opposed to the specific sequences of proteins selected by evolution) typically do not fold into well defined structures. Thus, the question of which physical principles govern the self-organized folding of proteins into their native, three-dimensional structure, is still unresolved. This is demonstrated from the very fact that, still, it is not possible to predict this structure solely from the (genetically determined and unique) amino acid sequence (“primary structure”) — although the protein structure is, under native conditions, fully determined by the sequence. This long standing problem is called “folding problem”, and by many considered the “holy grail” of theoretical molecular biophysics.

From an evolutionary point of view, however, it would be very surprising *not* to find well-defined protein structures: As soon as one accepts that the quite complex functions carried out by proteins (such as enzymatic catalysis, signal transduction, motor function, ion pumping and filtering and many more) require “gadgets” with at least similarly complex blueprints — i.e., if the analogy of a “nano-machine” holds — then one must indeed expect proteins to exhibit a high degree of structural and dynamic organization, built up from components the properties and motions of which are tailored towards the desired function.

Thus, a crucial selection criterium for proteins is apparently to fold (under physiological conditions) reliably, reproducibly, and in a self-organized manner into unique spatial structures. (As is often the case in biology, there are exceptions. Some proteins seem to perform their task without any well-defined spatial structure). From this picture it follows that already quite minute perturbations of protein structures, of the physico/chemical properties of single amino acids within a protein, or their intramolecular interactions, can compromise or even destroy the proper functioning of a particular protein. The large number of hereditary diseases and specific forms of cancer, which originate from modifications of single base pairs in the genes (and, therefore, from the exchange of a single amino acid within the encoded protein), demonstrates that this is in fact the case. The picture of a carefully aligned “molecular machine” is further supported by the observation that already minute changes in temperature, *pH*, or salt concentration, can destroy protein structure and function.

These examples also illustrate the essential role of collective structural transitions for protein function, so-called conformational transitions⁶. These differ from the thermal high-frequency vibrations, which fall into the picosecond range, in that they occur on much slower time scales, from nanoseconds to seconds. The opening and closing of ion channel

proteins in nerve cells is an example for such a functional conformational transition. It is characterized by millisecond rates and can be observed via an electric current through the channel by patch-clamp measurements⁷⁻⁹ on single proteins. Other examples are so-called allosteric effects (e.g., in hemoglobin), which are realized by conformational transitions, and induced fit motions upon ligand binding, e.g., of an antigen to an antibody. Last but not least we mention the elementary steps of motor proteins for muscle contraction, which also represent conformational transitions.

Functional processes in proteins occur on a large variety of different time scales, and so does protein dynamics and conformational dynamics, which covers a wide hierarchy of dynamical processes: the high frequency part of the thermal fluctuations of single atoms around their average position is characterized by reciprocal frequencies of several ten to several hundred femtoseconds; collective motions of smaller groups of atoms range up to several ten picoseconds. Larger structural changes occur on a hierarchy of time scales ranging from nanoseconds to hours. There is a convincing amount of experimental and theoretical evidence for “hierarchical substates”⁶ for protein conformations, which are linked by conformational transitions on a wide range of time scales¹⁰. These conformational dynamics apparently is a feature specific to proteins and other systems of comparable complexity¹¹.

The highly ordered, but heterogeneous structure of proteins in combination with the broad range of non-separating time scales in protein dynamics renders the application of established concepts of many body physics extremely difficult; work in this direction, therefore, is a challenging field in theoretical physics. Most current approaches, however, aim at describing and predicting quite general properties of protein dynamics of protein folding, and functional processes of specific proteins are typically not accessible.

Accordingly, the explicit, atomically resolved molecular dynamics simulation of the protein of interest is still to be considered the only reliable theoretical description. The lack of more elegant treatments has a high price: such protein dynamics simulations are computationally extremely demanding and, therefore, with presently available hardware (e.g., Linux clusters), restricted to relatively small simulations systems (several 100,000 atoms) and, most severely, to short time scales of several tens of nanoseconds. Nevertheless, the method is now widely and successfully employed for a large number of biological processes that fall into this category, and has yielded a large number of correct predictions and insights into protein function at the atomic level¹². The general method of molecular dynamics simulation is introduced by Mike P. Allen in a separate chapter of this book; here we shall, therefore, focus on the aspects specific for protein dynamics simulation.

3 Protein Dynamics Simulations

Molecular dynamics simulations compute the motion of every single atom within the simulation system (e.g., a protein solvated in physiological solvent (Fig.1, left), which is determined by the interaction forces (Fig.1, right) between all the n atoms of the system. Typically, these forces are described by a “force field” $V(\mathbf{x}_1, \dots, \mathbf{x}_n)$, which serves to compute a trajectory $\{\mathbf{x}_i(t_j)\}_{i=1\dots n; j=0\dots N}$ via the (classical) Newtonian equations of motion. This trajectory specifies the position $\mathbf{x}_i$ of each single atom in the system within a (discretized) period of time, $t_0 = 0, t_1 = \Delta t, t_2 = 2\Delta t, \dots, t_N = N\Delta t$, for N so-called integration steps. From such a trajectory one can subsequently calculate the observables of interest,

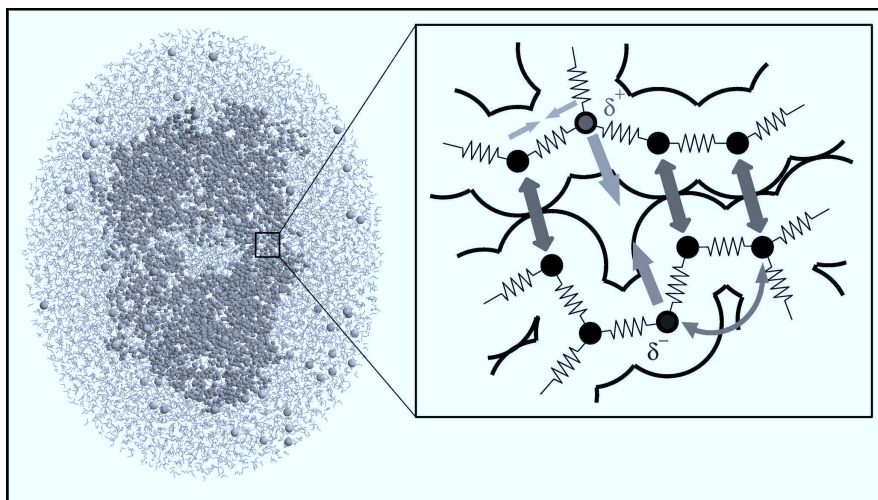


Figure 1. Left: A typical protein dynamics simulation system comprises a protein (dark grey atoms) solvated in water (light grey; only the O–H-bonds are shown). Salt ions (sparse light grey atoms) have been added to the aqueous solvent. Right: The inset shows part of the system, together with selected interatomic forces (e.g., chemical binding forces, Coulomb forces between atoms that carry partial charges, Pauli repulsion, van der Waals attraction) which determine the molecular motion.

often via averages.

Back in 1955, Fermi and co-workers have calculated numerically the motions of a one-dimensional chain of coupled anharmonic oscillators¹³. This first molecular dynamics simulation already highlights an essential aspect of the method: The simulation suggested an unexpected periodicity of the solution, and only with this simulation result in mind the authors could come up with a stringent derivation of a remarkable, non-linear partial differential equation (the Kortweg-de Vries-equation) for the continuum limit. Very similarly, the purpose of today's protein dynamics simulations is two-fold: first, to compute and predict observables; second, and often even more importantly, to provide the necessary insights and ideas which stimulate further studies.

Two years later, Alder and Wainwright studied systems of hard, two-dimensional discs¹⁴. For the first time, Rahman and Verlet used a more realistic interaction potential to study correlated motions and memory effects in liquid argon^{15,16} and, later, also in water¹⁷. The duration of those simulations was about 10 ps. The classical approximation of the nuclear motion also allowed to treat reactive scattering reactions in molecular beams such as the H+H₂-reaction¹⁸; these calculations served to compare the simulation results with transition state theories.

To facilitate the refinement of protein structure models to data from x-ray scattering experiments on protein crystals, and building upon earlier works on n-alkanes, cyclo-alkanes, and polypeptides¹⁹, Shneior Lifson suggested in 1969 to construct an interaction force field also for proteins²⁰. Combining infrared spectra, geometrical information from amide crystals, and quantummechanical calculations, a number of self consistent force field was developed and applied during the subsequent years.

The availability of these protein force fields finally triggered the first simulations of a protein^{21,22} (in gas phase, i.e., without surrounding solvent) for 10 ps. The static structure of hemoglobin, e.g., did not reveal the pathway via which oxygen atoms reach the binding site, which is deeply buried within the protein. The mentioned protein dynamics simulations provided the solution to this problem by suggesting that the thermal motion of the protein creates transient pathways for the oxygen molecule. This was the first of many subsequent examples for the relevance of protein dynamics at the atomic level for protein function²³.

Increased computer power and improved algorithms extended accessible system sizes and time scales. In 1998, the conformational dynamics of a small protein was simulated over one microsecond²⁴, requiring half a year of computer time on a 512 processor Cray T3E supercomputer. By joint use of donated computer time from more than 100,000 PCs located all over the world, recently a large number of trajectories could be accumulated for the study of protein folding processes, covering a total of several hundred microseconds²⁵.

For medium sized systems of about 100,000 atoms, currently simulations of several 10 nanoseconds durations are feasible at justifiable computational cost^{26,27}. One can estimate from this number that a further increase of computer power by a factor of about 10^4 will be required to cover most biochemical elementary reactions such as enzymatic catalysis or the transport of an ion across the membrane by ion pump proteins. For protein folding simulations, a 10^6 - to 10^8 -fold increase would be required. Assuming that the annual increase by a factor of 1.6 observed for the past 35 years continues into the future, these aims could be reached within 20 to 40 years. Today, 27 years after the first molecular dynamics simulation of a protein, and despite the impressive 10^5 -fold increase of computer power since then, the limited system size and, particularly, limited simulation lengths are our main obstacle in the attempt to derive the biochemical processes in proteins that enable our life from first principles physics laws.

4 Force Fields for Protein Dynamics

The forces between the atoms of a solvated macromolecule (Fig. 1 left) are diverse (inset, right). Chemical binding forces, symbolized by springs, enforce equilibrium distances or angles between chemically bound atoms (small arrows). Pauli repulsion forces (dark arrows) prevent atoms from moving through each other. Long ranged interactions, particularly electrostatic forces (light arrows) between partially charged atoms (δ^+ , δ^-) contribute substantially to the stability of protein structures and dominate the slow conformational dynamics. Hydrogen bonds, e.g., are mainly of electrostatic origin and only to a lesser extent a quantummechanical effect, and contribute significantly to the stability of α -helices and β -sheets²⁸.

All three forces (and several more, which are not discussed here) determine the three-dimensional structure of a protein as well as the motion of each single atom and, therefore, have to be considered in protein dynamics simulations. As mentioned before, this is achieved via the force field V , which describes the influence of the electronic dynamics on the motion of the nuclei. Most force fields are composed of empirically motivated interaction terms. Construction, parameterization, and testing of a force field is a challenging task, and substantial efforts will be required to further improve their accuracy.

Since the seventies, a number of force fields for biomolecules and, specifically for

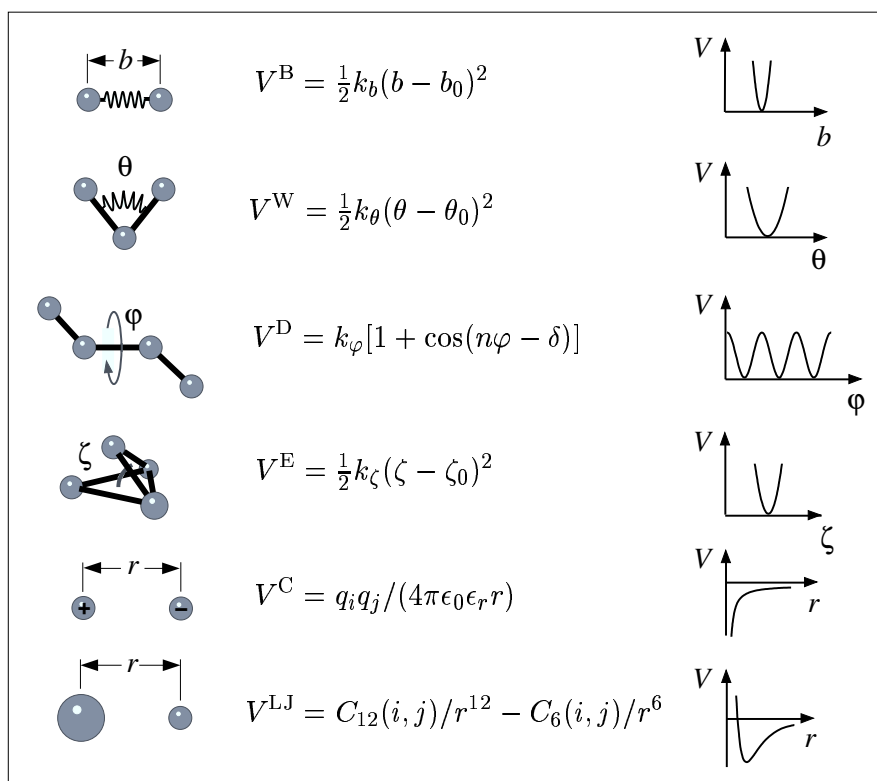


Figure 2. Interaction contributions to a typical force field (graphics adapted from Ref. 29). Bond stretch vibrations are described by a harmonic potential V^B , the minimum of which is at the equilibrium distance b_0 between the two atoms connected by chemical bond i [cf. eq. (1); the indices i, j etc. are not shown in the Figure]. Bond angles and out-of-plane angles are also described by harmonic potential terms, V^W and V^E , where Θ_0 and ζ_0 denote the respective equilibrium angles. Dihedral twists are subjected to a periodic potential V^D ; the respective force coefficients are denoted by k 's with appropriate indices. Non-bonded forces are described by Coulomb interactions, V^C , and Lennard-Jones potentials, $V^{LJ} = V^P + V^{\text{vdW}}$, where the latter includes the Pauli repulsion, $V^P \sim r^{-12}$, and the van der Waals interaction, $V^{\text{vdW}} \sim -r^{-6}$.

proteins and DNA/RNA have been developed. Well known and widely used are, e.g., CHARMM, GROMOS96, AMBER, and OPLS. Figure 2 shows interaction terms from which these force fields are typically composed. They describe interactions that arise from covalent chemical bonds as well as non-covalent interactions,

$$V = \sum_{\text{bonds } i} V_i^B + \sum_{\text{bond angles } j} V_j^W + \sum_{\text{dihedral angles } k} V_k^D + \sum_{\text{extraplanar angles } l} V_l^E + \sum_{\text{atom pairs } r,s} (V_{r,s}^C + V_{r,s}^P + V_{r,s}^{\text{vdW}}). \quad (1)$$

Covalent forces arise from changes of the length of chemical bonds ('B') and of the angle between two bonds ('W'), from twisting chemical bonds ('D'), and deviations of aromatic

carbon atoms from their in-plane positions ('E'). As non-covalent interactions the Coulomb interaction ('C'), the Pauli repulsion ('P') and the van der Waals interaction ('vdW') are included; the latter two are conveniently described by a Lennard-Jones-potential.

Many important details of protein dynamics simulations cannot be described and discussed at this basic level. Those include the treatment of the system boundaries or, alternatively, their periodic continuation, the "freezing" of fast bond vibrations to mimic their quantum mechanical character, the proper placement of salt ions in the vicinity of the protein, the treatment of protonable residues, simulations in canonical thermodynamic ensembles via appropriate coupling to heat and pressure baths, the treatment of non-polar hydrogen atoms through "compound atoms", pros and cons of explicit vs. implicit treatments of hydrogen bonds, the set-up of a simulation system using structures derived from x-ray crystallography and the problem of sufficiently long equilibration of the system, the implicit (or, in part, explicit) description of electronic polarizability, the efficient computation of the long-ranged Coulomb forces and the parallelization of the respective algorithms, as well as the proper choice of numerical integrators for the solution of the Newtonian equations of motion. Excellent reviews articles have been published on these topics³⁰⁻³²; for a full in-depth treatment, the reader is referred to the books of Gunsteren, Weiner and Wilkinson³³.

5 Force Probe Simulations^a

The recent years have seen dramatic improvements in single molecule experiments, particularly atomic force microscopy (AFM) methods, which triggered the force probe simulation technique described in this section.

In the experiments, the cantilever of the AFM microscope is used as a piconewton force sensor (Fig. 3 top right). The cantilever can be positioned to sub-nanometer accuracy, and its deflection — measuring the applied force — is detected via a laser beam with equally high precision.

Figure 3 (top left) illustrates the experiment, in which a ligand (biotin, light grey) is forced by the cantilever to leave the specific binding pocket of the receptor (streptavidin, dark grey). A polymer linker connects the proteins (with empty binding pockets) with a surface (left) as well as several biotin molecules with the tip of the cantilever (right). As the cantilever approaches the surface, several biotin/streptavidin complexes form, which dissociate one after the other upon subsequent retraction of the cantilever. Occasionally, one *single* complex remains intact until the very end of the experiment, in which case the force required to dissociate this last complex can be measured from the jump of the deflection of the cantilever to zero.

By repeating the experiment several hundred times one obtains a histogram of dissociation forces (Fig. 3 bottom right). Its maximum denotes the most probable dissociation force, in the shown example ca. 270 pN. This is the force that the non-covalent binding interaction between ligand and receptor can withstand at the time scale set by the experiment (typically milliseconds to seconds).

In a similar manner molecular forces have been measured recently for a number of systems. One example is the force generated by single motor proteins³⁶ which is used to

^aParts of this Section are adapted and translated from Ref. 34.

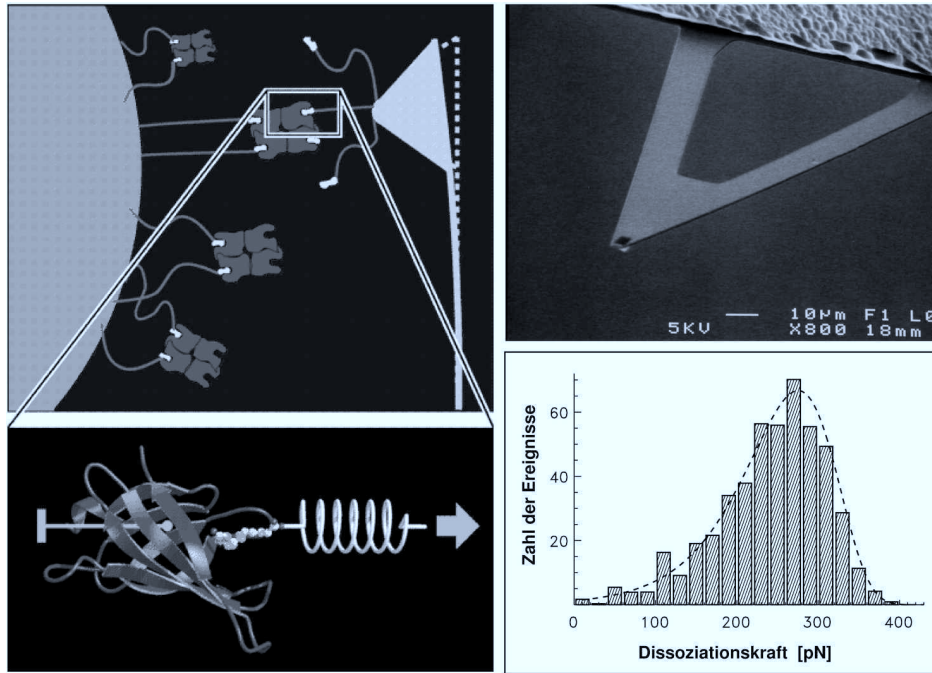


Figure 3. Top left: Principle of single molecule AFM experiments (see text); top right: electron microscopy picture of an AFM cantilever (picture kindly provided by Nanoscope); bottom left: simulation of single molecule AFM experiments (see text); bottom right: typical force histogram obtained from a series of single molecule unbinding experiments and fit to theory (dashed) [adapted from Rief et al.³⁵].

drive muscle contraction or to transport intracellular vesicles along filaments. An other one is the force generated by DNA polymerase upon transcription along the DNA primer³⁷.

However, in those experiments, the underlying atomistic dynamics and interactions that generate the observed forces could not be observed, which is the main motivation to simulate these experiments by means of protein dynamics simulations (Fig. 3 bottom left). To that aim, and modeling the effect of the cantilever, an additional harmonic “pulling potential”,

$$V(\mathbf{R}_i, t) = \frac{1}{2}k[(\mathbf{R}_i - \mathbf{R}_i^0) \cdot \hat{\mathbf{n}} - vt]^2 \quad (2)$$

is included within the molecular force field such that it acts on that particular atom i of the ligand molecule, which in the real experiment is covalently connected to the cantilever via the polymer linker. This ensures that in the simulation the atom is subjected to the same force as in the experiment. In the Figure, this pulling potential is symbolized by a spring. The normalized vector $\hat{\mathbf{n}}$ denotes the direction of pulling, and $\mathbf{R}_i^0$ the position of atom i at the start of the simulation. As can be seen from eq. (2), the minimum of the pulling potential (i.e., the equilibrium position of atom i) is subsequently moved with constant velocity v away from the binding pocket (arrow).

To avoid drift of the protein, the center of mass of the protein (consisting of n_p atoms

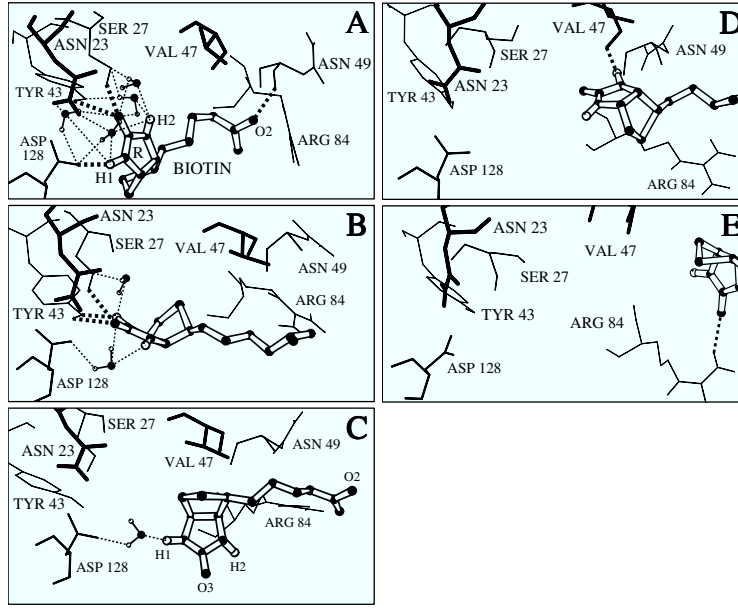


Figure 4. Snapshots of enforced dissociation of a biotin molecule (white, thick contours) from the streptavidin binding pocket (only the few relevant residues of the binding pocket are shown as stick-models). Hydrogen bonds are shown as bold dashed lines, few of the many water bridges with thin, dotted lines. The simulation length is one nanosecond; during this time the biotin molecule is moved by about one nanometer. (Adapted from Ref. 38).

with positions $\mathbf{R}_i, i = 1 \dots n_p$) is kept in place at $\mathbf{R}_{\text{fix}}$ by a second harmonic (but stationary) potential,

$$V_{\text{fix}}(\mathbf{R}_1, \dots, \mathbf{R}_{n_p}) = \frac{1}{2} k_{\text{fix}} \left[\mathbf{R}_{\text{fix}} - \frac{1}{n_p} \sum_{i=1}^{n_p} \mathbf{R}_i \right]^2, \quad (3)$$

with suitable spring coefficient k_{fix} . This ensures that the protein is free to undergo internal motions, such as induced-fit motions upon ligand unbinding, and that it can adopt to the pulling direction by rotations — as in the real experiment —, but prevents translational motions.

These force probe AFM simulations of enforced dissociation allowed one to compute dissociation forces, which agreed well with those measured in the AFM experiments. In such comparisons one has to carefully take into account the fact that the AFM measurements and the force probe simulations are carried out at quite different time scales, namely milliseconds versus nanoseconds. Therefore, in order to compare the respective dissociation forces, the computed force has to be re-scaled to the measured ones, e.g., for the simplest case using eq. (13), as will be described below.

Figure 4 shows a detailed model of streptavidin/biotin dissociation derived from force probe simulations, highlighting the sequence and type of stretching and subsequent rupture of single non-covalent bonds between the ligand and the receptor^{38,39}. These complex sequence of localized rupture events give rise to a complex “force profile” (Fig. 5), i.e., the

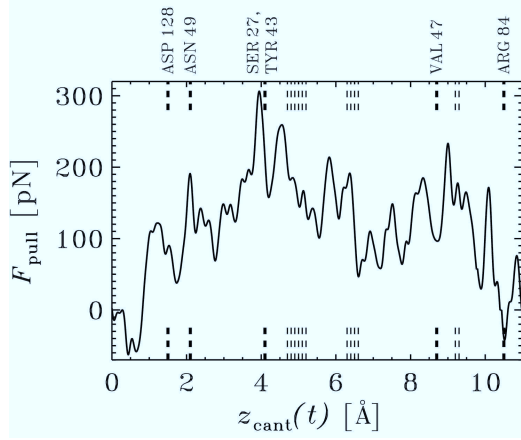


Figure 5. Force exerted onto the biotin ligand during enforced dissociation (“force profile”). A large number of local force maxima can be seen, each of which can be attributed to the rupture of single non-covalent bonds like hydrogen bonds or water bridges. The bold dashed lines at the top denote the respective residues, which are also shown in Fig. 4. the thin dashed lines denote rupture of water bridges (adapted from Ref. 38).

exerted force plotted during the simulation,

$$F_{\text{pull}} = \hat{\mathbf{n}} \cdot \nabla V(\mathbf{R}_i, t) = k \hat{\mathbf{n}} \cdot [(\mathbf{R}_i - \mathbf{R}_i^0) \cdot \hat{\mathbf{n}} - vt]. \quad (4)$$

This type of force probe simulations, also called “steered molecular dynamics”, is now widely used, e.g., to elucidate at the atomic level the structural changes that are induced by mechanically unfolding proteins like titin^{40,41} or by stretching various other polymers^{42–44}, and to explain the measured forces in terms of intramolecular interactions. For more extended review, we would like to refer the reader to Refs. 35,45.

6 Simple Non-Equilibrium Statistical Mechanics of Enforced Dissociation^b

For complex systems like proteins, the measured and computed dissociation or unfolding forces vary with the time scale at which the process is enforced to occur. Usually, the forces increase with shorter time scales, i.e., increasing loading rates. For thermodynamic equilibrium, such time scale dependency should not appear; hence, the underlying processes are likely to proceed far from equilibrium. This observation motivated the development of non-equilibrium theories of enforced unbinding reactions^{47,48,39,49–51}. One of these shall be sketched here.

We assume an effective free energy $G(x)$ along a (one-dimensional) reaction coordinate x (Fig. 6, solid bold line), e.g., the distance between the center of mass of the ligand and the one of the receptor, which governs the dissociation reaction.

As sketched in the two Figures, the bound state (left minimum) is separated from the dissociated state (right) by a more or less structured energy barrier (global maximum). The

^bParts of this section are translated from Ref. 46.

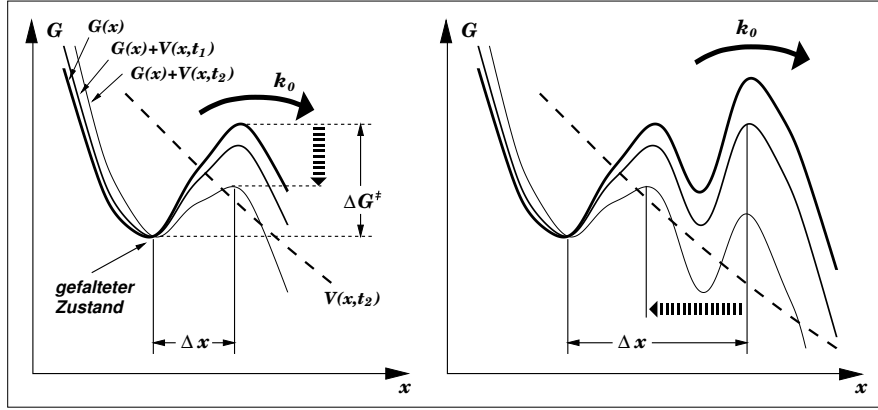


Figure 6. Time dependent Gibbs free energy in force probe experiments and simulations. A time-dependent potential $V(x, t)$ describes the force exerted by the cantilever and modifies the unperturbed energy landscape $G(x)$. As a result, the barrier that determines the spontaneous dissociation rate k_0 (black arrow) is reduced (left: dashed vertical arrow). For a more complex energy landscape (right), also the dissociation length Δx (dashed horizontal arrow), i.e., the distance between minimum and maximum of the free energy landscape, may change.

height of this barrier determines the rate coefficient k_0 for thermally activated spontaneous dissociation,

$$k_0 = \omega_0 \exp(-\beta \Delta G^\ddagger). \quad (5)$$

Here, $\beta = 1/(k_B T)$ denotes the reciprocal thermal energy and ω_0 the attempt frequency or Kramers' prefactor of the bound state^{52–54}. For streptavidin/biotin dissociation, e.g., the reciprocal rate coefficient is several months, whereas it can be as short as milliseconds for antibody/antigen dissociation. Note that if unbinding is enforced at a time scale that is slower than the reciprocal rate coefficient, no dissociation force is measured, since in this case the system has fallen apart already before any significant force has been applied.

As is the force probe simulations described above, the force exerted by the cantilever is assumed harmonic and moving with velocity v ,

$$V(x, t) = \frac{1}{2} k(x - vt)^2, \quad (6)$$

(dashed curve in Fig. 6), where k is the effective spring coefficient. Here we restrict our discussion to the case of small spring coefficients, i.e., soft springs, which covers most realistic situations. Therefore, $V(x, t)$ appears in the Figure as a nearly straight line, the slope of which increases linearly with time. Since the total potential $G(x) + V(x, t)$ now becomes time dependent, so does the barrier height $\Delta G^\ddagger$, and, hence, the dissociation coefficient k_0 . This effect is demonstrated by the thin, solid lines. As can be seen, the barrier height decreases with time. Dissociation occurs when the barrier has become small enough to be overcome thermally activated at the time scale $\tau = \Delta G^\ddagger / (kv\Delta x)$ set by the loading rate kv .

For simple and localized barriers (left picture), we will neglect the slight shift of the position of the minimum that also occurs. In this case, a simple two-state model⁴⁷ is appropriate, which neglects the details of $G(x)$ and assumes a linear decrease of the barrier

height with time,

$$\Delta G^\ddagger(t) = \Delta G^\ddagger - kvt\Delta x. \quad (7)$$

This implies a time dependent dissociation coefficient [cf. eq. (5)],

$$k_0(t) = \omega_0 e^{-\beta(\Delta G^\ddagger - kvt\Delta x)}, \quad (8)$$

which holds as long as $\Delta G^\ddagger - kvt\Delta x$ is larger than $k_B T$ and back reactions can be neglected. In analogy to the treatment of radioactive decay, the flux across the barrier decreases the probability $P(t)$ to find the system in the bound state at time t ,

$$\frac{dP(t)}{dt} = -P(t)k_0(t) = -P(t)\omega_0 e^{-\beta(\Delta G^\ddagger - kvt\Delta x)}, \quad (9)$$

with $P(t=0) = 1$. Solution of eq. (9),

$$P(t) = \exp \left[\frac{\omega_0}{\beta kv \Delta x} e^{-\beta \Delta G^\ddagger} (1 - e^{\beta kv t \Delta x}) \right], \quad (10)$$

yields a distribution $p(t_D)dt_D = -\frac{dP(t)}{dt}\bigg|_{t_D} dt_D$ of dissociation times t_D ,

$$p(t_D)dt_D = \omega_0 e^{-\beta(\Delta G^\ddagger - kvt_D \Delta x)} \exp \left[\frac{\omega_0}{\beta kv \Delta x} e^{-\beta \Delta G^\ddagger} (1 - e^{\beta kv t_D \Delta x}) \right] dt_D, \quad (11)$$

and dissociation forces $F_D = kvt_D$, respectively,

$$p(F_D)dF_D = \omega_0 e^{-\beta(\Delta G^\ddagger - F_D \Delta x)} \exp \left[\frac{\omega_0}{\beta kv \Delta x} e^{-\beta \Delta G^\ddagger} (1 - e^{\beta F_D \Delta x}) \right] dF_D. \quad (12)$$

As the main result, the maximum ($dP(F_D)/dF_D = 0$) denoting the most probable dissociation force, reads

$$F_{\max}(v) = \frac{\Delta G^\ddagger}{\Delta x} + \frac{1}{\beta \Delta x} \ln \frac{\beta kv \Delta x}{\omega_0} = \frac{1}{\beta \Delta x} \ln \frac{\beta kv \Delta x}{k_0}, \quad (13)$$

(where eq. (5) has been used) and thus increases logarithmically with the loading rate kv .

This result links the rupture length Δx to the slope of the loading rate dependent dissociation force and thus provides the basis for dynamic force spectroscopy (Fig. 7A) The larger Δx , the faster the energy barrier is decreased (cf. Fig. 6), and the stronger is the effect of the loading rate kv on the dissociation time t_D and dissociation force F_D . Furthermore, this result serves to rescale dissociation forces from force probe simulations to the much slower time scales of the AFM experiments.

For more complex energy landscapes as shown in Fig. 6 (right), the simple two-state model is not applicable. For the two barriers shown, there is a critical pulling force for which the left maximum becomes the global one (cf. the thin curve in the Figure) at the time of dissociation and, hence, the rupture length Δx jumps to smaller values. Accordingly, the logarithmic slope of the dissociation force increases as shown in Fig. 7B. For a more general treatment, the reader is referred to Refs. 50,51.

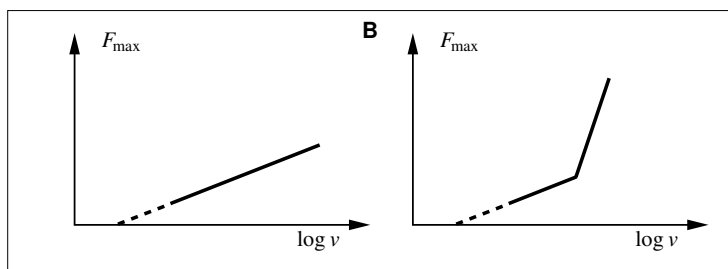


Figure 7. Dynamic force spectra show the increase of the dissociation force $F_{\max}$ with pulling velocity v of the cantilever or, equivalently, loading rate kv . A: For the case of a well-localized energy barrier in $G(x)$, the simple two-state model described in the text predicts a logarithmic increase. B: For more complex energy landscapes such as the two barriers sketched in Fig. 6 (right), the dynamic force spectrum falls into two regimes. For small v , the larger barrier in $G(x)$ dominates, and the resulting large Δx yields a slow increase of the spectrum with v . For large v , however, the (originally) smaller barrier to the left dominates at the time t_D of dissociation, resulting in a steeper logarithmic increase of $F_{\max}$ with v .

7 Primary Mechanical Energy Transfer Steps in the “Molecular Nano-Machine” F_1 ATP Synthase ^c

As an example we consider force probe simulations that aim at elucidating the primary mechanical energy transfer steps in F_1 ATP synthase. In fact, this protein is the smallest rotary motor known to date.

The mitochondrial membrane protein F_0F_1 -ATP synthase synthesizes adenosinetriphosphate (ATP), the universal currency of energy in the cell. In the human body, about 70 kg of ATP are synthesized by these proteins each day. The protein (Fig. 8a,b) utilizes the electrochemical potential of a transmembrane proton gradient for ATP synthesis⁵⁵. Remarkably, energy conversion involves mechano-chemical energy transport from a rotating asymmetric γ -“stalk” (red) over a distance of more than 2.5 nm to the three ATP synthesis sites of the protein (Fig. 8a)^{56–58}. X-ray structures of the F_1 -unit^{59,60} show its three β -subunits in different conformational states, TP, DP, and E, with bound ATP (analog), ADP, and empty binding site, respectively⁵⁹. The structures suggested that the energy consuming step in the ATP synthesis cycle, namely ATP release, is driven by concerted conformational motions within the $(\alpha\beta)_3$ -complex, which are dictated by the changing orientation of the asymmetric γ -subunit inside $(\alpha\beta)_3$ (Ref. 59), thus rendering F_0F_1 -ATP synthase a spectacular example of a “molecular machine”.

Our molecular dynamics simulations included the F_1 -unit, $(\alpha\beta)_3\gamma$ (except δ and b), bound nucleotides as shown in Fig. 8a, a large solvation shell (blue), and salt ions (red and yellow), comprising a total of 183 674 atoms. In the simulations, the lower part of the γ -stalk was enforced to rotate (Fig. 8c), thus resembling the rotation caused *in vivo* by protonmotive F_0 -rotation, and the subsequently triggered conformational motions have been studied.

The n backbone atoms of the lower part of the γ subunit with positions $\mathbf{x}_i$ ($\bullet$) were

^cThis Section is adapted and shortened from Ref. 27.

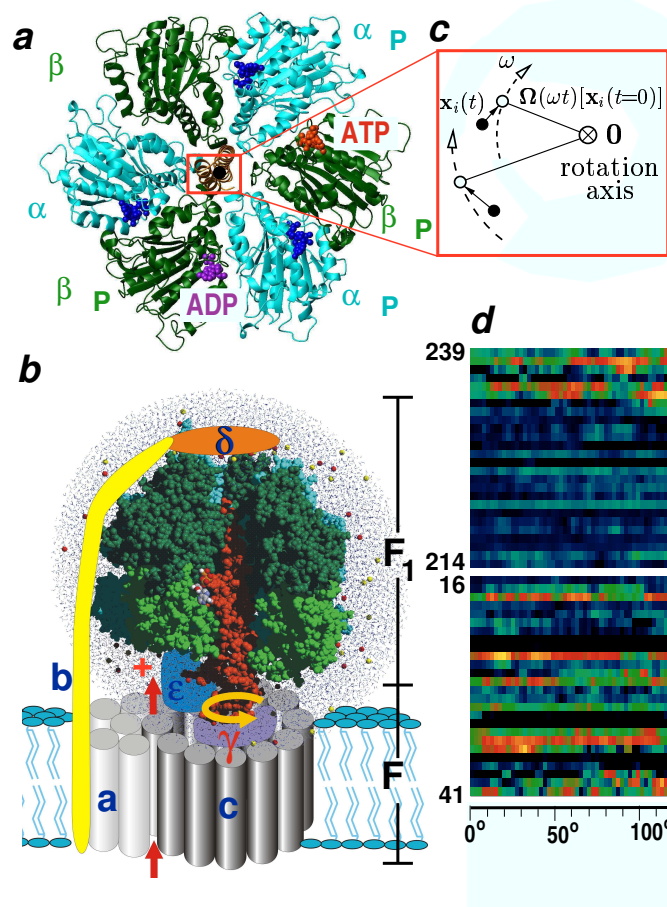


Figure 8. ATP synthase structure, function, and simulation setup. a: Slice through F_1 at the level of the active sites; b: F_0F_1 -ATP synthase with the atomically resolved F_1 simulation system. A proton flux through the membrane-bound F_0 -unit (red) drives rotation of the asymmetric γ -“stalk” (orange) with respect to the F_1 -unit (green, cyan) located in the mitochondrial matrix; co-rotation of F_1 is hindered by δ and b (yellow). This rotation triggers conformational changes mainly within the three β -subunits, DP, TP, and E (green), which bind ADP and phosphate, and subsequently drive ATP synthesis and release. These changes concern most notably the orientation of the “lower half” of the β -subunit (light green), which can tilt with respect to the “upper half” (dark green). Accordingly, the conformation seen for the empty subunit, β_E , is denoted “open”, whereas β_{TP} and β_{DP} are referred to as “closed”⁵⁹. After 120° rotation of γ , DP, TP, and E switch roles clockwise. c: In the simulations, a torque potential is used to enforce rotation of the γ subunit (see text). d: The torque required to move the individual residues of γ within $(\alpha\beta)_3$ (black: zero or negative torque; yellow: 0.7×10^{-18} Nm) allows to identify “hot spots” of mechanical energy transfer.

subjected to a rotating potential,

$$V = \frac{1}{2}k \sum_{i=1}^n [\Omega(\omega t)\mathbf{x}_i(t=0) - \mathbf{x}_i(t)]^2. \quad (14)$$

The rotation matrix Ω moved the minima of V ($\circ$) along concentric circles around the

symmetry axis of the complex, using a stiff spring coefficient k . The torques $\Theta_i(t)$ exerted on the individual atoms were computed from the deflection of the actual atomic positions $\mathbf{x}_i(t)$ from the respective minima,

$$\Theta_i(t) = \Omega(\omega t) \mathbf{x}_i(t=0) \times k[\Omega(\omega t) \mathbf{x}_i(t=0) - \mathbf{x}_i(t)]. \quad (15)$$

Throughout the simulations, co-rotation of the $(\alpha\beta)_3$ -complex was prevented by immobilizing the centers of mass of the putative δ contact region of two α -subunit head groups, thereby mimicking the effect of the δ “stator”.

With this approach we studied how the rotating γ -stalk induces conformational motions within the β subunits, how and along which route they propagate towards the nucleotide binding sites, and why and in what sequence the observed structural changes lower ATP binding affinity at β_{TP} , as is required for ATP to leave the binding pocket after synthesis.

Due to the huge computational cost of such simulations, the covered time span is limited to several nanoseconds, i.e., considerably shorter than the microseconds time scale at which the studied conformational motions likely take place *in vivo*⁶¹. Accordingly, torques that are larger by factors of 12 and 50, respectively, than the measured ones⁶² had to be applied in the simulations. Furthermore, slow (microsecond) conformational motions are unlikely to show up. For the faster motions that are actually seen in the simulations, however, the large amount of available experimental data enabled to assure that these motions are actually described correctly.

After equilibration, a 120° force probe rotation simulation in synthesis direction was carried out with an angular velocity of $\omega = 115^\circ \text{ ns}^{-1}$, as sketched in Fig. 8c. During a subsequent relaxation period of 3.5 ns, within which the lower part of the γ -subunit was kept restrained to avoid back-rotation, the induced conformational motions propagated further towards the active sites. The torques exerted onto the individual γ -residues during the 120° enforced rotation (Fig. 8d) show peaks, reflecting energy transfer to $(\alpha\beta)_3$, dissipation, and enforced crossing of associated energy barriers.

Figure 9 shows the main changes for β_{TP} , β_E , and β_{DP} during enforced rotation and subsequent free dynamics. The lower part of β_{TP} is pushed sideways towards α_{DP} , and downwards towards an open configuration (Fig. 9b). This observation is quantified by the change of the bending angle (inset) between the two domains separated by the dashed line in Fig. 9a, which approaches the one seen in β_E .

Simultaneously, an upwards rotation of the β_E lower half towards the closed x-ray conformation of β_{TP} and β_{DP} is observed (Fig. 9c). This motion is initially sterically obstructed by both the γ -subunit and β_{TP} and, therefore, proceeds synchronously to the removal of these obstacles. Much less changes are seen for β_{DP} (d), where the orientation of the lower part is nearly unchanged. Overall, the opening and closing motions of the lower halves of the β -subunits are nearly completed already within the 7 ns simulation time span without explicitly being forced to.

Focusing at the catalytic ATP binding site (Fig. 10), a sequential retraction of $\beta\text{Arg 191}$, $\alpha\text{Arg 373}$, and $\beta\text{Arg 189}$ from ATP is seen, quantified by the respective distance changes (inset, blue curves). This sequence is initiated, immediately after γ -rotation, by a 0.5 nm shift of $\beta\text{Arg 191}$ towards α_{TP} , which forms a new hydrogen bond to $\alpha\text{Gln 163}$, thereby weakening its interaction to ATP. Subsequently, initially strong hydrogen bonds (dashed lines) between ATP and residues $\alpha\text{Arg 373}$ and $\beta\text{Arg 189}$ rupture, thus further decreasing ATP binding affinity. The whole process is compatible with the “zipper-model”⁶³. Assum-

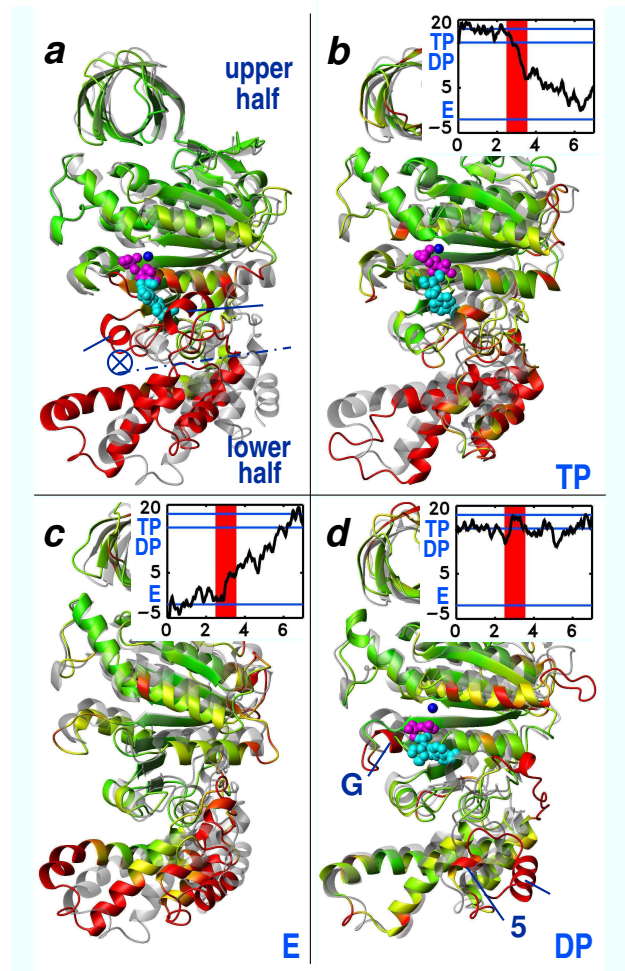


Figure 9. Structural changes of the β -subunits induced by enforced rotation of the γ -stalk. The size of these changes is color coded (green: small; red: large). a: Overlay of β_{TP} (colored) and β_E -subunits (grey) taken from the x-ray structure, which are in closed and open conformation, respectively. b–d: Changes in β_{TP} , β_E , and β_{DP} , as induced by enforced 120° rotation and subsequent free dynamics; shown are initial (x-ray) structures (grey) and final structures (colored). The insets show the angle between upper and lower part of the respective β -subunits in the course of the 7 ns simulation (axis: $\otimes$). The red bar indicates the period of enforced rotation, green lines denote the three β -subunit angles of the equilibrated x-ray structure.

ing that changes of ATP binding affinity are dominated by the interaction between ATP and the binding pocket, a continuous ATP affinity decrease of β_{TP} can be inferred from the increasing interaction energy estimate between ATP and the binding pocket (inset, black curve) as the net effect of the observed structural changes.

Taken together, the enforced γ -rotation suggests for this beautiful molecular machine a sequence of structural changes that effect primary mechano-chemical energy transfer steps within F_1 -ATP synthase. Many structural changes that are expected from the three

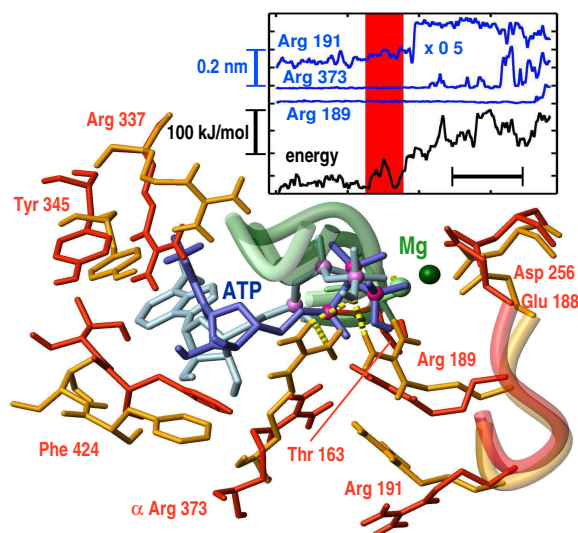


Figure 10. Changes at the β_{TPP} ATP binding site after the 120° and free dynamics. Shown are selected residues involved in ATP binding before (yellow, ATP in light blue) and after enforced rotation and subsequent free dynamics (red, ATP in blue). The backbones of part of the P-loop and of α are colored green and orange, respectively. Residue numbers refer to β_{TPP} if not annotated otherwise. The inset shows distances (blue) between the donor nitrogen atoms of three arginine residues involved in ATP binding and the respective ATP phosphates (reduced scale for Arg 191), as well as an estimate (black) for the change in interaction energy between ATP and the binding pocket in the course of the simulation. The period of enforced rotation is highlighted red.

x-ray intermediates are correctly reproduced. Because none of these motions have been artificially directed towards the expected target structure, the agreement shows that the necessarily accelerated rotation of the γ -stalk does not severely change their character.

Moreover, the simulations provide a *causal* picture of a sequence of structural motions, one triggering the next. Already within the simulation time span of 7 ns, the structural changes proceed sufficiently far to transfer a significant amount of energy to the β_{TPP} binding site, as estimated from its decreased ATP binding affinity. This decrease is affected gradually, by sequential and well concerted retraction of three charged arginines from the bound ATP.

Acknowledgments

I would like to thank Berthold Heymann and Rainer Böckmann for their contributions which are presented here.

References

1. Jeremy M. Berg, John L. Tymoczko, and Lubert Stryer. *Biochemistry*. W. H. Freeman and Company, New York, 2002.

2. Ken A. Dill and Sarina Bromberg. *Molecular Driving Forces: Statistical Thermodynamics in Chemistry and Biology*. Garland Publishing, London, 2003.
3. Thomas E. Creighton. Protein folding. *Biochem. J.*, 270:1–16, 1990.
4. F. C. Bernstein, T. F. Koetzle, G. J. B. Williams, E. F. Meyer, M. D. Brice, J. R. Rodgers, O. Kennard, T. Shimanouchi, and M. J. Tasumi. The protein data bank: A computer-based archival file for molecular structures. *J. Molec. Biol.*, 112:535–542, 1977.
5. J. C. Kendrew, G. Bodo, H. N. Dintzis, R. G. Parrish, H. Wyckoff, and D. C. Phillips. 3-dimensional model of the myoglobin molecule obtained by x-ray analysis. *Nature*, 181:662–666, 1958.
6. A. Ansari, J. Berendzen, S. F. Browne, H. Frauenfelder, I. E. T. Iben, T. B. Sauke, E. Shyamsunder, and R. D. Young. Protein states and protein quakes. *Proc. Natl. Acad. Sci. USA*, 82:5000–5004, 1985.
7. Erwin Neher and Bert Sakmann. Single-channel currents recorded from membrane of denervated frog muscle fibres. *Nature*, 260:799–802, 1976.
8. Bert Sakmann and Erwin Neher, editors. *Single-Channel Recording*. Plenum Press, 2nd edition, 1995.
9. Gary Yellen. The voltage-gated potassium channels and their relatives. *Nature*, 419:35–42, 2002.
10. Anjum Ansari, Joel Berendzen, David Braunstein, Benjamin R. Cowen, Hans Frauenfelder, Mi Kyung Hong, Icko E. T. Iben, J. Bruce Johnson, Pál Ormos, Todd B. Sauke, Reinhard Scholl, Alfons Schulte, Peter J. Steinbach, Joseph Vittitow, and Robert D. Young. Rebinding and relaxation in the myoglobin pocket. *Biophys. Chem.*, 26:337–355, 1987.
11. R. Elber and M. Karplus. Multiple conformational states of proteins: A molecular dynamics analysis of myoglobin. *Science*, 235:318–321, 1987.
12. H. J. C. Berendsen. Bio-molecular dynamics comes of age. *Science*, 271:954–955, 1996.
13. E. Fermi, J. Pasta, and S. Ulam. Studies in nonlinear problems, I. Los Alamos Report LA 1940, 1955.
14. B. J. Alder and T. E. Wainwright. Phase transition for a hard sphere system. *J. Chem. Phys.*, 27:1208–1209, 1957.
15. A. Rahman. Correlations in the motion of atoms in liquid argon. *Phys. Rev.*, 136:A405–A411, 1964.
16. Loup Verlet. Computer “experiments” on classical fluids. I. Thermodynamical properties of Lennard-Jones molecules. *Phys. Rev.*, 159:98–103, 1967.
17. A. Rahman and F. H. Stillinger. Molecular dynamics study of liquid water. *J. Chem. Phys.*, 55:3336–3359, 1971.
18. M. Karplus, R. N. Porter, and R. D. Sharma. Exchange reactions with activation energy. I. Simple barrier potential for (H, H₂). *J. Chem. Phys.*, 43:3259–3287, 1965.
19. H. A. Scheraga, R. A. Scott, and K. D. Gibson. Calculations of stable polypeptide conformations. *Fed. Proc.*, 25:345, 1966.
20. M. Levitt and Shneior Lifson. Refinement of protein conformation using a macromolecular energy minimization procedure. *J. Molec. Biol.*, 46:269–279, 1969.
21. J. A. McCammon, B. R. Gelin, and M. Karplus. Dynamics of folded proteins. *Nature*, 267:585–590, 1977.

22. W. F. van Gunsteren and H. J. C. Berendsen. Algorithms for macromolecular dynamics and constraint dynamics. *Molec. Phys.*, 34:1311–1327, 1977.
23. M. Karplus and J. A. McCammon. The dynamics of proteins. *Scient. Am.*, 4:30, 1986.
24. Y. Duan and P. A. Kollman. Pathways to a protein folding intermediate observed in a 1-microsecond simulation in aqueous solution. *Science*, 282:740–744, 1998.
25. Christopher D. Snow, Houbi Nguyen, Vijay S. Pande, and Martin Gruebele. Absolute comparison of simulated and experimental protein-folding dynamics. *Nature*, 420:102–106, 2002.
26. Bert L. de Groot and Helmut Grubmüller. Water permeation across biological membranes: Mechanism and dynamics of aquaporin-1 and GlpF. *Science*, 294:2353–2357, 2001.
27. Rainer Böckmann and Helmut Grubmüller. Nanoseconds molecular dynamics simulation of primary mechanical energy transfer steps in F₁-ATP synthase. *Nature Struct. Biol.*, 9:198–202, 2002.
28. L. Pauling, R. B. Corey, and H. R. Branson. The structure of proteins: Two hydrogen-bonded helical configurations of polypeptide chain. *Proc. Natl. Acad. Sci. USA*, 37:205, 1951.
29. Gunnar Schröder. Molekulardynamiksimulation der Flexibilität und Fluoreszenz-anisotropie eines an ein Protein gebundenen Farbstoffs. Master's thesis, Universität Göttingen, 2000.
30. Wilfred F. van Gunsteren and Herman J. C. Berendsen. Computer simulation of molecular dynamics: Methodology, applications, and perspectives in chemistry. *Angew. Chem. Int. Ed.*, 29:992–1023, 1990.
31. M. Parrinello. Simulating complex systems without adjustable parameters. *IEEE Computational Science & Engineering*, 2:22–27, 2000.
32. M. E. Tuckerman and G. J. Martyna. Understanding modern molecular dynamics: Techniques and applications. *J. Phys. Chem. B*, 104:159–178, 2000.
33. Wilfred F. van Gunsteren, Paul K. Weiner, and Anthony J. Wilkinson, editors. *Computer Simulation of Biomolecular Systems: Theoretical and experimental applications*, volume 1-3. Escom, Leiden, The Netherlands, 1989–1997.
34. Helmut Grubmüller. Computersimulation der Dynamik von Biomolekülen. Book Chapter for Bergmann/Schaefer, Lehrbuch der Experimentalphysik, Band 5: Vielteilchensysteme. In press.
35. Matthias Rief and Helmut Grubmüller. Kraftspektroskopie von einzelnen Biomolekülen. *Physikalische Blätter*, pages 55–61, Feb. 2001.
36. J. T. Finer, R. M. Simmons, and J. A. Spudich. Single myosin molecule mechanics: piconewton forces and nanometer steps. *Nature*, 368:113–119, 1994.
37. H. Yin, M. D. Wang, K. Svoboda, R. Landick, S. M. Block, and J. Gelles. Transcription against an applied force. *Science*, 270:1653–1657, 1995.
38. Helmut Grubmüller, Berthold Heymann, and Paul Tavan. Ligand binding: Molecular mechanics calculation of the streptavidin-biotin rupture force. *Science*, 271:997–999, 1996.
39. S. Izrailev, S. Stepaniants, M. Balsara, Y. Oono, and K. Schulten. Molecular dynamics study of unbinding of the avidin-biotin complex. *Biophys. J.*, 72:1568–1581, 1997.
40. H. Lu, B. Isralewitz, A. Krammer, V. Vogel, and K. Schulten. Unfolding of titin

- immunoglobulin domains by steered molecular dynamics simulation. *Biophys. J.*, 75:662–671, 1998.
41. P. Marszalek, H. Lu, H. Li, M. Carrion-Vazquez, A. F. Oberhauser, K. Schulten, and J. M. Fernandez. Mechanical unfolding intermediates in titin modules. *Nature*, 402:100–103, 1999.
 42. Matthias Rief, Filipp Oesterhelt, Berthold Heymann, and Hermann E. Gaub. Single molecule force spectroscopy reveals conformational change in polysaccharides. *Science*, 275:1295–1297, 1997.
 43. A. D. MacKerell and G. U. Lee. Structure, force, and energy of a double-stranded DNA oligonucleotide under tensile loads. *Europ. Biophys. J.*, 28:415–426, 1999.
 44. Berthold Heymann and Helmut Grubmüller. Elastic properties of poly(ethylene-glycol) studied by molecular dynamics stretching simulations. *Chem. Phys. Lett.*, 307:425–432, 1999.
 45. B. Isralewitz, M. Gao, and K. Schulten. Steered molecular dynamics and mechanical functions of proteins. *Curr. Opin. Struct. Biol.*, 11:224–230, 2001.
 46. Helmut Grubmüller. Kraftmikroskopische Experimente an einzelnen Biomolekülen. Book Chapter for Bergmann/Schaefer, Lehrbuch der Experimentalphysik, Band 5: Vielteilchensysteme. In press.
 47. G. I. Bell. Models for specific adhesion of cells to cells. *Science*, 200:618–627, 1978.
 48. Evan Evans and Ken Ritchie. Dynamic strength of molecular adhesion bonds. *Biophys. J.*, 72:1541–1555, 1997.
 49. U. Seifert. Rupture of multiple parallel molecular bonds under dynamic loading. *Phys. Rev. Lett.*, 84:2750–2753, 2000.
 50. Berthold Heymann and Helmut Grubmüller. Dynamic force spectroscopy of molecular adhesion bonds. *Phys. Rev. Lett.*, 84:6126–6129, 2000.
 51. G. Hummer and A. Szabo. Kinetics from nonequilibrium single-molecule pulling experiments. *Biophys. J.*, 85:5–15, 2003.
 52. H. Eyring. The activated complex in chemical reactions. *J. Chem. Phys.*, 3:107–115, 1935.
 53. H. A. Kramers. Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica (Utrecht)*, VII:284–304, 1940.
 54. Peter Hänggi, Peter Talkner, and Michal Borkovec. Reaction-rate theory: Fifty years after Kramers. *Rev. Mod. Phys.*, 62:251–341, Apr. 1990.
 55. P. D. Boyer. The binding change mechanism for ATP synthase — some probabilities and possibilities. *Biochim. Biophys. Acta*, 1140:215–250, 1993.
 56. T. M. Duncan, V. V. Bulygin, Y. Zhou, M. L. Hutcheon, and R. L. Cross. Rotation of subunits during catalysis by *Escherichia coli* F₁-ATPase. *Proc. Natl. Acad. Sci. USA*, 92:10964–10968, 1995.
 57. D. Sabbert, S. Engelbrecht, and W. Junge. Intersubunit rotation in active F-ATPase. *Nature*, 381:623–625, 1996.
 58. H. Noji, R. Yasuda, M. Yoshida, and K. Kinosita. Direct observation of the rotation of F₁-ATPase. *Nature*, 386:299–302, 1997.
 59. Jan Pieter Abrahams, Andrew G. W. Leslie, René Lutter, and John E. Walker. Structure at 2.8 Å resolution of F₁-ATPase from bovine heart mitochondria. *Nature*, 370:621–628, 1994.
 60. C. Gibbons, M. G. Montgomery, A. G. W. Leslie, and J. E. Walker. The structure

- of the central stalk in bovine F_1 -ATPase at 2.4 Å resolution. *Nature Struct. Biol.*, 7:1055–1061, 2000.
61. R. Yasuda, H. Noji, M. Yoshida, K. Kinoshita, and H. Itoh. Resolution of distinct rotational substeps by submillisecond kinetic analysis of F_1 -ATPase. *Nature*, 410:898–904, 2001.
 62. O. Pänke, D. A. Cherepanov, K. Gumbiowski, S. Engelbrecht, and W. Junge. Viscoelastic dynamics of actin filaments coupled to rotary F-ATPase: Angular torque profile of the enzym. *Biophys. J.*, 81:1220–1233, 2001.
 63. G. Oster and H. Y. Wang. Reverse engineering a protein: the mechanochemistry of ATP synthase. *Biochim. Biophys. Acta*, 1458:482–510, 2000.

Mechanical Functions of Proteins

Klaus Schulten

Beckman Institute and Department of Physics,
University of Illinois at Urbana-Champaign, U.S.A.
E-mail: kschulte@ks.uiuc.edu

Numerous functions of cells involve mechanical properties of biopolymers. Recent studies of single molecule measurement techniques like atomic force microscopy and optical tweezers have greatly advanced knowledge of these properties that govern cell motion, cell adhesion, intra-cellular motor-driven transport, or the expression of the genome. Steered molecular dynamics (SMD) complements these observations through computational modeling and provides atomic level descriptions of the underlying events¹. For this purpose SMD applies external forces to mimic experimental set-ups or to manipulate biomolecules to probe mechanical properties.

The combination of the experimental and computational methods yields information on the energy landscape that controls mechanical function of biopolymers¹ or conduction through channels^{2,3}, on the force bearing components of proteins⁴, as well as on the underlying physical mechanisms^{5,6}. The new field of cellular mechanics centers around the key role of mechanical forces that arise in cellular processes as input, as output, and as signals, forces playing much the same role as chemicals play in cellular physiology.

This lecture will illustrate cellular mechanics through several examples and introduce the non-equilibrium work relation as a key tool for quantitative analysis. We review the relation, also known as the Jarzynski's equality, and discuss its implications in relation to the second law of thermodynamics and computer simulations. We illustrate how potentials of mean force and thermodynamic potentials can be obtained from SMD calculations .

References

1. Barry Isralewitz, Mu Gao, and Klaus Schulten. Steered molecular dynamics and mechanical functions of proteins. *Curr. Op. Struct. Biol.*, 11:224–230, 2001.
2. Morten Ø. Jensen, Sanghyun Park, Emad Tajkhorshid, and Klaus Schulten. Energetics of glycerol conduction through aquaglyceroporin GlpF. *Proc. Natl. Acad. Sci. USA*, 99:6731–6736, 2002.
3. Paul Grayson, Emad Tajkhorshid, and Klaus Schulten. Mechanisms of selectivity in channels and enzymes studied with interactive molecular dynamics. *Biophys. J.*, 85:36–48, 2003.
4. Mu Gao, David Craig, Olivier Lequin, Iain D. Campbell, Viola Vogel, and Klaus Schulten. Structure and functional significance of mechanically unfolded fibronectin type IIII intermediates. *Proc. Natl. Acad. Sci. USA*, 100:14784–14789, 2003.

^aListed publications are available at <http://www.ks.uiuc.edu/Publications/Papers/SMD/>

5. Sanghyun Park, Fatemeh Khalili-Araghi, Emad Tajkhorshid, and Klaus Schulten. Free energy calculation from steered molecular dynamics simulations using Jarzynski's equality. *J. Chem. Phys.*, 119:3559–3566, 2003.
6. Sanghyun Park and Klaus Schulten. Calculating potentials of mean force from steered molecular dynamics simulations. *J. Chem. Phys.*, 2004. In press.

Already published:

**Modern Methods and Algorithms of Quantum Chemistry -
Proceedings**

Johannes Grotendorst (Editor)
Winterschool, 21 - 25 February 2000, Forschungszentrum Jülich
NIC Series Volume 1
ISBN 3-00-005618-1, February 2000, 562 pages
out of print

**Modern Methods and Algorithms of Quantum Chemistry -
Poster Presentations**

Johannes Grotendorst (Editor)
Winterschool, 21 - 25 February 2000, Forschungszentrum Jülich
NIC Series Volume 2
ISBN 3-00-005746-3, February 2000, 77 pages
out of print

**Modern Methods and Algorithms of Quantum Chemistry -
Proceedings, Second Edition**

Johannes Grotendorst (Editor)
Winterschool, 21 - 25 February 2000, Forschungszentrum Jülich
NIC Series Volume 3
ISBN 3-00-005834-6, December 2000, 638 pages

**Nichtlineare Analyse raum-zeitlicher Aspekte der
hirnelektrischen Aktivität von Epilepsiepatienten**

Jochen Arnold
NIC Series Volume 4
ISBN 3-00-006221-1, September 2000, 120 pages

**Elektron-Elektron-Wechselwirkung in Halbleitern:
Von hochkorrelierten kohärenten Anfangszuständen
zu inkohärentem Transport**

Reinhold Löwenich
NIC Series Volume 5
ISBN 3-00-006329-3, August 2000, 146 pages

**Erkennung von Nichtlinearitäten und
wechselseitigen Abhängigkeiten in Zeitreihen**

Andreas Schmitz
NIC Series Volume 6
ISBN 3-00-007871-1, May 2001, 142 pages

**Multiparadigm Programming with Object-Oriented Languages -
Proceedings**

Kei Davis, Yannis Smaragdakis, Jörg Striegnitz (Editors)
Workshop MPOOL, 18 May 2001, Budapest
NIC Series Volume 7
ISBN 3-00-007968-8, June 2001, 160 pages

**Europhysics Conference on Computational Physics -
Book of Abstracts**

Friedel Hossfeld, Kurt Binder (Editors)
Conference, 5 - 8 September 2001, Aachen
NIC Series Volume 8
ISBN 3-00-008236-0, September 2001, 500 pages

NIC Symposium 2001 - Proceedings

Horst Rollnik, Dietrich Wolf (Editors)
Symposium, 5 - 6 December 2001, Forschungszentrum Jülich
NIC Series Volume 9
ISBN 3-00-009055-X, May 2002, 514 pages

**Quantum Simulations of Complex Many-Body Systems:
From Theory to Algorithms - Lecture Notes**

Johannes Grotendorst, Dominik Marx, Alejandro Muramatsu (Editors)
Winter School, 25 February - 1 March 2002, Rolduc Conference Centre,
Kerkrade, The Netherlands
NIC Series Volume 10
ISBN 3-00-009057-6, February 2002, 548 pages

**Quantum Simulations of Complex Many-Body Systems:
From Theory to Algorithms- Poster Presentations**

Johannes Grotendorst, Dominik Marx, Alejandro Muramatsu (Editors)
Winter School, 25 February - 1 March 2002, Rolduc Conference Centre,
Kerkrade, The Netherlands
NIC Series Volume 11
ISBN 3-00-009058-4, February 2002, 194 pages

**Strongly Disordered Quantum Spin Systems in Low Dimensions:
Numerical Study of Spin Chains, Spin Ladders and
Two-Dimensional Systems**

Yu-cheng Lin
NIC Series Volume 12
ISBN 3-00-009056-8, May 2002, 146 pages

**Multiparadigm Programming with Object-Oriented Languages -
Proceedings**

Jörg Striegnitz, Kei Davis, Yannis Smaragdakis (Editors)
Workshop MPOOL 2002, 11 June 2002, Malaga
NIC Series Volume 13
ISBN 3-00-009099-1, June 2002, 132 pages

**Quantum Simulations of Complex Many-Body Systems:
From Theory to Algorithms - Audio-Visual Lecture Notes**

Johannes Grotendorst, Dominik Marx, Alejandro Muramatsu (Editors)
Winter School, 25 February - 1 March 2002, Rolduc Conference Centre,
Kerkrade, The Netherlands
NIC Series Volume 14
ISBN 3-00-010000-8, November 2002, DVD

Numerical Methods for Limit and Shakedown Analysis

Manfred Staat, Michael Heitzer (Eds.)
NIC Series Volume 15
ISBN 3-00-010001-6, February 2003, 306 pages

**Design and Evaluation of a Bandwidth Broker that Provides
Network Quality of Service for Grid Applications**

Volker Sander
NIC Series Volume 16
ISBN 3-00-010002-4, February 2003, 208 pages

**Automatic Performance Analysis on Parallel Computers with
SMP Nodes**

Felix Wolf
NIC Series Volume 17
ISBN 3-00-010003-2, February 2003, 168 pages

**Haptisches Rendern zum Einpassen von hochaufgelösten
Molekülstrukturdaten in niedrigaufgelöste
Elektronenmikroskopie-Dichteverteilungen**

Stefan Birmanns
NIC Series Volume 18
ISBN 3-00-010004-0, September 2003, 178 pages

Auswirkungen der Virtualisierung auf den IT-Betrieb

Wolfgang Gürich (Editor)
GI Conference, 4 - 5 November 2003, Forschungszentrum Jülich
NIC Series Volume 19
ISBN 3-00-009100-9, October 2003, 126 pages

NIC Symposium 2004

Dietrich Wolf, Gernot Münster, Manfred Kremer (Editors)
Symposium, 17 - 18 February 2004, Forschungszentrum Jülich
NIC Series Volume 20
ISBN 3-00-012372-5

All volumes are available online at <http://www.fz-juelich.de/nic-series/>.