NIC Symposium 2001

edited by
Horst Rollnik
Dietrich Wolf

Symposium, 5. - 6. December 2001
Forschungszentrum Jülich
Proceedings

organized by
John von Neumann Institute for Computing

NIC Series Volume 9

ISBN 3-00-009055-X
Preface

Horst Rollnik
Physikalisches Institut, Universität Bonn
53115 Bonn, Germany

Dietrich Wolf
Institut für Physik, Gerhard-Mercator-Universität Duisburg
47048 Duisburg, Germany

With this book the John von Neumann Institute for Computing (NIC) presents an account of its recent scientific activities. It is addressed both to scientists working in computational science and to the public interested in the advancement of computer based modern science. High performance computing has become an indispensable research tool for a large spectrum of disciplines, a truly cross-disciplinary link. Therefore we hope that this book will become a useful reference for researchers and stimulate the dialogue between the disciplines.

It should be kept in mind that the breathtaking evolution of electronic computing hardware - from the early beginnings 60 years ago to the Tera-Flops machines of today - implies a continuous redefinition of “high performance” computing. A lively exchange between the users of supercomputers about algorithmic developments, ideas, methods and breakthroughs is required in order to stay at the forefront of this technological and scientific development. Therefore this book is the first in a series.

High performance computing is expensive. Therefore scientific progress based on this key technology should not only be counted as successes of the respective disciplines, but should also justify further investments in first class computing environments in Germany. This book should serve both purposes.

In 1987 the three German National Laboratories: Forschungszentrum Jülich (FZJ, at that time KFA Jülich), DESY (Hamburg) and GMD (St. Augustin) founded the Höchstleistungsrechenzentrum (HLRZ) as the first nationwide accessible supercomputing centre in Germany. Forschungszentrum Jülich provided the first supercomputer for this centre, a CRA Y X-MP 48 vector computer. This and the following Cray supercomputers were operated by the Zentralinstitut für Angewandte Mathematik (ZAM) of FZJ. The excellent support that was given by ZAM from the beginning has contributed significantly to the rapid growth of the user community of the supercomputers in Jülich.

In 1998 the John von Neumann Institute for Computing was founded by FZJ and DESY as the successor to the former Höchstleistungsrechenzentrum, and both partners contributed significantly to the computing facilities available at NIC. After German unification new facilities were established in Zeuthen near Berlin, where the APE-computers provided by DESY are operated, while the supercomputers provided by FZJ continue to be operated by the ZAM in Jülich.
The increasing demand for computing resources prompted the foundation of three additional supercomputing centres in Germany. NIC and these three centres have a very close and fruitful cooperation, which is important to coordinate the efficient use of the facilities and to continue the “innovation spiral”. This is necessary to assure international competitiveness of Germany, particularly in view of the national support programs in USA and Japan.

Presently the NIC is equipped with two massively parallel Cray T3E computers with 512 processors each and a Cray SV1ex vector supercomputer in Jülich and with a topical APEmille facility at Zeuthen. The APEmille provides presently a peak performance of ca. 400 Giga-Flops. Thus altogether more than 1.3 Tera-Flops peak performance are available for the users of the NIC. The comparison with the 840 Mega-Flops peak performance available at the beginning of HLRZ 15 years ago illustrates the impressive progress in high performance computing. This year a further step will be taken: A new massively parallel computer with SMP-architecture will be installed in Jülich. Within the next two years it will be expanded to a computing power of 10 Tera-Flops.

From the beginning it has been the objective of NIC to be an institution which supports a nationwide community working in computational science. The users of the NIC computing facilities come from universities and research institutions all over Germany, as well as some international ones, as is documented by the affiliations of the authors of the contributions in this book. In 2001 the peer review board of NIC granted 119 projects from 45 universities and research institutions. Most of the projects are from the fields of condensed matter, material and polymer science and elementary particle physics, but during the last years the fraction of other fields, in particular astrophysics and chemistry, has grown continuously.

NIC not only provides computing resources for scientific research projects, but it also supports two in-house research groups with extensive guest programs, in order to develop and propagate competence and new ideas in high performance computing. These groups work in the fields of “Complex Systems” and “Elementary Particle Physics”. (A third research group is currently being established in the field of biophysics/bioinformatics.) Distinguished researchers are appointed as directors of these groups, currently Prof. Peter Grassberger for the “Complex Systems Group” and Dr. Karl Jansen for the “Elementary Particle Physics Group”. Their reports on the research activities of their groups constitute the first chapter of this book. The remaining chapters contain 38 representative research reports ranging from astrophysics to polymer science. These reports were coordinated by distinguished referees, who also wrote introductions to their respective chapters. They moreover chose one or two contributions per chapter which they regarded as most interesting for a general audience and which were presented by the authors at the first NIC Symposium on the 5th and 6th December 2001 in Jülich.

We thank the referees for helping provide an account of such high quality. Most of them are also members of the NIC peer review board. We also thank the contributors whose work documents the high quality and forward looking nature of computational science and may increase the public interest and support for the key technology of high performance computing. None of the works presented in this book would have been possible without the provision of the high quality computing environments by DESY and the FZJ. In particular the competence and commitment of Prof. F. Höpfel, the director of the ZAM in Jülich,
has been of crucial importance for the NIC from the very beginning, and even before, when the idea of a German supercomputing centre was born. We also thank Dr. Manfred Kremer for coordinating the evaluation procedure and the progress of this book. Without his dedication and timely preparation of decisions, the NIC would not be run so smoothly. Mrs. Beate Herrmann at ZAM was of immense help in the editing process of this book. Dr. Bernd Krahl-Urban and his team from the conference office of FZJ made the organization of the NIC Symposium a real pleasure. Mrs. Helga Frank (NIC) and Mrs. Martina Kamps (NIC) took excellent care of all the organizational details for the conference. We are very grateful to all of them.

Jülich, Easter 2002

D. Wolf          H. Rollnik
Contents

Activities of the NIC Research Groups

Advanced Sequential Monte Carlo Methods in Physics
P. Grassberger 1

Research of a Research Group: The Case of the Elementary Particle Physics Group at NIC
K. Jansen 13

Astrophysics
Introduction by P. L. Biermann 25

Chemodynamical Evolution of Dwarf Irregular Galaxies
G. Hensler, A. Rieschick 27

Supernova Explosions of Massive Stars

Double Black Holes in Galactic Nuclei – Do they Merge or not?
M. Hemsendorf, R. Spurzem, S. Sigurdsson 51

Chemistry
Introduction by S. Peyerimhoff 61

Binding Properties in Protein Nucleic Acids
B. Dietrich, T. A. Hupp, B. Engels 63

Molecules as Mechanical Machines. Genetic Algorithms in the Search of Appropriate Descriptions of Molecular Catalysts
G. Huttner, J. Friedrich, A. Frick, V. Schulz, J. Karas, J. Hunger, St. Beyreuther 73

Efficient Calculation of Rovibrational Energy Levels of General Tetratomic Molecules
M. Mladenović 85

Computersimulation of Free Liquid Surfaces
J. Oberbrodhage, H. Morgner 95

Transition Metal Centers in Biological Matrices: Structure and Function of Vanadate in Vanadium Haloperoxidases and as Phosphate Analog
W. Plass 103
Massively Parallel Individually Selecting Configuration Interaction: A Progress Report
P. Stampfuß, W. Wenzel

Modeling Si/B/N/(C) Ceramic Materials
M. Gastreich, S. Reinhardt, M. Doerr, Ch. M. Marian

Elementary Particle Physics
Introduction by G. Münster

Quarks in a Box
T. Lippert, B. Orth, K. Schilling

Fundamental Parameters of QCD
ALPHA Collaboration, R. Sommer, H. Wittig

Hadronic Structure from Lattice QCD
G. Schierholz

QCD Thermodynamics: The Numerical Study of Strongly Interacting Matter under Extreme Conditions
F. Karsch, E. Laermann, P. Petreczky, S. Stickan, I. Wetzorke

Numerical Simulation of High Quark Densities in QCD with two Colours
I. Montvay, S. Hands, L. Scorzato, J. I. Skullerud

From Enemies to Friends: Chiral Symmetry on the Lattice
P. Hernández, K. Jansen, L. Lellouch

Numerical Simulation of Supersymmetric Yang-Mills Theory
I. Montvay, F. Farchioni, A. Feo, T. Galla, C. Gebert, R. Kirchner, G. Münster, R. Peetz, A. Vladikas

Material Science
Introduction by R. O. Jones

Application of Implicit Density Functionals to 3d Transition Metal Monoxides
R. N. Schmid, E. Engel, R. M. Dreizler

Dislocation Processes and Deformation Twinning in Nanocrystalline Al
V. Yamakov, D. Wolf, S. R. Phillpot, H. Gleiter
Polymers
Introduction by K. Kremer

Phase Transitions in Macromolecular Systems: Monte Carlo and Molecular Dynamics Simulations of Coarse Grained Models
K. Binder, J. Baschnagel, M. Müller, W. Paul

Critical Phenomena and “Crossover Scaling” in Model Systems for Soft Condensed Matter
K. Binder, J. Houdayer, E. Luijten, M. Müller

Computer Modeling of Charged Polymers
Ch. Holm, K. Kremer, M. Deserno, H. J. Limbach

Polyelectrolyte Brushes: A Molecular Dynamics Study
Ch. Seidel

Dissipative Particle Dynamics Simulations of Planar Amphiphilic Bilayers
J. C. Shillcock, R. Lipowsky

Fluid Dynamics
Massively Parallel Direct Numerical Simulation of Turbulent Combustion
M. Lange, J. Warnatz

Environment
Introduction by C. Simmer

Broadening of Convective Cells During Cold Air Outbreaks: A Numerical Study Using a Parallelized Large-Eddy Simulation Model
S. Raasch, M. Schröter

Sensitivity Studies with Numerical Models of Medium Resolution of the Atlantic Ocean
J.-O. Beismann, R. Redler

Modelling of Atmospheric Chemistry-Transport Processes
R. Wolke, O. Knoth, E. Renner, W. Schröder, J. Weickert

Nuclear and Atomic Physics
Introduction by P. Grassberger

Few-Baryon Systems
W. Glöckle, H. Witała, J. Golak, H. Kamada, A. Nogga
Computer Science and Numerical Mathematics

Introduction by M. Griebel

Hash Based Adaptive Parallel Multilevel Methods with Space-Filling Curves
M. Griebel, G. Zumbusch

High Performance FEM Simulation via FEAST and Application to Parallel CFD via FEATFLOW
Ch. Becker, S. H. M. Buijssen, S. Kilian, St. Turek

P. Bastian, K. Johannsen, St. Lang, S. Nägele, Ch. Wieners, V. Reichenberger, G. Wittum, Ch. Wrobel
Advanced Sequential Monte Carlo Methods in Physics

Peter Grassberger

Complex Systems Research Group
John von Neumann Institute for Computing
Forschungszentrum Jülich, 52425 Jülich, Germany

We describe a general strategy for sampling configurations from a given (Gibbs-Boltzmann or other) distribution, not based on the standard Metropolis (Markov chain) strategy. It uses the fact that nontrivial problems in statistical physics are high dimensional and often close to Markovian. Therefore, configurations are built up in many steps, and long range correlations between these steps are often weak. Using biased sampling for each step, each configuration carries its weight which changes at every step. If the bias is close to optimal, all weights are similar and importance sampling is perfect. If not, “population control” is applied by cloning/killing partial configurations with too high/low weight. This is done such that the final (weighted) distribution is unbiased. We apply this method (which is also closely related to diffusion type quantum Monte Carlo) to several problems of polymer statistics, reaction-diffusion models, sequence alignment, and percolation.

1 Introduction

For many statistical physicists, “Monte Carlo” (MC) is synonymous for the Metropolis strategy\(^1\) where one sets up an ergodic Markov process which has the desired Gibbs-Boltzmann distribution as its unique asymptotic state. There exist numerous refinements with more efficient transitions or with modified distributions such as multicanonical MC or simulated tempering. But all of these schemes remain entirely within the framework of the general Metropolis strategy.

On the other hand, stochastic simulations not based on the Metropolis strategy have been used from early times on. Well known examples are evolutionary (in particular genetic) algorithms\(^2\), diffusion type quantum MC simulations\(^3\), and several algorithms devised for the simulation of long chain molecules\(^4\)–\(^8\).

As these methods were developed independently in different communities, it was not generally recognized that most of them are realizations of a common strategy. Indeed, this strategy was first discussed by Herman Kahn in 1956\(^9\) who called it “Russian Roulette and Splitting” and attributed it to unpublished work by von Neumann and Ulam. During the last decade, this strategy has been used intensively in the mathematical statistics community, mainly for Bayesian estimates in time series analysis and similar inference problems\(^10\),\(^11\).

There it is usually called (biased) sequential importance sampling (SIS) with resampling. A thorough error analysis of one variant of this sampling method was given by Aldous and Vazirani\(^12\) who coined the maybe most catchy name for it, “go with the winners” algorithms.

In the present review we shall only discuss applications to physics. But within this limitation we shall cover a very wide range, demonstrating thereby its potential for further applications. More details can be found in\(^13\),\(^14\).
2 The Basic Strategy

2.1 Sequential Importance Sampling

Our aim is to estimate a partition sum of some problem defined by a hamiltonian $H(x)$, by drawing configurations $x$ randomly from some distribution $p(x)$. If $p(x)$ is not uniform but biased, we have to compensate this bias by giving each configuration a weight $\propto 1/p(x)$.

$$Z = \sum_x e^{-\beta H(x)} \approx M^{-1} \sum_{\alpha=1}^M e^{-\beta E(x^\alpha)} / p(x^\alpha) \equiv \hat{Z},$$

leading to a total weight $W(x) = e^{-\beta E(x)}/p(x)$. In importance sampling we try to choose $p(x) \propto e^{-\beta E(x)}$, so that all weights become equal.

We now assume that we can break up the construction of a configuration into $N$ single steps $x_n, n = 1, 2, \ldots N$. For a polymer, the $n$-th step would e.g. be the placement of the $n$-th monomer. Then the weight $W$ is obtained recursively as

$$W_n = W_{n-1} \frac{e^{-\beta(E(x_1,\ldots,x_n)-E(x_1,\ldots,x_{n-1}))}}{p_n(x_n|x_1\ldots x_{n-1})}.$$

In statistics this is called sequential importance sampling (SIS).

In some cases, “natural” values for the $p_n(x_n)$ are easy to guess. In the Rosenbluth method for simulating a self-avoiding walk (SAW), e.g., one chooses uniformly among the free neighbours. But this is not optimal, a better choice for SAWs is provided by markovian anticipation. In general, for choosing $p_n(x_n|x_1\ldots x_{n-1})$ one has to depend on heuristics, except in the case of diffusion quantum MC where perfect importance sampling ($W_N \equiv 1$) is possible if the ground state wave function is known. Specific choices will be discussed together with applications.

2.2 Population Control

The main drawback of SIS is that the distribution of weights can become extremely wide, if the $p_n(x_n)$ are not perfectly chosen. If long range correlations are weak (as e.g. for SAWs), $\log W_N$ is roughly a sum of independent terms. This suggests to use some “resampling” strategy whose effect will be to eliminate ‘bad’ configurations already at an early stage, and to replace them by clones of ‘good’ ones, similar to the competition in an evolutionary algorithm. In contrast to the latter, we do this however in a well controlled way which leaves unchanged the estimator of the partition sum. There is a wide range of possible such resampling methods discussed in the literature, we shall concentrate on the following:

If at any step $n$ the weight $W_n$ is above a suitably chosen threshold $W^+_n$, we make an additional copy of the configuration $x_1, \ldots x_n$, and give both copies the weight $W_n/2$. Both are then grown independently (with eventual later copyings) up to full length $a$. In this way high weights are suppressed and precious “good” configurations are less likely to be lost entirely by bad subsequent moves. In Ref. a similar strategy (but not based on weights) was called ‘enrichment’.

Footnote: In some cases (e.g. at low temperatures, where Boltzmann factors are huge), it might be necessary to make several copies and to distribute the weight evenly among them.
On the other hand, if $W_n^+$ falls below another threshold $W_n^-$, we draw a random number $r \in [0, 1]$. If $r < 1/2$ we kill the configuration and start a new one. If $r > 1/2$ we keep it and double its weight.

Obviously, for any choice of the thresholds, neither the cloning nor the pruning introduce any additional bias. Thus we can, in principle, use any choice for $W_n^+$ and $W_n^-$, and we can change them ad libitum during the simulation. Bad choices will, however, lead to inefficiency, just as do bad choices for $p_n(x)$. Usually, a good choice is to take $W_n^+$ and $W_n^-$ proportional to the current estimate $\hat{Z}_n$ of the partition sum. Details and tricks to avoid excessive cloning at very low temperatures where weights can become huge due to the large Boltzmann factors can be found in [8, 16, 17].

### 2.3 Depth First Versus Breadth First

As described above, the algorithm is most efficiently implemented in a depth first fashion, and as such was called PERM (pruned-enriched Roenbluth method) in [8]. In a depth first approach [18], we follow one copy until its end before we take up the other copy. In breadth first search, on the other hand, we treat all copies in parallel. We first handle the $n$-th steps of all copies before we go to $n + 1$.

Evolutionary algorithms are usually implemented breadth first. One puts up a population of $M$ replicas which are evolved simultaneously, and population control is exercised such that $M$ stays constant during the evolution. The same is true for most implementations of the SIS strategy. This has several advantages:

- Breadth first approaches are well adapted for massively parallel computers. One simply puts one configuration on each processor.
- One has no problem with keeping the number of replicas constant.
- One can use more general population control strategies [11].

But the last two points seem minor in most application we have studied. On the other hand, the main advantage of depth first is the elegance and efficiency of the codes. The most easy implementation is by means of recursion (for a pseudocode see [9]). Copies (or rather instructions to make copies) are then put on a stack which is maintained automatically if recursive function calls are used. Storage use is minimized, in particular for non-Markovian problems where the entire configuration has to be kept in memory anyhow (as only a single copy and its history is kept in memory), and communications are also less than in breadth first.

### 3 A Lamb in Front of a Pride of Lions

The first example is a very idealized problem from population dynamics [19, 20]: Consider a ‘lamb’, represented by a random walker on a 1-dimensional lattice with discrete time and hopping rate $\lambda$ per time unit, leading to a diffusion constant $D_{\text{lamb}}$. It starts at time $t = 0$ at $x = 0$. Together with it, there start also $N$ ‘lions’, $n_L$ of them at $x_i = -2$ ($i = 1, \ldots, n_L$) and $n_R = N - n_L$ at $x_i = +2$. They also perform random walks, but with a diffusion
constant $D_{\text{lion}}$ which may differ from $D_{\text{lamb}}$. Lions don’t interact with each other. But if a lion and the lamb meet at the same site, the lamb is eaten and the process is finished.

For all $(n_L, n_R)$, one can prove\textsuperscript{19} that the survival probability decays for large $t$ as

$$P_{n_L,n_R} \sim t^{-\alpha_{n_L,n_R}},$$

(3)

For $N \leq 2$ the decay exponents are known exactly. They depend continuously on $D_{\text{lamb}}/D_{\text{lion}}$ for $N = 2$. For $N > 2$ no exact result is known and one needs simulations\textsuperscript{19}.

Using PERM, such simulations were performed for $D_{\text{lamb}} = D_{\text{lion}} = 1$. The jumping probabilities $p_l$ (left) and $p_r$ (right) were essentially unbiased except for the closest lamb-lion pairs which preferentially jumped away from each other. Decay exponents for $n_L = N, n_R = 0$ are shown in the left panel of Fig.1, for $n_L = n_R = N/2$ they are shown in the right panel. To estimate them, we had to measure probabilities down to $\approx 10^{-60}$ which is possible only due to resampling.

The lower straight lines in Fig.1 are predictions based on the observation that fluctuations should become small for $N \to \infty$ and $t \to \infty$\textsuperscript{20}. They are obviously far off the MC points, but the latter show a slight curvature which is indeed significant, suggesting that the MC curves will asymptotically be indeed parallel to them. Similarly huge corrections to the asymptotic behaviour were also found in the reaction $A + B \to B$ with uniformly distributed diffusing $B$ particles\textsuperscript{21}, and in the Donsker-Varadhan\textsuperscript{22} problem where the $B$ particles are stationary traps\textsuperscript{23}. Also in these cases, sampling from probabilities as low as $10^{-30} - 10^{-100}$ with high accuracy was crucial and could hardly have been done with conventional methods.

4 Multiple Spanning Percolation Clusters

Let us now consider percolation on a large but finite rectangular lattice in $2 \leq d < 6$. We single out one direction as “spanning”. In this direction boundaries are open, while periodic b.c. are used in the other direction(s). For a long time it was believed that there is at most one spanning cluster (which touches both open boundaries) in the limit of large lattices, keeping the aspect ratio fixed ($L_i = x_i L, L \to \infty, i = 1, \ldots d$).
Since there is no spanning cluster for subcritical percolation and exactly one in supercritical, the only relevant case is critical percolation. There it is now known that the probabilities $P_k$ to have exactly $k$ spanning clusters are all non-zero in the limit $L \to \infty$. In $d = 2$ they are known exactly from conformal invariance, but for $d \geq 3$ no exact results are known. But there is a conjecture by Aizenman\textsuperscript{24}, stating that for a lattice of size $L \times \ldots \times L \times (rL)$ ($rL$ is the length in the spanning direction) $P_k \sim e^{-\alpha r}$ with

$$\alpha \propto k^{d/(d-1)} \quad \text{for} \quad k \gg 1. \quad (4)$$

For $d = 2$ one has $\alpha \sim k^2$, in agreement with Eq.(4). A generalization in $d = 2$ consists in demanding that clusters are separated by more than one lattice unit. More precisely, we demand that there are at least $q$ disjoint spanning paths on the dual lattice\textsuperscript{25}. In that case, and for periodic transverse b.c.,

$$\alpha = \frac{2\pi}{12} \left[ ((q + 1)k)^2 - 1 \right] . \quad k \geq 2, d = 2 \quad (5)$$

In order to test Eqs. (4),(5) for a wide range of values of $k$ and $r$, one has to simulate events with tiny probabilities, $\ln P_k \sim -10^2$ to $-10^3$. It is thus not surprising that previous
To demonstrate that such rare events can be simulated with PERM, we show in Fig. 2 a lattice of size $500 \times 900$ with 5 spanning clusters which keep distances $\geq 2$. Eq.(5) predicts for it $P_k = \exp(-336\pi/5) \approx 10^{-92}$. This configuration was obtained by letting 5 clusters grow simultaneously, using a standard cluster growth algorithm\textsuperscript{28}, from the left border. Precautions were taken that they grow with the same speed towards the right, i.e. if one of them lagged behind, the growth of the others was stopped until the lagging cluster had caught up. If one of them died, or if two came closer than two lattice units, the entire configuration was discarded. If not, it was cloned if the weight $W_n$ exceeded $3\hat{Z}_n$. Note that here the growth was made without bias, and therefore no pruning was necessary.

In this way we could check Eq.(5) with high precision, proving the correctness of our algorithm.

More interesting is the test of Eq. (4) for $d = 3$. Values of $\alpha$ extrapolated from various lattices up to $128 \times 128 \times 2000$ are shown in Fig. 3\textsuperscript{29}. They are based on estimated probabilities as small as $10^{-300}$! The dashed line there is a fit $\alpha = 2.76(k^2 - 0.61)^{3/4}$ which is in perfect agreement with Eq.(4).

5 Polymers

One of the main applications of the go-with-the-winners strategy is configurational statistics of long polymer chains. For further polymer problems not discussed below see\textsuperscript{16, 35, 15, 36–40}. For a breadth first algorithm which otherwise is very similar to PERM see\textsuperscript{7}.

5.1 $\Theta$-Polymers

PERM is particularly efficient near the so-called ‘theta-’ or coil-globule transition. In a bad solvent, a single polymer forms a collapsed globule. If temperature is raised the entropy gained by opening wins at $T > T_\theta$, and a SAW like coil is reached. According to the generally accepted scenario, the theta-point is tricritical with upper critical dimension $d_c = 3\textsuperscript{30}$. At $T_\theta$, Rosenbluth bias correction and Boltzmann factors nearly cancel in $d = 3$. Therefore, polymers have essentially random walk configurations with small (logarithmic) corrections. Therefore, an non-reversing random walk (U-turns are forbidden) for SIS is already sufficient to give good statistics with very few pruning and enrichment events. In Ref.\textsuperscript{8} chains made of up to 1,000,000 steps could be sampled with high statistics within modest CPU time. They were done in finite volumes (“dense limit”) and verified that the $\Theta$-point indeed is a second order transition in the sense that the monomer density inside a large globule approaches zero when $T \rightarrow T_\theta$ from below. The most precise verification of logarithmic corrections came from chains with $N = 10,000$ in infinite volume. The deviations from random walk behaviour turned out to be much stronger than the leading-log corrections predicted from the renormalization group\textsuperscript{31}, but agreement improves substantially when higher order corrections are included in the latter\textsuperscript{32}. 

6
5.2 Critical Unmixing

A related problem is the unmixing of semidilute polymer solutions. For any finite chain length $N$ this is in the Ising universality class. But in addition to the Ising scaling laws, there are further universal scaling laws for parameters and amplitudes which, from the Ising point of view, would be non-universal. In particular, the critical temperature should approach $T_\theta$ when $N \to \infty$, $T_c - T_\theta \sim N^{-1/2}$, and the critical monomer concentration should tend to zero,

$$\phi_c \sim N^{-1/2}.$$  \hspace{1cm} (6)

The exponents here are mean field, appropriate for $d = 3$. Indeed one should also expect logarithmic corrections\textsuperscript{31}. Previous experiments had suggested an exponent $0.38 \pm 0.01$ in Eq.(6). This would be very hard to understand and has stirred a lot of theoretical activity (for a review see\textsuperscript{41,33}). Simulations using PERM\textsuperscript{33} showed that this is wrong: The deviations from Eq.(6) can be understood most easily as logarithmic corrections.

5.3 DNA Melting

DNA in physiological conditions forms a double helix. Changing the pH value or increasing $T$ can break the hydrogen bonds between the base pairs, and a phase transition to an open coil occurs. Experiments suggest it to be first order\textsuperscript{42}. While a second order transition would be easy to explain\textsuperscript{43,44}, no previous model had been able to give a first order transition.

The model studied in\textsuperscript{34} lives on a simple cubic lattice. A double strand of DNA with length $N$ is described by a diblock copolymer of length $2N$, made of $N$ monomers of type $A$ and $N$ monomers of type $B$. All monomers have excluded volume interactions, i.e. two monomers cannot occupy the same lattice site, with one exception: The $k$-th $A$-monomer and the $k$-th $B$-monomer, with $k$ being counted from the center where both strands are joint together, can occupy the same site. If they do so, then they even gain an energy $-\epsilon$. This models the binding of complementary bases.

![Figure 4](image-url)

Figure 4. Histograms of the number of contacts, for single strand length $N = 500, \ldots, 3000$, at $\epsilon = \epsilon_c$. On the horizontal axis is plotted $n/N$ as is appropriate for a first order transition.
The surprising result of simulations of chains with $N$ up to 4000 is that the transition is first order, but shows finite scaling behaviour as expected for a second order transition with cross-over exponent $\phi = 1$. To demonstrate this, we show in Fig. 4 energy histograms for different chain lengths. One sees two maxima, one at $n = 0$ and the other at $n \approx N/2$, whose distance scales proportionally to $N$. But in contrast to usual first order transitions the minimum in between does not deepen with increasing $N$. This is due to the absence of any analogon to a surface tension. The same conclusion is obtained from specific heat data and from end-to-end distances.\(^{34}\)

In Ref.\(^{34}\) we studied also similar models with (partially) switched off excluded volume effects. They show that excluded volume is the main force making the transition first order, as also confirmed by subsequent analytic calculations.\(^{45}\)

5.4 Native Configurations of Toy Proteins

Predicting the native ($\approx$ ground) states of proteins is one of the most challenging problems in mathematical biology.\(^{46}\) It is difficult because of the many local energy minima.

In view of this, there exists a large literature on finding ground states of artificially constructed heteropolymers. Most of these models are formulated on a (square or simple cubic) lattice and use only few monomer types. The best known example is the HP model if K. Dill\(^{47}\) which has two types of amino acids: hyrophobic (H) and hyrophilic (polar, P) ones. With most algorithms, one can find ground states typically for random chains of length up to $\sim 50$.

In Ref.\(^{17}\) we have used PERM to study several sequences, of the HP model and of similar models, which had been discussed previously by other authors. In all cases we found the known lowest energy states, but in several cases we found new ones. A particularly impressive example is a chain of length 80 with two types of monomers, constructed such that it should fold into a bundle of four “helices” with an energy $-94$. Even with a specially designed algorithm, the authors of\(^{48}\) were not able to recover this state. With PERM we not only found it easily, we also found several lower states, the lowest one having energy $-98$ and a completely different structure.

6 Lattice Animals (Randomly Branched Polymers)

Consider the set of all connected clusters of $n$ sites on a regular lattice, with the origin being one of these sites, and with a weight defined on each cluster. Lattice animals are defined by giving the same weight to each cluster. This distinguishes them from percolation clusters where the weight depends on the ‘wetting’ probability $p$. In the limit $p \rightarrow 0$ this difference disappears, and the two statistics coincide. It is believed that lattice animals are a good model for randomly branched polymers.\(^{49}\) While there existed no efficient algorithm for estimating the animal partition sum there exist very simple and efficient Leath-type\(^{28}\) algorithms for percolation clusters.

Our PERM strategy\(^{14,50}\) consists in starting off to generate subcritical percolation clusters by a (breadth first) Leath method, re-weighing them as animals while they are still growing, and in making clones of “good” ones. Since we work at $p < p_c$, we do not need pruning. The threshold $W_+$ for cloning is chosen such that it depends both on the present animal weight and on the anticipated success for further growth.
In this way we obtained good statistics for animals of several thousand sites, independent of the dimension of the lattice. A typical 2-d animal with 8000 sites is shown in Fig. 5. We also simulated animal collapse (when each nearest neighbor pair contributes \(-\epsilon\) to the energy), and animals near an adsorbing surface\(^{50}\).

7 Conclusion

We have seen that MC simulations not following the Metropolis scheme can be very efficient. We have illustrated this with a wide range of problems. Conspicuously, the Ising model was not among them. It simply would be very hard to beat, say, the Swendsen-Wang algorithm. In principle, the go-with-the-winners strategy has as wide a range of applications as the Metropolis scheme. Its only requirement is that instances (configurations, histories, ...) are built up in small steps, and that the growth of their weights during the early steps of this build-up is not too misleading.

The method is not new. It has its roots in algorithms which are regularly used since several decades. Some of them, like genetic algorithms, are familiar to most scientists, but it is in general not well appreciated that they can be made into a general purpose tool. And it seems even less appreciated how closely related are methods developed for quantum MC simulations, polymer simulations, and optimization. I firmly believe that this close relationship can be made use of in many more applications to come.

Among these are significance tests for sequence alignment, where one needs large samples of random pairs of sequences in order to check whether an observed alignment is significant. Instead of really draw random pairs, one can use PERM to draw biased pairs which are more similar than random ones, enhancing thereby the interesting high-score region\(^{51}\).

Another application is to epidemic models where one can follow the fate of epidemics which have a very low chance of survival since, e.g., they started in a very hostile environment which they first have to adopt to their needs. Here simulations with PERM\(^{52}\) allowed to verify with very high statistics the claim of\(^{53}\) that no power laws result, in contrast to previous suggestions.

I am indebted to all my collaborators, in particular to U. Bastolla, G.T. Barkema, M.S. Causo, B. Coluzzi, H. Frauenkron, E. Gerstner, H.-P. Hsu, V. Mehra, W. Nadler and L. Schäfer.
References

50. P. Grassberger and W. Nadler, to be published.
51. R. Bundschuh and P. Grassberger, to be published.
52. P. Grassberger, to be published.
We give a short overview of the activities of the NIC research group elementary particle physics. We then concentrate on two conceptual advances in lattice field theory, the improvement of discretization errors and the usage of finite size effects. An example, structure functions on the lattice, will serve to demonstrate how these concepts work in practical applications. We finally report about the status of the APE machines and the planned installations.

1 Introduction

Elementary particle physics covers a broad range of phenomena like spontaneous symmetry breaking, dynamical mass generation, confinement, asymptotic freedom and phase transitions to name only a few. The physics of these phenomena is often of inherent non-perturbative nature. In such cases the approach of lattice field theory can help to understand physical phenomena better and test model predictions. In the lattice approach our usual space-time is made discrete and a non-vanishing value of a lattice spacing $a$ is introduced. In this way physics problems can be made understandable in form of a program code to a computer. This allows for first principle calculations of physical observables having as only input a theoretical model. Of course, the discreteness of the space-time structure is only an approximation to the real world and eventually this systematic error has to be removed through a well-controlled continuum limit where the lattice spacing $a$ is sent to zero.

Lattice field theory has reached a quite mature stage. Many areas of physics are addressed by this approach ranging from 2-dimensional spin systems, the standard model of elementary particle interactions and even investigations of quantum gravity (see the proceedings of the annual lattice symposia). Still, the main activity of the lattice physicists community is the understanding of quantum chromodynamics (QCD), i.e. our model for the strong interactions. This is also reflected in the activities of the NIC research group elementary particle physics where various aspects of QCD are investigated. Topics are

- determination of fundamental parameters of QCD like the coupling strength and quark masses
- the recent development of chiral invariant formulations of QCD
- structure functions and hadron spectrum of QCD
- confinement, monopoles and topology
- high precision studies of lower dimensional spin models
• finite temperature phase transitions

• improvement of algorithms for dynamical quark and chiral invariant fermion simulation

• participation in the array processor experiment (APE), leading to the development of APE computers that are dedicated to QCD applications

It is clear that in this contribution not all the above topics can be covered. I will rather concentrate on a particular example that demonstrates how two new concepts in lattice field theory are incorporated in practical applications. I refer to a few contributions to this symposium where also other topics of the above list are addressed. Besides the author himself, at the time of this write-up the team that works in the NIC research group on the topics above consist of S. Capitani (Postdoc), T. Chiarappa (Ph.D. student), M. Hasenbusch (Postdoc), C. Hoelbling (Postdoc), R. Horsley (Postdoc), T. Kovacs (EU fellow) D. Pleiter (Postdoc), G. Schierholz (permanent member), C. Urbach (Diploma student) and I. Wetzorke (Postdoc).

2 Scary Animals on the Lattice

When working in lattice field theory, two rather scary animals are met: the first is the non-vanishing lattice spacing, the second is the finite volume. If we keep the physical size of the box fixed, say $L = 1\text{fm}$, corresponding to the diameter of the proton, then this physical length is realized by putting $N$ points for a given value of the lattice spacing $a$ such that $L = N \cdot a$. If we start with $a = 0.1\text{fm}$, we would have $N = 10$. Now, if we want to reach the continuum limit, the value of $a$ has to be decreased while $L$ is to be kept fixed. Hence, for $a = 0.01\text{fm}$ we would need $N = 100$, for $a = 0.001\text{fm}$, $N = 1000$ and so on. Since we are working in four space-time dimensions, the number of lattice points to be kept in the computer grows like $N^4$ in this so-called “naive” continuum limit. Clearly, one runs easily out of computer resources, with respect to available computer power as well as memory.

What is to do? The idea of keeping the value of the lattice spacing finite is not a very good one: severe artifacts of the non-vanishing lattice spacing are detected leading to systematic errors in physical observables that are difficult to control. Also, keeping the physical length smaller than 1fm gives rise to large finite size effects inducing again systematic errors that are hard to control. We seem to end up with a real dilemma forcing us to go the brute force way and just increase the number of lattice points while decreasing $a$.

However, modern conceptual advances have paved the way to partly circumvent these problems. We nowadays have the tool of the improvement programme to at least diminish the effects of a non-vanishing lattice spacing. In addition, we have learnt to actually make use of finite size effects to extract physical information of the infinitely large system. For an example see another paper. In the following we will discuss these concepts a little more detailed (more extended reviews can be found in).
3 Improvement Programme

When physical observables are computed in lattice QCD it is noticed that the dependence of these observables may depend rather strongly on the lattice spacing, rendering any continuum extrapolation difficult and costly. The reason for this problem is seen directly from the action of lattice QCD. The standard form is

$$S_{\text{old}} = S_G + S_{\text{wilson}}$$

where $S_G$ is the pure gluonic part of the action while $S_{\text{wilson}}$ is the fermionic part. As indicated in eq. (1), in contrast to the pure gluonic part of the action, the fermions induce a linear dependence on the lattice spacing, resulting in sizable lattice artifacts that are problematic to extrapolate away.

A (part) solution is to add new terms to the action. If all possible symmetries on the lattice and the equations of motion are used, it is found that there is only one more term that can appear at $O(a^2)$. The action then reads

$$S_{\text{new}} = a^5 \sum_x c_{\text{sw}} \bar{\Psi}(x) \frac{i}{4} \hat{F}_{\mu\nu}(x) \Psi(x).$$

The particular form of the new term is not of interest here. What is important, however, is the fact that the parameter $c_{\text{sw}}$ is freely tunable. In particular, $c_{\text{sw}}$ can be tuned such that the additional term in eq. (2) exactly cancels the $O(a)$ effects from the original action, eq. (1), thus leaving only errors of $O(a^2)$ behind.
Figure 2. A typical collision of two hadrons.

Figure 1 shows how this procedure works in practice. A particular physical observable (a step scaling function, see next section) is shown whose physics interpretation is not of importance here. We only remark that first of all this quantity has a well defined continuum limit. Second, we see that the evaluation with the standard fermion action, eq. (1) (dashed line), leads to large lattice artifacts while the use of the improved action (solid line), eq. (2), seems to eliminate these artifacts almost completely. In this way, an acceleration to the continuum limit is achieved and the systematic errors coming from discretization errors are much better controlled.

4 Finite Size Effects and Structure Functions

Let us now come to the second dangerous animal on the lattice, finite size effects. While in the case of discretization errors we could only achieve an improvement in reaching the continuum limit, the situation with finite size effects is completely different: here we “turn around the table” and instead of trying to eliminate finite size effects, we are actually going to use them to extract physical information.

To illustrate this, we will discuss the important field of structure functions. Let us consider a typical particle collision as performed at large accelerators, e.g. at DESY, see figure 2.

Here a hadron A hits a hadron B, a jet is formed and some remnants are left over. We denote by \( x_A \) a fraction of the momentum of hadron A (and similar of the hadron B). Now, what we want to know is the internal structure of the hadron which is believed to consist of quarks and gluons. Our aim is to know the probability \( f_a(x_A) \) to find a quark \( a \) in the hadron that carries a momentum fraction \( x_A \) of the hadron A. Similarly, we can ask for the probability function of the quark \( b \) and, of course, of the gluons. The complete knowledge of the probability functions \( f(x) \) –also called structure functions– would give us the desired information about all the internal constituents of the hadron.

Unfortunately, the structure functions are very complicated to compute on the lattice.
However, in experiment, not only the structure functions themselves but also their moments can be determined. The $n$th moment is defined by

$$M^{(n)}_a = \int_0^1 dx x^{n-1} f_a(x) , \ n = 1, 2, \cdots .$$  \(3\)

For example, the 2nd moment corresponds to the average momentum $\langle x \rangle$ that is carried by the quark. Knowing all the moments, the complete structure function maybe reconstructed. In practice, it is presumably only possible to compute a few moments on the lattice. Estimates indicate, however, that even with $n \leq 3$ a rather good approximation of the full structure function can be achieved\(^{24}\). The advantage of considering moments of structure functions is that these are related to local operators that can be computed with lattice techniques using standard methods\(^8\).

The scattering process of figure 2 is a quantum process. This means that the simple picture of figure 2 is not what actually occurs in nature. Rather, many quantum fluctuations disturb the picture and a steady generation and annihilation of virtual quark-antiquark pairs and gluons take place. This leads to the phenomenon that the moments depend on the energy scale $\mu$ at which the collision process is observed since the probability of generating virtual quarks and gluons depends on the energy. Therefore,

$$M^{(n)}_a = M^{(n)}_a(\mu) .$$  \(4\)

In quantum field theory we speak of a *scale dependent renormalization* of the moment.

In practice this means that we have to evaluate a moment not as a single number but as a function of the energy scale $\mu$. The energies we have to consider reach from the inverse W-mass to the confinement scale, while at the same time we should eliminate the discretization and the finite volume errors:

$$a \ll \frac{0.001 \text{fm}}{1/M_W} \ll \frac{1 \text{fm}}{\text{conf.scale}} \ll L .$$  \(5\)

Clearly, covering all the distances of eq. (5) in a single lattice would lead to a totally unrealistic number of lattice points.

The way out of this problem comes in two steps. In the first step we identify the scale with the inverse box length, $\mu = 1/L$, thus using the finite box length as a probe for the physics taking place at this scale. In the second step, we break the problem into many smaller steps each of which can be evaluated with moderate computer resources.

Let us be a bit more concrete and focus on the expectation value of the average momentum, $\langle x \rangle(\mu_0)$, assuming that we know this expectation value at a certain energy scale $\mu_0$. This scale will be chosen such that $\langle x \rangle(\mu_0)$ is easily accessible for lattice computations. Then the expectation value at a different scale $\mu$ is obtained by applying the so-called step scaling function $\sigma(\mu/\mu_0)$, i.e.

$$\langle x \rangle(\mu) = \sigma(\mu/\mu_0) \cdot \langle x \rangle(\mu_0) .$$  \(6\)

The step scaling function describes the full scale evolution of the a scale dependent quantity of interest. Note that different observables need a different step scaling function. A very important property of the step scaling function is that it can be split into several steps. Identifying the scale $\mu = 1/L$, this means that (choosing steps of size 2)

$$\sigma(L_0/L) = \sigma(2L_0/L_0) \cdot \sigma(4L_0/2L_0) \cdot \cdots \cdot \sigma(2L/L) .$$  \(7\)
The aim in the end is to have the step scaling function, eq.(7), and the running of $\langle x \rangle (\mu)$ non-perturbatively. It is exactly at this point that we depart from our continuous space-time structure and introduce a discrete lattice with non-vanishing lattice spacing $a$ allowing us to resort to numerical simulations. Of course, in the end the discretization error has to be removed in a well-controlled continuum limit. As an important side remark we note that $\sigma(2L/L)$ has indeed a well-defined continuum limit such that the whole procedure makes sense.

On the lattice we introduce the lattice step scaling function

$$\Sigma(a) = \sigma(2L/L, a).$$

$\Sigma(a)$ is the quantity that we want to (and can) compute with lattice methods. The contin-
uum step scaling function is then reached through

$$\sigma = \lim_{a \to 0} \left. \Sigma(a) \right|_{\mu^{-1} = L \text{ fixed}}.$$  \hspace{1cm} (9)

An important remark is that in the \( a \to 0 \) limit the physical scale \( L \) is kept fixed. This can be achieved by fixing the value of the renormalized gauge coupling \( \bar{g} \).

We show in figure 3 the approach to the continuum limit of \( \Sigma(a) \), eq.(8), for two ways of discretizing fermions on the lattice, once with standard fermions (dashed lines) using the action of eq.(1) and the non-perturbatively improved action (full lines) of eq. (2). It is reassuring that both \( \Sigma \)'s extrapolate to the same continuum step scaling function showing the universality of the continuum limit. The individual graphs are obtained at fixed values of the running coupling \( \bar{g} \) corresponding to a fixed scale \( 1/L \).

Figure 3 demonstrates that the approach to the continuum limit of the lattice step scaling function is well controlled. At this point we then have the behaviour of the non-perturbatively evaluated step scaling function \textit{in the continuum} and can safely forget that we ever worked on the lattice. In a very similar fashion, also matrix elements themselves, like \( \langle x \rangle(\mu_0) \) can be computed at a reference scale \( \mu_0 \) in the continuum limit (see ref.8). What we have achieved then are results in the continuum theory from fully non-perturbative calculations – which is nothing else but the goal of lattice calculations. The price we have to pay for such a result is besides having to use powerful computers – that we end up with statistical errors and not with an analytical prediction. But, with increasing computer power these errors can eventually be rendered as small as required by the precision of the experimental data.

Let us discuss shortly, the result of the above described computation for the average momentum in a pion at a scale of \( \mu = 2.4 \)GeV. One result that is interesting for lattice physicists is that there can be rather large discretization errors. Changing the lattice spacing from about \( a = 0.1 \)fm to zero, we find

$$\langle x \rangle(a = 0.093 \text{fm}) = 0.30 \to \langle x \rangle(a = 0) = 0.20.$$  \hspace{1cm} (10)

This means that discretization errors can lead to about 50% systematic errors and demonstrates that a controlled continuum limit is mandatory. Another result that is interesting for the theoretical physicist is that the effects of a non-perturbative renormalization procedure can be at the order of 10% to 15% as compared to perturbative results.

Finally, the result that interests everybody is the comparison to experiment:

$$\langle x \rangle^{\text{experiment}}(\mu = 2.4 \text{GeV}) = 0.23(2)$$

$$\langle x \rangle^{\text{quenched}}_{\text{MS}}(\mu = 2.4 \text{GeV}) = 0.30(3).$$  \hspace{1cm} (11)

We indicated in eq. (11) that the theoretical values are obtained in the commonly used \textsc{ms} scheme and in the quenched approximation, where the internal generation of (virtual) quarks inside the hadron is neglected. The still somewhat large error of both the experimental and the theoretical results saves us from really claiming a discrepancy. Of course, we do not have a very safe control about the systematic error of using the quenched approximation and it will be very interesting to repeat the calculation presented here in the full theory with dynamical quarks. As a final comment we mention that the main result of a computation outlined above are \textit{renormalization group invariant} (RGI) quantities. The reason is that RGI quantities can be used directly in other regularization schemes used e.g.
in perturbation theory. We give the definition of the RGI matrix element and the result from the lattice in appendix A for the interested reader.

5 Machines

Concepts like the improvement of discretization errors or the usage of finite size effects as described above are certainly very important. Nevertheless, without powerful supercomputers such physics projects would not be able to perform. Now, the numerical problem that is in the heart of applications in QCD is fortunately rather simple: roughly—and somewhat over-simplified—speaking one needs mainly complex normal operations, $a + b \ast c$, and a communication between only nearest neighbours on the lattice. This simple structure makes it possible to design and develop special purpose machines that are highly efficient for lattice QCD calculations.

The NIC research group is involved in these developments, in particular in the APE (Array Processor Experiment) machines. APE computers are custom made and have a long history already with a first machine installed in Italy around 1985. The first machine used in the high energy physics community in Germany was the APE100. This machine appeared to be extremely successful. It is a massively parallel SIMD machine with a fast interconnecting network. APE100 ran very stable in practice and became the workhorse for many groups performing lattice gauge theory simulations.

The APEmille computer is the successor of APE100 and hence already the third generation of APE machines. It is based on a 3-dimensional mesh of nodes connected by a synchronous communication network, linking nearest neighbours of nodes. We give its specification in table 1.

<table>
<thead>
<tr>
<th>Specification</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peak performance</td>
<td>528 MFlops/proc</td>
</tr>
<tr>
<td>Clock frequency</td>
<td>66 MHz</td>
</tr>
<tr>
<td>FP registers</td>
<td>512 (32-bit)</td>
</tr>
<tr>
<td>Data memory</td>
<td>32 MByte/proc</td>
</tr>
<tr>
<td>Communication BW</td>
<td>66 MByte/s/direction</td>
</tr>
<tr>
<td>I/O BW per master</td>
<td>6 MByte/s</td>
</tr>
<tr>
<td>Power consumption</td>
<td>28 W/GFlops</td>
</tr>
<tr>
<td>Price</td>
<td>2.5 Euro/MFlops peak</td>
</tr>
</tbody>
</table>

Table 1. Key parameters of APEmille.

Power consumption of APEmille systems is very low (less than 30 W/GFlops) and the footprint of a two-crate rack is about $0.7 m^2$. For these reasons, APEmille machines are simply air cooled and do not need complex infrastructure.

At present, several APEmille installations exist Europe wide at various places (Rome I, II, DESY/NIC Zeuthen, Pisa, Milano, Bari, Paris Sud, Bielefeld, Swansea and INFN-LNGS). In the near future all planned installations will be finished and an integrated peak performance of about 2 Teraflops will be reached. Typical applications reach an efficiency of 32% - 47% depending on the distribution of the lattice onto the mesh of nodes. The stability of the existing APEmille platforms as is experienced today strongly indicates that
APEmille has the potential to become the workhorse to the lattice physicists in Europe in the next years – as APE100 was in the past.

But, this is not the end of the story. A follow-up machine, apeNEXT is being envisaged and actually already quite far in its developing stage. Major differences to earlier APE machines are that apeNEXT runs in 64-bit precision while its predecessors had only 32-bit words. It is a SPMD machine and runs asynchronously giving new challenges to the APE collaboration. Details of the architecture can be found in\textsuperscript{15} where also the present status of apeNEXT is described. The final installations (intended for 2003) will achieve 2-3Tflops for a stand alone system with price/performance ratio of 0.5Euro/Mflop (peak). These parameters will then meet the requirements formulated by an ECFA panel\textsuperscript{25} for the performance needs of lattice field theory in the next years.

6 Conclusion

Lattice calculations benefit from

- conceptual theoretical developments like improving discretization errors and using finite size effects,
- improvements of the algorithms employed,
- further and further increase of available computer power as it is provided by centers like NIC, housing commercial supercomputers and custom made machines like APE.

It is the interplay between all these three areas that furthers the precision of lattice field theory computations to a stage where direct contact to experiment will be achieved. If the progress in lattice field theory we have observed in the last years can be maintained in the coming years, lattice field theory will certainly be a major player in analyzing and interpreting experimental data coming from future accelerators. The NIC research group works in all the three areas of theoretical, algorithmic and machine developments and we hope that we can contribute to the ambitious aim mentioned above.

Appendix A

Let us discuss here a theoretical point that is meant to be for the experts. The step scaling function of the main text is –in the perturbative regime– of course nothing else but the Callen-Symanzik function. This allows to define renormalization group invariant quantities if the scale $\mu$ is sent to large enough values

$$\langle x \rangle^\text{ren}_{\text{INV}} = O^\text{ren}_{\text{SF}}(\mu) \cdot f^\text{SF}(\tilde{g}^2(\mu))$$

with

$$f^\text{SF}(\tilde{g}^2(\mu)) = (\tilde{g}^2(\mu))^{-\gamma_0/2b_0} \exp \left\{ - \int_0^{\tilde{g}(\mu)} dg \left[ \frac{\gamma(g)}{\beta(g)} - \frac{\gamma_0}{b_0 g} \right] \right\}$$

where $g$ is the coupling strength, $\gamma(g)$ is the anomalous dimension function, $\beta(g)$ is the $\beta$-function and SF stands for the Schrödinger functional renormalization scheme\textsuperscript{23} that
is particularly suitable for lattice calculations. Now, in perturbative computations in the continuum theory schemes different from the Schrödinger functional one are used. The renormalization group invariant quantities now help to make contact to these schemes used in the continuum theory:

\[
\langle x \rangle^{\text{MS}}_{\text{ren INV}} = \langle x \rangle_{\text{MS}}^{\text{my preferred}} = O_{\text{ren INV}}^{\text{my preferred}} f_{\text{MS}}(\bar{g}^2(\mu))
\]

where MS stands for a standard scheme, often used in perturbative calculations and my preferred is self-explanatory. From these considerations it should become clear that the computation of the RGI quantities are the most important ones to come out from the non-perturbative lattice calculations.

For completeness, we will give the value of the RGI matrixelement as obtained form the numerical simulations discussed in the main text, i.e. \( O_{\text{INV}}^{\text{ren}} = 0.222(24) \).

References

4. P. Hernández, K. Jansen and L. Lellouch, these proceedings.
7. S. Capitani, hep-lat/0108028.
10. V. Bornyakov et.al, hep-lat/0111042; V. Bornyakov and M. Müller-Preussker, hep-lat/0110209; R. Horsley, T.G. Kovacs, V. Linke, D. Pleiter and G. Schierholz, hep-lat/0111030; V. Bornyakov et.al, hep-lat/0103032.
16. R. Sommer and H. Wittig, these proceedings.
17. G. Schierholz, these proceedings.
18. F. Karsch, these proceedings.
The detection of the first three waves in the spatial fluctuation spectrum of the cosmic microwave background has allowed us to determine the parameters of any given cosmological model with very high accuracy; already the data analysis requires supercomputers. The modelling of the early phases of the universe, with structure formation, the development of the first galaxies, as well as the first black holes also require the highest level of numerical sophistication and the largest computers. As a result we know today the parameters of today’s standard concept for the cosmology of our universe, like the Λ-parameter often referred to as dark energy, the dark matter content, and the baryonic matter content with great accuracy - and yet we do not even come close to a physical understanding: It is generally hoped that particle physics, in this case a cosmic particle physics will provide testable predictions, where the entire universe will disprove or prove a future physical deep understanding as yet eluding us. Two key aspect of young galaxy evolution serve here as examples for the development of our approach to the physics of the early universe: First, the interaction of black holes and the surrounding stars in the cores of galaxies can serve as a testbed for our growing awareness and physical understanding of the environments of the ubiquitous black holes in the centers of galaxies; and second, the probable surviving building blocks for galaxy evolution, dwarf galaxies, their evolution of the stellar population as well as the hot and cold gas content with their chemical abundance distribution is the other example. Both require hierarchical numerical codes used on supercomputers, and in both cases the modelling can be compared in rather great detail with observations at many wavelengths, from the radio to TeV gamma rays across the entire electromagnetic spectrum. It can be expected that the detailed interaction between technological development, observations with the newly developed instruments, and supercomputers modelling together with analytic phenomenological reasoning will lead us further to a much deeper understanding of the world around us.

The very rapid development of new highly sensitive detectors to measure the spatial fluctuations of the cosmic microwave background has advanced a great deal in 2001, with the simultaneous publication of the first three overtones or wave amplitude peaks, by three experiments at the same time; two of these experiments were balloon borne instruments. Here the advance in balloon technology simultaneous with the rapidly improving detector sensitivity was key. The results were outstanding: Using supercomputers it has been possible to test quantitatively the predictions of the standard model of the structure of our universe, and determine the numerical values of a great many parameters simultaneously: We now know that the universe not only contains baryonic matter like stars, gases and people; there is almost an order of magnitude more baryonic matter hidden in black stars or invisible matter - the suggestions have ranged recently from a very large number of white dwarfs to a rather warm gas component, unobservable due to the ultraviolet absorption in interstellar space. Next we know that the universe contains dark matter of a very differ-
ent form, matter that is clumping and provides the basic gravitational field structure in the universe, usually called “cold dark matter”. And then, using presumed standard candles like supernovae of type Ia, we can find at once that the universe apparently manifests a flat geometry, but also contains “dark energy”, accelerating the universe’s expansion. Here we require supercomputers to model the structure formation, the soap-bubble like web of the universe, visible today both in the galaxy distribution, and also in the radio emission from relativistic electrons. The discovery of the missing solar neutrinos by the Sudbury Mine experiment has clinched the case built from the atmospheric neutrinos observed by the Super-K experiment, suggesting that the neutrino mass is finite, but that at the same time neutrinos almost certainly do not provide the key ingredient of dark matter.

Apart from pushing ever deeper into measuring ever smaller amplitudes and shorter spatial waves in the cosmic microwave background, determining its polarization hopefully soon, we can also focus in towards the building blocks of our optically visible universe, the galaxies and their centers. The two lectures, by Rainer Spurzem and by Gerhard Hensler, are thus examples of this pioneering work.

The centers of galaxies like our own almost always show evidence for the presence of a massive black hole, with a mass of order $2.5 \times 10^6$ solar masses in the case of our own Milky Way, up to $3 \times 10^9$ solar masses in a small number of cases in our cosmic neighborhood, such as in the galaxy M87 in the nearest large cluster of galaxies, the Virgo cluster. As galaxies often merge during their evolution, and so two black holes are expected to spiral in towards each other, strongly modifying the stellar distribution. The gravitational interaction between the stars around the black hole with the black hole itself, the interaction with a second black hole from a nearby galaxy is a great test for our understanding. The resulting cusps in the stellar distribution are observable, and so may turn into tell-tale signs after a detailed numerical exploration of the gravitational interaction of the one or two black holes, their surrounding stars, and quite possibly the distribution of dark matter particles, whatever they may be, with the lightest supersymmetric particle one often favored example.

As a corollary we note that the final stage of two black holes spiralling in towards each other should result in a burst of gravitational waves - not yet observed; observations are planned to begin in 2002.

In the same vein, large galaxies are believed to develop from many mergers, and so it is hoped, that dwarf galaxies may be the purest form of a simple galaxy. Dwarf galaxies should tell us the most about early galaxies, their dark matter distribution, the exchange between hot and cold gas, the gaseous mass loss from a galaxy, and infall, the energy input from supernovae and stellar winds from massive stars, stellar mass loss and star formation out of cold gas, and the slow and fast build-up of the chemical elements. These cycles of activity form also the conceptual building blocks for larger galaxies, when mergers and interaction with the environment may play a bigger and thus complicating role. The early environment around galaxies should be reflected in properties of dwarf galaxies, and so they may help us understand when and how galaxies first formed. Numerical modelling is required to advance our understanding of all these physical processes, and helps us to comprehend which physical concepts are key.

These are just two examples out of the important numerical work being done on supercomputers today; they play an increasingly important role in the advance of knowledge about the universe around us.
Due to their structural and chemical peculiarities dwarf irregular galaxies are of particular interest in astrophysics. Internal and environmental effects which exceed the low gravitational energy of these type of galaxies have a large influence on the evolutions of these low-mass systems. Dwarf irregular galaxies serve as an ideal laboratory for understanding plasmaphysical as well as astrophysical processes.

In order to understand the evolution of this type of galaxies an appropriate description must take large-scale dynamical and small-scale thermal processes as well as chemical ingredients and their influence on the processes into account. Since the numerical treatment has to adapt properly to the temporal and spatial resolution of these processes and is therefore most time consuming, only in 2d are feasible without loosing insight into small-scale processes. Here we present numerical simulations of dwarf irregular galaxies performed with our chemo-dynamical galactic evolution code CoDEx. The models not only represent successfully observed signatures but also give an insight into dynamical and physical processes working also in all other kinds of galaxies. By the evolution of a representative dwarf galaxy model we briefly discuss evolutionary phases and typical chemical abundances.

1 Introduction

As the sun is not a unique star, each galaxy like our Milky Way (MWG) is a unicate in the universe. These huge astronomical systems consist of stars and gas with total masses between $10^6$ and $10^{12} \, M_\odot$ and show a great variety of morphological types. Their sizes range from around 2 to 40 kpc.

Two main parameters to classify galaxies are the total mass and the angular momentum. Massive galaxies with masses above $10^{10} \, M_\odot$ have already been divided by Edwin Hubble in the 30s into flattened spiral galaxies (gSs) like our MWG and elliptical systems (gEs). Although, in particular, gSs look very spectacular and complex, also dwarf galaxies (DGs) with masses below $10^{10} \, M_\odot$ are particularly interesting objects. On the one hand, they are the most numerous galaxy type. On the other hand, from cosmological reasons they are expected to serve as the building blocks of galaxy formation, but form at all cosmological epochs, even at present. In addition, dwarf systems are distributed over a wide range of appearances, from dwarf ellipticals with high velocity dispersions and their low-mass end as dwarf spheroidals, to rotationally supported and gas rich dwarf irregular galaxies (dIrr) with a patchy structure due to star formation (SF) sites.

In summary, DGs are most interesting astrophysical objects and serve as an ideal laboratory to investigate processes significant for the chemical and dynamical evolution of galaxies. (See Ferguson & Bingelli\(^3\) for a review about the different morphological types of DGs.)

---

\(^a\) M_\odot = solar masses is a common astrophysical mass unit; \(1 \, M_\odot = 1.99 \cdot 10^{30} \, kg\)

\(^b\) kpc = 1000 pc; pc, i.e. parsec is the usual astrophysical length unit and equals to 3.26 light years; 
\(1 \, pc = 3.08 \cdot 10^{16} \, m\)
1.1 Irregular Dwarf Galaxies

A typical example of a dIrr galaxy is the Large Magellanic Cloud (LMC) on the southern sky, being a small companion to our MWG, 20-40 times lighter and around 5 times smaller. Fig. 1 shows four images of the LMC at different wavelengths. As a typical feature of dIrrs, the visible stellar bar of the LMC is deeply embedded in a large HI disk. While the gas fraction in the MWG is about 10% of the baryonic mass, in dIrrs sometimes not more than 60% of the initial gas mass is transformed into stars yet by means of star forming processes. According to their lower gas mass also their SF rate amounts to $10^{-3}$ to $10^{-1}$ $\frac{M_\odot}{yr}$ only compared to 2.0 $\frac{M_\odot}{yr}$ in the MWG. The old stellar population is usually widely distributed, while the current SF regions show a patchy structure as in Fig. 1.

Because of their lower binding energy internal energetical processes like supernova (SN) explosions, stellar winds and stellar radiation strongly influence the evolution of dIrrs and make their figure irregular. They are also strongly affected by external influences like infall of intergalactic gas and encounters with nearby galaxies. Most dIrrs have undergone various epochs of enhanced SF. Some galaxies with very bright, blue and compact SF centers are in an extreme state of SF, a so called starburst (SB), probably triggered e.g. by
Due to the low gravitation of these low-mass galaxies a continuous mass outflow driven by SNe has to be assumed. To keep the baryonic matter gravitationally bound an additional dark matter component is required, whose nature is still unknown. Remarkably, dIrrs have low chemical element abundances (always lower than the sun and sometimes down to 1/100 of the solar value) with a large scatter. Often abundance peculiarities are observed.

1.2 Problems in Understanding Chemical Evolution

Our universe started from the primordial gas mixture consisting mainly of hydrogen and helium with tiny portions of other light elements (Li, B, Be). All heavier chemical elements have been produced by nucleosynthesis processes in the hot interior of stars.

The element mixture blown back into the galactic gas at stellar death depends mainly on stellar masses but also on individual conditions, like e.g. initial chemical composition or the interaction with a companion in a binary star system.

All chemical elements heavier than helium are referred to by astronomers as metals. This metallicity has increased from zero to the solar value of 2% in mass or even higher values in some galaxies. It traces the evolution of galaxies. Analytically, in a simple model of an isolated system the metallicity increases with the logarithm of gas consumption by SF. This simple picture is to some extent not fulfilled in dIrrs with respect to the existence of underlying old stellar populations but still low metallicities.

It is more likely to assume scenarios of very low SF rates and sporadic events like SB under which conditions metal-rich gas can be expelled from the galaxy. The infall of primordial intergalactic gas clouds could additionally dilute the dIrr’s gas mixture. Huge halos of neutral gas observed around many dIrrs could serve as a reservoir of pristine gas.

Additionally to the metallicity the ratio of chemical element abundances is of special interest, in particular, nitrogen ($N$) and oxygen ($O$) are well observable even in distant and faint galaxies. $N$ is mainly released by intermediate mass stars to the warm cloudy interstellar gas (CM), while the high-mass stars produce O and expel it into the hot intercloud medium (ICM). These two gas phases are dynamically decoupled but exchange matter due to evaporation and condensation processes on the surfaces of gas clouds. Only a detailed comparison of both dynamical and mixing timescales can provide an insight into the amount of stellar nucleosynthesis products exchanged between the gas components. The following processes are interesting with respect to their influence on chemical abundances:

- Oxygen is lost from the galaxy by means of galactic winds if not mixed into the warm gas phase before. This leads to a lower O and higher $N/O$ value.
- Short-living massive stars in SBs decrease the $N/O$-rate by $O$ release, until long-living intermediate mass stars dominate the metal production and release.
- Infall of pristine gas leaves the $N/O$-ratio unchanged but leads to generally lower metallicities.

Detailed comparisons with observations and, even more important, a close look on the detailed effects of astrophysical processes and their timescales are necessary to decide about

\[ ^{6}\text{by supernova explosion in the case of massive stars or by the formation of planetary nebulae for stars with masses below ca. } 10 \, M_{\odot} \]

\[ ^{4}\text{with masses between } 1 \, M_{\odot} \text{ and nearly } 10 \, M_{\odot} \]
their importance for the galactic evolution. Significant processes which affect the evolution of these galaxies are:

- large-scale streaming motions,
- local gas-phase mixing processes,
- long cooling timescales due to low mass densities,
- starburst-trigger events which also fuel the dIrr with pristine gas,
- star-gas interactions with a large energetic input relative to the energy content of the whole dIrr,
- loss of metals due to galactic winds and the resulting change of chemical abundances.

2 Numerical Models

Astrophysics is unique among all physical disciplines since it has to suffer from the impossibility of hand-made experiments. Astronomical observations, even with enormous and sophisticated instruments, like e.g. the Hubble Space Telescope or the European Very Large Telescope, offer always a ‘snapshot’ of the extremely long-lasting astrophysical evolutionary progression. The evolution of galaxies is counted in billion years and begins soon
- two dimensions, cylindrical symmetry
- staggered grid, logarithmically stretched
- explicit/implicit code
- spatial resolution 40 – 50 pc
- total grid size 20 kpc
- operation splitting, van Leer advection scheme
- 5 component: 3 stellar, 2 gaseous
- differential equations for each component
- 7 variables per component (hot gas: 5)
- 22 materialistic and energetic transition rates

Table 1: Numerical characteristics of the cd galactic evolution code CoDEx

after the ’birth’ of our universe, the Big Bang, around 15 billion years ago. Evolutionary simulations of galaxies try to reproduce the history of galaxies in order to explain individual signatures of galaxies and to study the influences of evolutionary effects.

One approach to treat the dynamical evolution are N-body simulations where the galactic components are represented by thousands of mass points. These calculations can properly deal with 3d structures, like e.g. bars or external perturbations by means of galactic encounters, and are at present in progress to handle the chemo-dynamical interaction scheme (Berczik et al.1). Nevertheless, they lack of sufficient spatial resolution for small-scale effects, although the particles concentrate where mass is accumulating.

The main purpose of hydrodynamical grid codes is the proper treatment of gas dynamics in galaxies even on smaller scales. If interaction processes, however, are treated on small scales, they are limited yet to 2d. Chemical evolution calculations are trying to trace the metallicity and chemical abundance ratios.

The chemodynamical (cd) evolutionary description applied in our models are the most complex and sophisticated numerical calculations concerning the entire galaxy. A 2d hydrodynamical grid code is combined with a complete set of plasmophysical and astrophysical processes as shown in Fig. 2. The chemical evolution is taken into account during the lives of stars by means of by means of a treatment simulating the nucleosynthesis of chemical elements inside stars. Since dynamical and chemical processes are closely coupled only these cd simulations allow to get a self-consistent picture of the SF history and also the chemical evolution of galaxies. The restriction to two dimensions due to the complexity of these calculations does not falsify the results as comparisons with less complex and resolved 3d simulations show.

Table 1 shows an overview of the numerical characteristics of the cd galactic evolution code CoDEx. This modeling of galactic evolution is properly treated by the cd prescription. Its formulation in 1d and 2d dynamics with the ”materialistic” and ”energetic” equations can be found for interested readers e.g. in Theis et al.8 and Samland et al.6, respectively.

2.1 A Representative Dwarf Irregular Galaxy Model

The model galaxy presented here starts with a gaseous mass of $10^9 \, M_\odot$ within the total numerical grid size of $20 \times 20 \, kpc^2$. The matter is distributed according to a Kuzmin-
Plummer-model (Satoh) and is thereby in virial equilibrium until a collapse of this protogalactic cloud is initiated by dissipative processes like radiative cooling or cloud-cloud collisions. Even after the collapse is finished most of the initially given matter remains in the form of CM gas and envelopes the galaxy as a gas reservoir for later infall. Merely $10^8 M_\odot$ form the core region of the galaxy where most of the SF takes place. Additionally to the gas matter a static dark matter halo with a total mass of $10^{10} M_\odot$ ($10^9 M_\odot$ in the core region) is given with a density distribution according to Burkert.

3 Results

3.1 Evolutionary Phases

The evolutionary history of the model galaxy (as shown in Fig. 3) is traced for a time intervall of 10 Gyr since the beginning of the collapse and in characterized by five distinct epochs which differ with regard to the dynamics of the two gas phases and the SF history as well as to the morphological structure. (See Rieschick & Hensler for details.) The phases are:

- Collapse phase ($0 - 0.3$ Gyr): collapse of protogalactic cloud; increasing SF activity,
- Post-collapse phase ($0.3 - 0.8$ Gyr): reduction of SF rate due to self-regulation and reexpansion,
- Transitional phase ($0.8 - 2.0$ Gyr): formation of final galactic structure: center and thick disk,
- Turbulent phase ($2.0 - 5.8$ Gyr): strong variation in SF rate, galaxy grows slightly,
Figure 4. Evolutionary tracks of 2d \( cd \) models for \( 10^9 \) M_⊙ galaxies with (upper dashed curve) and without DM halo (lower full line) in comparison with N/O vs. O/H measurements of SF regions in different objects. The symbols are: yellow filled circle: the solar value, stars: quasi-stellar objects (QSOs), blue triangles: damped Lyman alpha galaxies (DLAs), green diamonds: Lyman limit systems (LLS), red squares: dwarf galaxies (DGs).

- Irregular phase \( (> 5.8 \text{ Gyr}) \): equilibrium is reached; short-scale fluctuations dominate.

3.2 Gas Mixing Cycles

During the whole evolution matter is exchanged between the gaseous components leading to a mixing of matter and especially to a redistribution of the metals produced by stars on different cycles. About 90% of the metals released in PNe and 25% of the SN products stay in a small-scale local cycle. On the other hand, gas is blown out from the galactic body with high initial velocities of about 100 km/s and reaches typical distances from the galactic center of 5 kpc. Although this velocity exceeds the gravitational binding, due to energy losses by expansion, interaction with clouds and work against external pressure the gas returns to the inner parts of the galaxy as CM infall. This forms a second, so called galactic cycle.

On even longer ranges a global cycle occurs where the galactic gas mixes with the intergalactic medium. This metal-enriched gas can, since it is not longer gravitationally bound, only fall back in the case of external infall events. Only the minor part of metals leaves the galactic gravitation. From classical simulations only the local and the global cycle are known. Cd simulations show that rather the galactic mixing cycle which includes the gas reservoir, is the most important one.
3.3 Reproduction of Observed N/O Chemical Abundences

Fig. 4 shows different evolutionary tracks up to a galactic age of \(3 \cdot 10^9\) yr. In contrast to the metal enrichment in classical chemical evolutionary models, here the evolutionary track in the N/O-O diagram does not circumvent the regime of the observed values for dIrrs along a horizontal track at almost solar N/O ratios, but passes from low N/O and O/H values through the regime of dIrr observations. At higher ages the tracks would leave the regime of the observed values to larger O value, unless sporadic infall of primordial or sparsely enriched gas would reduce the metallicity again.

The interested reader is referred to Hensler et al.\(^4\) for further details about chemical implications from \(cd\) models without assuming artificial mixing and outflow processes.

4 Conclusions

As a great success the \(cd\) models reproduce already convincingly the observed abundance peculiarities of dIrr in a self-consistent way. Since the \(cd\) treatment is appropriate for sensitively balanced systems of low gravitational potential energy like the disks of spiral galaxies or, in this case, dIrrs, its models can provide a fundamental insight into interaction processes of small-scale energetics and large-scale dynamics.

Contrary to the often used assumption in chemical evolutionary simulations, the \(cd\) calculations show only a small selective outflow of metal-enriched SN expelled gas. A small fraction of metals is mixed locally, while the larger part enters the long-range galactic cycle coupled to the enveloping neutral gas reservoir. While some scenario assume that the time needed by the metals to return to the inner regions of the galaxy is determined by the dynamics of the hot gas, the \(cd\) models show that the timescale depends solely on the infall velocity of the cold CM onto which the metals condense in shorter times. The mixing process will change if infall of large amounts of intergalactic matter, gravitational interaction with a nearby galaxy or even mergers are assumed, that drive gas from the reservoir into the galactic body. In this case the deposited metals flow into the SF regions instantaneously and thereby would be accessible to observations.

The \(cd\) models indicate that instead of gaseous outflows infall events of pristine material lead to the reduction of metals. By this infall also the SF history with current peaks and ancient episodes, determined from underlying old stellar components can be explained. \(Cd\) models suggest that infall events are essential to explain the trigger of enhanced SF, since the usually self-regulated SF does not permit large variations of the stellar populations.

We have shown that only self-consistent, and hence, numerically expensive simulations which take all coupled non-linear processes into account, can provide a detailed picture about the mixing processes and the metal enrichment of galaxies.

Up to now the models are calculated with a program version optimized for vector computers. Because of the large fraction of time consuming local processes parallization of the code seems to be reasonable. Programming and first testing activities concerning this item are in progress now.
Acknowledgments

We gratefully acknowledge cooperation and discussions with P. Berczik, J. Köppen and Ch. Theis. This work is supported by the Deutsche Forschungsgemeinschaft under grant no. He 1487/5-3 and He 1487/23-1.

References

Supernova Explosions of Massive Stars

Hans-Thomas Janka¹, Markus Rampp, Konstantinos Kifonidis, and Robert Buras

Max-Planck-Institut für Astrophysik
Karl-Schwarzschild-Str. 1, 85741 Garching, Germany
¹E-mail: thj@mpa-garching.mpg.de

Supernova explosions belong to the most energetic phenomena in the Universe. When a massive star is disrupted at the end of its life, a spectacular outburst of light is emitted which can reach the brightness of a whole galaxy. Heavy elements, produced by the star during millions of years of quiet nuclear burning, and radioactive nuclei, freshly created during the early moments of the explosion, are swept into the circumstellar space to form the seed of a new generation of stars and planets. As brilliant as it may be, such a cosmic catastrophe is only a weak side effect of an even more violent event: The iron core of the massive star collapses to a neutron star or a black hole. The gravitational binding energy released during this process is carried away by neutrinos, which are abundantly produced in reactions of energetic particles. These neutrinos play a crucial role for the dynamics of the stellar collapse and neutron star formation. Only a fraction of about one per cent of their energy is sufficient to cause the disruption of the star. The principle possibility of this neutrino-driven mechanism has been verified by analytic arguments and numerical experiments. Its viability, however, has not yet been convincingly demonstrated by self-consistent hydrodynamical models with a satisfactory treatment of all aspects of the relevant physics. In the project described here, we intend to perform such simulations by combining multi-dimensional hydrodynamics with an accurate handling of the neutrino transport and a state-of-the-art description of neutrino-matter interactions. These simulations will attempt to answer the fundamental question whether supernova explosions are caused by neutrino energy deposition behind the shock, aided by the effects of convection inside the nascent neutron star and by convective overturn in the neutrino-heating region.

1 Supernovae: The (Astro-)Physicist’s Interests

Roughly every second one supernova explodes in any galaxy of the Universe. Something like a 100 million supernovae have enriched the gas of the Milky Way with heavy elements, the oxygen we breathe, the iron in our blood cells, the calcium in our bones and the silicon in the rocks beneath our feet. Their gigantic release of energy and momentum helped shaping the galaxies in the early Universe and affects the formation of new generations of stars. Supernova shocks, plowing through the interstellar medium, are considered to be responsible for the acceleration of high-energy cosmic rays, which continuously bombard the Earth. Moreover, explosions of massive stars are astrophysical sources of neutrinos and can produce gravitational waves which are potentially detectable by the new generation of laser interferometer experiments that are on the verge of starting data taking right now.

Due to their enormous brightness and unexpected appearance as “new stars” on the sky, supernovae have always fascinated human beings and have attracted the particular interest of astronomers. This has not changed until now, but in addition we have realized that a broad variety of physical phenomena are connected with supernova explosions. Extremely interesting physical processes play a role during the collapse and the explosion of a massive star and the associated formation of a neutron star or black hole. Supernova research therefore requires input from many other fields such as atomic physics, nuclear physics, and particle physics. Vice versa, a deeper understanding of supernova explosions
Photodisintegration of Fe Nuclei and Dense Core

\[
\nu_e + p \rightarrow n + \nu_e
\]

and Photodisintegration of Fe Nuclei

"White Dwarf" (Fe−Core)

Progenitor (~ 15 \(M_\odot\))
(Lifetime: 1 ~ 2 \(\times 10^3\) y)

\(10^3\) cm

Late Protoneutron Star
(R ~ 20 km)

\(10^6\) cm

Hot Extended Mantle
(Protoneutron) Star

\(\sim 1\) Sec.

Collaps of Core (~1.5 \(M_\odot\))

30000 − 60000 km/s
(R ~ 10000 km)

Figure 1. Evolution of a massive star from the onset of iron core collapse to a neutron star. The star has developed a typical onion-shell structure with layers of increasingly heavier elements surrounding the iron core at the center (upper left corner). This iron core (enlarged on the lower left side) collapses to a proto-neutron star within a fraction of a second. This gives rise to a strong shock wave which disrupts the star (lower right). The neutron star is initially very extended (enlarged in the upper right corner), and contracts to a more compact configuration while accreting more matter within the next second of its evolution. This phase as well as the subsequent cooling and neutronization of the remnant is driven by the emission of neutrinos of all flavors. (Figure adapted from Ref. [1].)

has important consequences not only for astrophysics and our knowledge of astronomical objects, but can also have far-reaching consequences for other areas of physics.

Perhaps the most fundamental problem of supernova research is the question about the cause of the explosion. What is the mechanism that accelerates the stellar debris up to a tenth of the speed of light? Besides being of fundamental interest, this question needs to be answered before we will be able to explain how the energy of the explosion varies with the mass of the progenitor star, how the production of radioactive nuclei like \(^{56}\)Ni, \(^{57}\)Ni or \(^{44}\)Ti changes with the progenitor, when the compact remnant is a neutron star and when it is a black hole, and how the explosion can obtain a big anisotropy and the neutron stars the observed large kick velocities. Moreover, the question whether and how r-process elements can be created in supernovae is linked to a better understanding of the explosion mechanism and the early evolution of the newly formed neutron star. This, of course, is
also true for reliable predictions of the neutrino signal which hopefully will be measured in very much detail in case of a Galactic supernova in the near future.

The historical detection of about 20 neutrinos from Supernova 1987A in three underground experiments in Japan, U.S.A. and Russia was in overall agreement with expectations from models and confirmed the theoretical perception that neutrinos play a crucial role during stellar core collapse and neutron star formation. Neutrinos dominate these events energetically and carry away the gravitational binding energy of the nascent neutron star. About one per cent of their energy, however, is already sufficient to account for the kinetic energy of a supernova (which again is 100 times larger than the energy emitted in electromagnetic radiation). Therefore neutrinos have been proposed to be responsible for the explosion of the star. Physical arguments as well as computer simulations support this idea, but explosions by the “neutrino-driven mechanism” have so far been obtained only in simulations with substantial and sometimes questionable approximations of the treatment of the input physics. Because of the extreme complexity of the problem, no finally convincing hydrodynamical simulations have been performed yet, although the idea of neutrino-driven explosions has been around for more than 30 years now, and despite of significant progress in supernova modelling has been achieved. A standard model for the explosion of massive stars therefore does not exist.

2 This Project: A Rigorous Approach

In the work described here, we attempt to perform the first numerical simulations which combine all ingredients that have been recognized as relevant for the explosion. The crucial neutrino physics, neutrino transport as well as neutrino-matter interactions, will be treated by the most reliable and accurate method applied to this problem so far. Convective mixing inside the neutron star and hydrodynamic instabilities in the supernova ejecta will be taken into account by performing the simulations in more than one spatial dimension. According to simplified models, these processes can decide between explosion or failure by boosting the neutrino emission of the neutron star on the one hand, and by increasing the efficiency of the energy transfer from neutrinos to the stellar matter on the other. Therefore multidimensional models are necessary to give an answer to the question whether supernova explosions can be obtained by the neutrino-driven mechanism.

Of course, a rigorously accurate and sophisticated approach has its price. Such simulations are very time consuming and require the use of top-end supercomputing facilities like those provided by the John von Neumann Institute for Computing (NIC).

The rest of this article is structured as follows: In Section 3, an overview of the evolution from stars to neutron stars will be given. Section 4 contains a more detailed discussion of the role of neutrinos in supernovae and their importance for the explosion mechanism. In Section 5, hydrodynamic instabilities will be described as an important phenomenon which seems to be essential for understanding supernova explosions and their observable properties. Section 6 will outline the objectives and the scientific progress that will be achieved by the computational project presented here.
3 From Massive Stars to Neutron Stars

After several 10 million years of quiet hydrostatic evolution, stars with initial masses of more than about 8 solar masses end their lives by supernova explosions. These stars have developed an “onion-shell” structure in a sequence of nuclear burning phases by which increasingly heavier elements were created. The heavier nuclei are contained in shells closer to the center of the star, while the lightest elements like helium and hydrogen form the outer layers of the stellar mantle and envelope (Fig. 1, upper left corner).

The core of the evolved star contains iron and iron-group elements, which are the most tightly bound atomic nuclei. No further energy release by nuclear fusion in the center of the star is possible at this phase of the evolution. Therefore the core cannot escape gravitational collapse once it has reached a mass near the critical Chandrasekhar limit.

During collapse the stellar matter experiences a dramatic compression, and protons begin to capture electrons to be converted to neutrons (Fig. 1, lower left corner). The collapse does not stop before the density of nuclear matter is reached at the center. This happens only a fraction of a second after the onset of the gravitational instability. At this moment repulsive nuclear forces stiffen the equation of state, and the core resists further compression. A strong shock wave is launched and starts propagating outward. Ultimately, it will disrupt the star in the supernova explosion and will eject most of the stellar gas into the circumstellar space (Fig. 1, lower right corner).

The newly formed neutron star which is left behind at the center is more massive than our Sun but will ultimately have a radius of only 10 km. A gigantic amount of gravitational energy was stored in its interior as thermal and degeneracy energy of neutrons, protons, and electrons during the collapse from the initial iron core with a radius of several 1000 km to the final, much more compact configuration. This energy is now released in the form of neutrinos within a period of several seconds (Fig. 1, upper right corner and center).

Ongoing electron captures on protons produce about $10^{57}$ electron neutrinos, which diffuse out of the star and thus drive its slowly progressing neutronization. In addition, the nascent neutron star is a very hot object, and roughly $10^{58}$ neutrinos and antineutrinos are created as pairs of all flavors in approximately equal numbers. The emission of these “thermal” neutrinos cools the star on its way to the final, degenerate remnant.

A small fraction of only one per cent of the energy that is carried away by the neutrinos is very well sufficient to power a typical supernova event. But how can this energy reservoir be tapped efficiently enough? How can the energy for the explosion be transferred from the dense neutron star to the more dilute stellar gas around it? Despite of more than 30 years of intense theoretical research, and despite of significant progress in our understanding of the physics going on in dying stars, a finally convincing answer to this fundamental question has not been found yet.

4 Neutrinos and the Explosion

In fact, neutrinos play a crucial role during all phases of stellar collapse, supernova explosion, and neutron star formation. Six characteristic stages can be discriminated during the evolution from the pre-collapse star to the neutrino-transparent, cool neutron star. Figure 2 visualizes these phases, showing the dynamical state, the corresponding nuclear composition of the stellar medium, and the basic character of the neutrino emission.
Figure 2. Schematic representation of the processes that occur in a collapsing stellar iron core on the way to the supernova explosion. The diagrams (from top left to bottom right) visualize the physical conditions at the onset of core collapse, neutrino trapping, shock formation, propagation of the prompt shock, shock stagnation and revival by neutrino heating, and r-process nucleosynthesis in the neutrino-driven wind of the newly formed neutron star, respectively, as suggested by current computer simulations. In the upper parts of the figures the dynamical state is shown, with arrows indicating the flow of the stellar fluid. The lower parts of the figures contain information about the nuclear composition of the stellar plasma and the role of neutrinos during the different phases.
During core collapse electron neutrinos are produced by electron captures on protons and nuclei. Initially, these neutrinos escape essentially unhindered. When the central density increases to roughly a hundredth of the density of nuclear matter, however, neutrino scatterings off nuclei become so frequent that neutrino “trapping” sets in. From this moment on neutrinos are carried along with the infalling stellar plasma.

The production and escape of neutrinos reduces the electron fraction and the pressure in the collapsing stellar core. This accelerates the infall and decreases the size of the subsonic “inner core”. The outer edge of the latter marks the position where a shock front
forms shortly after nuclear matter density has been reached and the core bounces, resisting further compression. The sound waves that are created at this moment cannot travel into the supersonically falling outer layers but steepen to a hydrodynamical shock wave near the boundary of the inner core. This shock begins to propagate outward in mass and in radius.

Doing so, it suffers from severe energy losses because iron-group nuclei are disintegrated to free nucleons in the hot medium behind the shock. Free protons immediately capture electrons and produce electron neutrinos in large numbers. When the shock reaches a density so low that these neutrinos can stream faster than the shock is able to move, a very luminous burst of electron neutrinos is radiated. This moment occurs only a few milliseconds after shock formation and is called the “shock-breakout through the neutrinosphere”. The additional neutrino losses are disastrous for the shock and lead to its stagnation before it has reached the outer edge of the iron core. All current simulations agree in the fact that the bounce shock is unable to directly cause the explosion of the star.

While during this very early post-bounce phase neutrinos drain energy from the layers behind the shock, the situation changes only fractions of a second later. As the shock is slowly pushed to larger radii by the hot stellar matter that is accreted through the shock and piles up on the forming neutron star, the density and temperature behind the shock begin to drop. At the same time, the flow of neutrinos and antineutrinos that diffuse out from deeper layers and the mean neutrino energies increase. This favors the formation of a “gain layer” behind the shock where energetic neutrinos start heating the stellar gas. This happens mainly by the absorption of electron neutrinos and antineutrinos on neutrons and protons, respectively. If the corresponding energy deposition is strong enough, it can revive the stalled supernova shock and can provide the energy for the explosion of the star. Since the timescale for this to happen is a tenth of a second or more, which is much longer than the propagation of the prompt hydrodynamic shock, neutrino-driven explosions were named “delayed” explosions.

The neutron star that is formed at the center of the explosion neutronizes and cools by the emission of neutrinos over a period of several seconds. Neutrinos can diffuse out of the dense and hot inner regions of the compact remnant only slowly. Also at late times, they still heat the surface-near layers of the star and thus cause a continuous outflow of baryonic matter, the so-called “neutrino-driven wind”. This wind can become a very neutron-rich environment and can therefore provide favorable conditions for the production of heavy elements such as gold, lead or uranium through rapid neutron-capture reactions by seed nuclei. Whether such an r-process can occur or not is currently a matter of vivid debate. It requires a suitable combination of a number of properties which characterize the neutrino-driven wind, e.g., its neutron excess, entropy and expansion velocity. These parameters depend on the neutron star mass and radius and are mainly determined by the interaction of the outflowing matter with the intense flux of neutrinos from the forming neutron star.

Because of the central role of neutrinos for the supernova problem, their transport and interactions with matter have to be described with particular care and the highest possible accuracy. This important aspect is therefore treated with high priority in the described project.
Supernova 1987A was a true milestone of supernova research. Exploding in the Large Magellanic Cloud, a small satellite galaxy of the Milky Way, and thus in our immediate cosmic neighbourhood, this supernova offered us the unique possibility of a relatively close view of the events that accompany the death of a massive star. Due to its proximity and the advantages of modern observational technology, Supernova 1987A provided us with an unprecedented wealth of data from nearly the first moment of the explosion to even several
years later.

Spectral information could be obtained in very much detail and provided us with new insights into the dynamical processes during the explosion of a star. In particular, Supernova 1987A showed an unexpectedly early emission of X-rays and gamma-rays, at a time when the expanding star was still very opaque and radiation could only escape from its hydrogen envelope. This meant that radioactive nuclei must have been transported from the site of their creation near the newly formed neutron star into the hydrogen envelope of the exploding star. Indeed, Doppler features of iron lines indicated that iron was distributed over a large range of velocities. Some of the iron was expanding with a speed of up to 4000 km/s, much faster than predicted by spherically symmetric models where the heavier elements are found deeper inside the star and move with smaller velocities. The observations could only mean that the onion-shell structure of the progenitor star had been destroyed during the explosion.

More recently, advanced observational instruments, especially high-resolution X-ray telescopes and gamma-ray detectors, are used to study the morphology and chemical composition of the gaseous remnants of supernovae which exploded hundreds or thousands of years ago. Data obtained by the X-ray satellite *Chandra* in three different spectral bands show a spatially inhomogeneous and anisotropic distribution of nucleosynthesis products in the supernova remnant Cas A³ (Figure 3). Iron-rich filaments seem to be located at the outer edge of the remnant, although iron was created deepest inside the supernova. X-ray images of the Vela supernova remnant reveal fast-moving fragments of hot gas that have overtaken the shock front and form Mach cones by their supersonic propagation through the surrounding medium² (Fig. 4). Reconstruction of the directions of their motion points to a common origin near the center of the remnant, consistent with the assumption that these clumps might have been formed during the stellar explosion. Such findings are in conflict with expectations from spherically symmetric models. They are therefore interpreted as strong indications that large-scale mixing processes and macroscopic anisotropies seem to be a generic feature of exploding stars.

On grounds of theoretical considerations convection was long thought to be of potential importance in stellar explosions⁵. However, the large radial extent of the mixing as inferred from properties of the lightcurve and the spectral lines for the first time in case of Supernova 1987A and meanwhile found in a number of other supernovae, too, was unexpected. Radioactive nickel is produced in the deep interior of the supernova, in the immediate vicinity of the nascent neutron star. This nickel is accelerated to very high velocities during the first second of the explosion. It seems to be able to retain much of its initial velocity, thus penetrating even into the helium and hydrogen layers of the star. This requires that multi-dimensional phenomena are present already during the very early stages of the explosion.

Computer simulations in two and three dimensions have indeed shown that the region of neutrino energy deposition behind the shock is convectively unstable⁶–⁹. Rising mushrooms transport heated gas closer to the shock while downflows of cold matter replace the hot gas near the gain radius, where the heating is strongest⁶–⁹ (Fig. 5). This increases the efficiency of the neutrino-driven mechanism and helps pushing the shock farther out. Convective processes were discovered to be present also inside the nascent neutron star and speed up its depletonization and cooling¹⁰ (Fig. 5). The corresponding boost of the neutrino luminosity also strengthens the neutrino heating and thus supports the shock revival.
Figure 5. Convection inside the neutron star (top) and in the neutrino-heating region behind the supernova shock. The pictures show the velocity field (top left) and the fluctuations of the lepton-to-baryon ratio (top right) in a two-dimensional (axially symmetric), hydrodynamic simulation one second after neutron star formation. Neutron-rich gas sinks in while proton-rich matter rises due to buoyancy forces. The convective motions reach velocities of several 1000 km/s. In the lower figure, the entropy distribution outside the neutron star is displayed at 0.1 seconds after shock formation. The shock has reached a radius of about 500 km at this time. Neutrino-heated plasma appears in deep red, the neutron star is visible as a black circle at the center.

When the shock finally moves out through the onion-shell structure of the progenitor star, hydrodynamic instabilities begin to grow after the shock passage at the boundaries between layers of different chemical composition. The anisotropies that were created in the neutrino-heating region during the first second after bounce act as seed perturbations for these later instabilities. As a consequence of the large seed amplitudes, heavy elements are carried outward with high velocities and helium is mixed inward very efficiently (Fig. 6). The convective processes around the neutron star during the shock-revival phase can therefore naturally explain the inhomogeneities and the mixing observed in stellar explosions.
Figure 6. Snapshot of a two-dimensional supernova simulation about 20 minutes after the formation of the supernova shock in a 15 solar-mass star. The shock has already left the displayed volume, which has a radius of more than two million kilometers, and is now propagating through the hydrogen layer. On the left side, the density distribution is color coded with white marking the highest densities. On the right side, the partial densities of radioactive nickel (red and purple), silicon (green) and oxygen (blue) are largest in the dense filaments and clumps that expand with very high velocities into the more dilute gas of the helium shell.
The nickel velocities found in the calculations are in good agreement with measurements for some supernovae\textsuperscript{11}.

\section{Towards a Standard Model of Supernova Explosions}

Spherically symmetric models with a state-of-the-art description of the input physics do not yield explosions. This result was confirmed by simulations with the most advanced description of the neutrino-matter interactions applied so far, where the neutrino transport was handled by solving the Boltzmann equation. Simulations of this kind have become possible and the corresponding codes have been developed only recently\textsuperscript{12-14}. This new generation of supernova models has reached an unprecedented level of sophistication in the treatment of neutrino effects. For the first time, the numerical inaccuracies are now smaller than the uncertainties associated with the approximations of the microphysics.

On the other hand, multi-dimensional simulations demonstrate the potentially crucial role of convective processes for the success of the neutrino-heating mechanism. Models which do not show explosions in spherical symmetry have been found to explode when convection was taken into account\textsuperscript{8}. The neutrino physics in such comparative studies and in all multi-dimensional simulations performed so far\textsuperscript{8,9,15-17}, however, was described by making serious simplifications and partly by using questionable approximations. Although these simulations are enlightening, they do not yield an answer to the question whether supernova explosions can be explained by the neutrino-heating mechanism, when the microphysics in the supernova core is described according to our best current knowledge.

The current project attempts to take a major step towards a solution of this important question. In a new computer code we have combined our highly accurate Boltzmann neutrino transport method with a multi-dimensional treatment of the hydrodynamics. Also for the first time, the neutrino-matter interactions include a detailed description of the reaction kinematics and the correlation effects of nucleons in the dense neutron star medium.

In all previous simulations these effects have been ignored, thereby accepting errors of the order of several ten per cent around the neutrinosphere up to factors of a few at several times nuclear matter density. First calculations with the new implementation of the neutrino-nucleon interactions show significant and interesting differences.

The multi-dimensional neutrino-hydrodynamics problem has been coined in a form such that it can be handled on current supercomputers with an acceptable amount of processor time, although the requirements are appreciable. The numerical algorithms employed for solving the set of complex equations allow for an optimal performance on shared-memory vector-parallel computers like the Cray T90 of the NIC.

In fact, a shared memory architecture offers significant advantages for dealing with the considered problem. On the one hand, it was possible to generalize the chosen methods easily from spherical symmetry to more than one dimension, with rather limited requirements of manpower for the necessary adaptations to a multi-processor shared-memory environment. On the other hand, the neutrino transport, which is done with an implicit time integration, implies solving big matrix problems. Breaking up the latter into pieces which do not need intense communication and at the same time guarantee a high computational efficiency, is very hard to achieve. In contrast, algorithms are known and have already been developed, which exhibit a nearly ideal scaling behavior on shared-memory platforms.
Acknowledgments

It is a great pleasure for us to acknowledge the input from K. Takahashi and C. Horowitz for upgrading our treatment of the neutrino-nucleon interactions. We are very grateful for conversations with J. Beacom, A. Burrows, M. Liebendörfer, A. Mezzacappa, G. Raffelt, S. Reddy, J. Pons, and T. Thompson. We also thank the Institute for Nuclear Theory at the University of Washington for its hospitality and the Department of Energy for support during a visit of the Summer Program on Neutron Stars. HTJ thanks Patrick Inderst for preparing the graphics of Figs. 1 and 2. This work was also supported by the Sonderforschungsbereich 375 on “Astroparticle Physics” of the Deutsche Forschungsgemeinschaft. The majority of the computations with the new Boltzmann solver for the neutrino transport was only possible because of a grant of computer time on the Cray T90 of the NIC in Jülich.

References

14. M. Liebendörfer et al., Probing the gravitational well: No supernova explosion in spherical symmetry with general relativistic Boltzmann neutrino transport,


Double Black Holes in Galactic Nuclei –
Do they Merge or not?

Marc Hemsendorf\textsuperscript{1,2}, Rainer Spurzem\textsuperscript{2}, and Steinn Sigurdsson\textsuperscript{3}

\textsuperscript{1} Department of Physics & Astronomy
Rutgers University, 136 Frelinghuysen, Piscataway N.J. 08854-8019 U.S.A.
E-mail: marchems@physics.rutgers.edu

\textsuperscript{2} Astronomisches Rechen-Institut, Mönchhofstrasse 12-14, 69120 Heidelberg, Germany
E-mail: spurzem@ari.uni-heidelberg.de

\textsuperscript{3} Department of Astronomy & Astrophysics, 525 Davey Lab.
Penn State University, University Park PA 16802 U.S.A.
E-mail: steinn@najma.astro.psu.edu

We investigate with high accuracy direct numerical $N$-body integrations the time-evolution of dense stellar systems harboring a massive black hole. This is a model for galactic nuclei in centres of galaxies after they merge in the context of cosmological galaxy formation and evolution. An important astrophysical question is how fast these black holes come close to each other due to dynamical friction and superelastic scatterings with field stars. At a certain critical separation they would spiral-in and merge induced by gravitational radiation. The computer programmes \textsc{EUROSTAR} and \textsc{NBODY6++} developed for use with mpi libraries maintain high accuracy over billions of integration timesteps, allowing a detailed analysis of the orbit of the black hole(s), its interactions with field stars, and the two-body relaxation in the surrounding dense stellar system.

1 Binary Black Holes in Galactic Nuclei

Massive black holes are very likely to reside in the centres of galaxies as a fossil of earlier activity\textsuperscript{16,18}. For their formation collisionless dynamical general relativistic collapse or dissipative processes during galaxy formation have been proposed\textsuperscript{32}, but a complete and quantitative understanding does not yet exist (compare, however, some related work in that direction in\textsuperscript{17,4}). Nowadays there is strong evidence that the formation of central black holes in galaxies can at least qualitatively be understood in semi-analytical models of galaxy formation in the framework of hierarchical cosmological build-up of structure\textsuperscript{12,15} (see also earlier work of e.g. Eisenstein & Loeb\textsuperscript{7}). These studies are complemented on the observational side by strong correlations found between the central black hole mass and global quantities of the surrounding galaxy or galactic bulge (luminosity or mass, and velocity dispersion\textsuperscript{8,9}).

Following Begelman, Blandford & Rees\textsuperscript{5}, the central black holes of two galaxies will ultimately coalesce with strong gravitational radiation emission, after their mother galaxies have merged. During the early stages of the merger, the stellar component will form a nearly spherical system within the short timescale of violent relaxation. After that, the two supermassive black holes move through the stellar component with a velocity similar to the initial relative motion between the two galaxies. From this moment on, both massive bodies will feel dynamical friction from the surrounding stars. This friction leads the black holes to the newly-formed galactic centre, while the frictional force becomes more efficient
with increasing density. Through this process, the black holes must inevitably ‘find’ each other and form a binary system\textsuperscript{21}.

After it forms, the bound binary hardens first through dynamical friction. Once the separation of the binary is too small to be further affected by dynamical friction, there are still flybys and resonant scatterings with individual field stars, the latter providing a further hardening source. In an idealized situation, where the binary is at rest, loss-cone stars on orbits subject to a resonant scattering will be depleted and the hardening of the black hole binary would stall, hardening time scales, may become very long, as was noted by\textsuperscript{5}. Close, superelastic three-body scatterings between the black hole binary and single stars, however, will generate a strong recoil on the binary, which thus starts a stochastic walk around in the nucleus, and may always find enough interaction partners for further hardening. Direct \textit{N}-body simulations with high particle numbers are a very good tool to study this effect quantitatively\textsuperscript{23, 33, 13}.

We follow the sinking of two massive black holes in a spherical stellar system. The massive particles become bound under the regime of dynamical friction. Once bound, the binary hardens by three body encounters with surrounding stars. Unlike other assumptions the massive system moves inside the core providing an enhanced supply of reaction partners for the hardening. These are the first results from simulations applying a hybrid “self consistent field” (SCF) and direct Aarseth \textit{N}–body integrator (NBOY6), which synthesises the advantages of the direct force calculation with the efficiency of the field method. The code is aimed for use on parallel architectures and is therefore applicable for collisional \textit{N}–body integrations with extraordinarily large particle numbers (> 10\textsuperscript{5}). It opens the perspective to simulate the dynamics of globular clusters with realistic collisional relaxation, as well as stellar systems surrounding a supermassive black hole in galactic nuclei.

These problems attract much attention in the astrophysical community\textsuperscript{11}. Work on this has been done either by solving the perturbed two and three body problem in simplified models\textsuperscript{37, 26} or by \textit{N}-body simulations\textsuperscript{25, 33, 23, 21}. As this shows, the modelling of binary black hole hardening turns out to be extremely challenging, algorithmically and computationally. This work introduces some details of the problem of a sinking black hole binary.

## 2 Results

To give an example of one of our models, we have followed the shrinking of two black holes of

\[
M_\bullet := 1.00015 \cdot 10^7 \, M_\odot, \tag{1}
\]

in a galactic nucleus of \(M_{\text{tot}} = 10^9 \, M_\odot\), using 65,000 or in some case 128,000 stellar particles in the simulations. So the mean mass of a particle is \(\bar{M} = 1.52588 \cdot 10^4 \, M_\odot\). This choice means that every stellar particle with mass \(M_*\) represents a compact star cluster with the order of \(10^4\) particles. The chosen mass for the black hole particle has approximately the same mass as the central black hole of M31\textsuperscript{18}. In a typical model, the initial distance between the black holes is 355.39 pc. They become bound after approximately 40 million years. The total simulated time is approximately 190 million years. At the end of the simulation the black holes’ distance varies between 1 pc at the apo-center and 0.2 pc at the peri-center. The semi-major axis of the black hole binary at the time it becomes first
bound is 21 pc. For dimensional analysis we can identify the time the binary becomes first bound with the time it has the first critical separation \(a_h \approx GM/\sigma^2\), where \(\sigma\) is the velocity dispersion of the stellar system outside of the cusp dominated by the black hole, and \(M\) is the total mass of the black hole binary. At this stage the dominant interaction process with the stellar system changes from dynamical friction to three-body encounters. The minimum separation of 0.2 pc at the end of our simulation is precise of the order of the second critical separation \(a_{\text{crit}} \approx 0.012 \text{pc} \cdot M_8^{0.8}\), where \(M_8\) is the black hole mass in units of \(10^8 M_\odot\). At a semi-major axis of the order of \(a_{\text{crit}}\) quick gravitational radiation merger sets in; \(a_{\text{crit}}\) is approximately 0.01 \(a_h^{24}\), and since the semimajor axis of our black hole binary at the end of the simulation is still 0.6 pc, we are only a factor of three away from the gravitational radiation merger. This is closer than any previously published models of that kind, and similar to the models of Milosavljevic & Merritt\(^\text{30}\), who use the same code for the central region than we.

The final eccentricity of the black hole binary in our model varies around 0.7. It depends on the initial eccentricity; we start with more eccentric orbit of the black hole than Milosavljevic & Merritt’s models\(^\text{30}\), whose black hole binary starts on a circular orbit, and they obtain smaller final eccentricities of only 0.3. This depends on the merging history of the nucleus and its density structure. Eccentricity is of crucial importance for the final fate of the black hole binary since the gravitational radiation time scale becomes much smaller for high eccentricities. The evolution of the orbital data of the black hole binary can be seen in Fig. 1. More details are published in Hensendorf, Sigurdsson & Spurzem\(^\text{13}\).

In total we have performed a number of (up to now) ten different runs used to acquire a small statistical data basis. As one example we show the time evolution of the semi-major axis of the black hole binary in an ensemble averaged sense (error bars are intrinsic 1-\(\sigma\) scatter of the data). We also find a smaller than expected decrease of the motion of the black hole binary’s centre of mass with increasing particle number. Despite of expectation that Brownian motion of the black hole binary should decrease with increasing \(N\), we find that our results are consistent with no dependency on particle number. That would mean this effect is not a classical Brownian motion, but other effects, e.g. induced by the superelastic scatterings or collective interactions with the stellar system’s core play a role. Our statistical data (compare Figs. 5 and 6 of Hensendorf, Sigurdsson & Spurzem\(^\text{13}\) for the available data on statistically averaged motion of the black hole binary) are not yet complete and reliable enough to give further conclusions here.

This motion prevents the binary from easily evacuating surrounding stars, which establishes an efficient hardening even at the late stages when dynamical friction becomes less important for this. Movies in MPEG format (1.6MB) of this process are available at ftp://ftp.ari.uni-heidelberg.de/pub/staff/marc/MPEG/simulation600.mpeg

This movie (as well as simulations800.mpeg and simulations400.mpeg) illustrates in the first stages the standard shrinking of the binary black hole to the centre by dynamical friction, then, in the second stages, the feedback effects it has on the core as it moves through the nucleus, and the sometimes rather chaotic motion of the black hole binary, which in our interpretation is responsible for the relatively high eccentricity.

How our present results scale to the case of real particle number of galactic nuclei is the subject of future work. Within the next years and subject to appropriate computing equipment we will be able to follow the black hole binary into its gravitational radiation
merger phase. We have not examined yet the question what happens if a third black hole comes in before the first binary merges. Conventional wisdom says (e.g. Valtonen\textsuperscript{38}) that slingshot ejections would eject single or even binary black holes. Recent huge direct $N$-body models performed by Makino (personal communication) using a GRAPE-6 special purpose computer suggest however, that before that happens, there is a large chance that two black holes in the resonant three-body interaction come very close to each other (eccentricity 0.99) to merge quickly. We have, however, much more carefully than any other study examined the black hole motion in a self-consistent study with very large particle number.
EuroStar proved to be capable of integrating the binary black hole problem in galactic centres as a point mass system. The simulations introduced in this work are the first fully collisional simulations in this field, which could only be carried out on the up-to-date parallel computers accessible to the authors. The new code is going to be applied for simulating collisional dynamics for large $N$ spherical systems including very massive particles.

3 Algorithmic and Computational Aspects

Assume a set of $N$ particles with positions $\vec{r}_i(t_0)$ and velocities $\vec{v}_i(t_0)$ ($i = 1, \ldots, N$) is given at time $t = t_0$, and let us look at a selected test particle at $\vec{r} = \vec{r}_0 = \vec{r}(t_0)$ and $\vec{v} = \vec{v}_0 = \vec{v}(t_0)$. Note that here and in the following the index $i$ for the test particle $i$ and also occasionally the index 0 indicating the time $t_0$ will be dropped for brevity; sums over $j$ are to be understood to include all $j$ with $j \neq i$, since there should be no self-interaction. Accelerations $\vec{a}_0$ and their time derivatives $\vec{a}_0$ are calculated explicitly:

$$\vec{a}_0 = \sum_j Gm_j \frac{\vec{R}_j}{R_j^3} ; \quad \ddot{\vec{a}}_0 = \sum_j Gm_j \left[ \frac{\ddot{V}_j}{R_j^3} - \frac{3(V_j \cdot \vec{R}_j)\ddot{R}_j}{R_j^5} \right],$$

where $\vec{R}_j := \vec{r} - \vec{r}_j$, $V_j := |\vec{R}_j|$, $\dot{V}_j := |\vec{V}_j|$. By low order predictions,

$$\ddot{x}_p(t) = \frac{1}{6}(t - t_0)^3 \vec{a}_0 + \frac{1}{2}(t - t_0)^2 \ddot{\vec{a}}_0 + (t - t_0)\vec{v} + \vec{x};$$
$$\ddot{v}_p(t) = \frac{1}{2}(t - t_0)^2 \vec{a}_0 + (t - t_0)\ddot{\vec{a}}_0 + \vec{v},$$

new positions and velocities for all particles at $t > t_0$ are calculated and used to determine a new acceleration and its derivative directly according to Eq. 2 at $t = t_1$, denoted by $\vec{a}_1$ and $\ddot{\vec{a}}_1$. On the other hand $\vec{a}_1$ and $\ddot{\vec{a}}_1$ can also be obtained from a Taylor series using higher derivatives of $\vec{a}$ at $t = t_0$:

$$\vec{a}_1 = \frac{1}{6}(t - t_0)^3 \vec{a}_0^{(3)} + \frac{1}{2}(t - t_0)^2 \vec{a}_0^{(2)} + (t - t_0)\ddot{\vec{a}}_0 + \vec{a}_0,$$
$$\ddot{\vec{a}}_1 = \frac{1}{2}(t - t_0)^2 \vec{a}_0^{(3)} + (t - t_0)\vec{a}_0^{(2)} + \vec{a}_0. \quad (4)$$

If $\vec{a}_1$ and $\ddot{\vec{a}}_1$ is known from direct summation (from Eq. 2 using the predicted positions and velocities) one can invert the equations above to determine the unknown higher order derivatives of the acceleration at $t = t_0$ for the test particle:

$$\frac{1}{2}\frac{\vec{a}_0^{(2)}}{t - t_0} = -3 \frac{\vec{a}_0 - \vec{a}_1}{(t - t_0)^2} - \frac{2\ddot{\vec{a}}_0 + \vec{a}_1}{(t - t_0)};$$
$$\frac{1}{6}\frac{\vec{a}_0^{(3)}}{t - t_0} = 2 \frac{\vec{a}_0 - \vec{a}_1}{(t - t_0)^3} - \frac{\ddot{\vec{a}}_0 + \vec{a}_1}{(t - t_0)^2}, \quad (5)$$

This is the Hermite interpolation, which finally allows to correct positions and velocities at $t_1$ to high order from

$$\dddot{x}(t) = \dddot{x}_p(t) + \frac{1}{24}(t - t_0)^4 \dddot{\vec{a}}_0 + \frac{1}{120}(t - t_0)^5 \dddot{\vec{a}}_0;$$
$$\dddot{v}(t) = \dddot{v}_p(t) + \frac{1}{6}(t - t_0)^3 \dddot{\vec{a}}_0 + \frac{1}{24}(t - t_0)^4 \dddot{\vec{a}}_0. \quad (6)$$

55
Taking the time derivative of Eq. 6 it turns out that the error in the force calculation for this method is $O(\Delta t^4)$, as opposed to the widely used leap-frog schemes, which have a force error of $O(\Delta t^2)$. Additional errors induced by approximate potential calculations (particle mesh or T\text{REE}) create potentially even larger errors than that. However, it can be shown that the above Hermite method used for a real $N$-body integration sustains an error of $O(\Delta t^4)$ for the entire calculation$^{19}$. Many persons in the world know as Aarseth scheme (in particular the code version NBODY5$^1$) an integrator of the same order as the Hermite scheme, but using only accelerations on four time points instead of $\ddot{a}$ and $\dot{\ddot{a}}$ on two time points. As is shown by Makino$^{19}$, the Aarseth scheme is $O(\Delta t^4)$ as well, but for the same number of time steps the absolute value of the energy error (not its slope) is clearly smaller in the Hermite scheme. This means that for a given energy error the Hermite scheme allows timesteps which are larger by some factor of order unity depending on the parameters of the system under study. The Hermite scheme has been commonly adopted during the past years$^2$, because it needs less memory, and allows slightly larger timesteps. More importantly, after the addition of a hierarchical (as opposed to individual) time step scheme it is well suited for parallelization on modern special and general purpose high performance computers$^{35}$. The timestep scheme will be discussed now.

4 Choice of Timesteps – Parallelization

Aarseth$^1$ provides an empirical timestep criterion

$$\Delta t = \sqrt{\eta \frac{|\ddot{a}| |\dddot{a}|^2 + |\dddot{a}|^2}{|\dddot{a}| |\dddot{a}|^2}}.$$  \hspace{1cm} (7)

The error is governed by the choice of $\eta$, which in most practical applications is taken to be $\eta = 0.01 - 0.04$. It is instructive to compare this with the inverse square of the curvature $\kappa$ of the curve $\ddot{a}(t)$ in coordinate space

$$\frac{1}{\kappa^2} = \frac{1 + |\dddot{a}|^2}{|\dddot{a}|^2}. \hspace{1cm} (8)$$

Clearly under certain conditions the time step choice Eq. 7 becomes similar to choosing the timestep according to the curvature of the acceleration curve; since it was determined just empirically, however, it cannot generally be related to the curvature expression above. Makino$^{19}$ suggests a different time step criterion, which appears simpler and more straightforwardly defined, and couples the timestep to the difference between predicted and corrected coordinates. The standard Aarseth time step criterion Eq. 7 has been used in most $N$-body simulations so far (but compare the discussion by Sweatman$^{36}$).

Since the position of all field particles can be determined at any time by the low-order prediction Eq. 3, the time step of each particle (which determines the time at which the corrector Eq. 6 is applied) can be freely chosen according to the local requirements of the test particle, practically, however, for the purpose of efficient parallelization (originally: vectorization) a hierarchical quantized time step scheme is used, where each particle can only obtain a time step out of a finite set of values$^{20}$. In practice the timestep is taken from the set $\{2^i | i = -n, 0\}$ with $n \leq 32$.  

56
Figure 2. CPU time needed for one \( N \)-body time unit as a function of particle number \( N \) using NBODY6++ on the CRAY T3E. The collection of data points includes runs with varying average neighbour number and processor/pipeline number, starting from 8 for low \( N \) up to 512 for the largest \( N \), which are not individually discriminated in the figure.

Another refinement of the Hermite or Aarseth “brute force” method is the two-time step scheme, denoted as neighbour or Ahmad-Cohen scheme\(^3\). For each particle a neighbour radius is defined, and \( \ddot{a} \) and \( \ddot{\ddot{a}} \) are computed due to neighbours and non-neighbours separately. Similar to the Hermite scheme the higher derivatives are computed separately for the neighbour force (irregular force) and non-neighbour force (regular force). Computing two timesteps, an irregular small \( \Delta t_{\text{irr}} \) and a regular large \( \Delta t_{\text{reg}} \), from these two force components by Eq. 7 yields a timestep ratio of \( \gamma := \Delta t_{\text{reg}} / \Delta t_{\text{irr}} \) being in a typical range of 5–20 for \( N \) of the order \( 10^3 \) to \( 10^4 \). The reason is that the regular force has much less fluctuations than the irregular force.

If the two-body force between any pair of particles becomes dominant their (perturbed) relative motion is integrated in special regularized coordinates (taking into account perturbations from field particles), in which the singularity of the two-body motion is transformed into a slowly varying parameter (the binding energy) and does not occur in the integration variables. The rest of the \( N \)-body simulation generally regards the regularized pair as a compound particle located at the position and moving with the velocity of its centre of mass, except in the case when a perturber moves very close to a regularized pair (in such cases the pair is resolved). An excellent account of regularization, historically and scientifically has been given by Mikkola\(^{27}\). Most recent developments are the slow-down treatment of tight binaries\(^{28}\) and a new method to gain accuracy and exact solutions in the unperturbed case using Stumpff functions\(^{29}\).

For the binary black hole models of galactic nuclei described in the previous section a hybrid code was used, which embeds the direct \( N \)-body region in a larger system, where the potential is approximately computed by a serious evaluation; the latter piece is denoted

57
by some authors as SCF ("self-consistent fields\textsuperscript{14}). In Fig. 3 the speed-up of the combined code as a function of the number of nodes is displayed. For parallelization we use a hand-made MPI implementation, parallelizing strategic loops, which need most of the CPU time. The direct $N$-body portion of the code (particle-particle interactions) is the most dominant one still, and it uses a parallel force computation for the long-range and short-range forces in a loop over those particles which are due for integration in the individually blocked time step scheme. No domain decomposition is used here yet, all particle data are sent in copy to all PE's. Here, a typical integration time needed is about 1 hour wall clock for 128k particles on 128 processors of the CRAY T3E. A model for the binary black hole problem requires for each individual run about 60 time units. Models for the globular star clusters are much more complicated, because they need a few hundred time units in physical time (long relaxation time). Also regularized binaries are present in a larger number, and there are physical hints from star formation that we should include in our models thousands of close binaries from the very beginning\textsuperscript{10}. Future work on improvement of our implementation is therefore directed in two directions: first an efficient parallel integration of the regularized pairs (work nearly completed now with S.J. Aarseth), and second, a domain decomposition and division of the force for the very distant particles, in order not to lose accuracy but gain efficiency. The recent papers of Makino\textsuperscript{22} and Dorband, Hensendorf & Merritt\textsuperscript{6} both discuss future prospects regarding the use of hypersystolic algorithms and non-blocking communication.
Acknowledgements

The authors would like to thank S. Aarseth and D. Merritt for fruitful help and discussion. This project is funded by Sonderforschungsbereich (SFB) 439 “Galaxies in the Young Universe” at the University of Heidelberg. Support and computer resources awarded by NIC Jülich and HLRS in Stuttgart are gratefully acknowledged.

References

35. Spurzem, R., 1999, in Riffert, H., Werner, K. (eds), Computational Astrophysics,
   The Journal of Computational and Applied Mathematics (JCAM) 109, Elsevier Press,
   Amsterdam, 407.
Computational Chemistry, which was started by a handful of quantum chemists in the 1950s and 1960s, has achieved in the last decade its firm position among the other disciplines of chemistry, in molecular physics and pharmacy. In some areas, such as the determination of molecular structures or relative stabilities of molecules, it has become a complementary tool to experimental investigations. In many other areas, however, in which details of the microscopic behaviour must be known in order to understand the mechanisms and course of a chemical reaction, computational chemistry may be the only tool to study binding interactions or properties of short-lived intermediates. These investigations should be performed in close connection with measurements which in general observe only starting material and final products depending on the experimental situation. There are even some questions which can only be answered by computations, for example if molecules cannot be synthesized or have very short life times under given conditions; typical such study areas are reactions in the atmosphere or in interstellar space. Many of such investigations are undertaken on powerful workstations. For large molecules, however, in particular approaching the biological regime, and for very complex molecules, whose investigation requires unusual high accuracy, high-performance computers such as available at the NIC are required.

The following articles represent typical examples for to-days chemical investigations. The computational methods employed range from simple force-field molecular mechanics (MM) methods to more or less routine quantum chemical (QC) procedures such as density functional theory (DFT) and self-consistent-field (SCF) calculations with additional Möller-Plesset (MP2) perturbation corrections, all the way to high-level QM methods such as coupled cluster (CCSD(T)) procedures or multi-reference configuration interaction (MR-CI) procedures. For large systems combined treatments are in use, combining quantum mechanics for the description of subtle details with molecular mechanics (QM/MM) for the less important part of the systems. Simulation methods such as molecular dynamics (MD) are also essential and have been improved for certain applications.

The various contributions show very clearly the great importance of the close cooperation between the experimentalist and the computational chemist. Most of the results would not have been obtained if each of the researchers had stayed within his own discipline.

Binding properties in protein nucleic acids, in particular hydrogen (pairing) bonds and van der Waals (stacking) properties are studied by modifying the backbone of a DNA strand in order to be able to better differentiate between these types of interactions. The synthesis of the modified duplex, in which the complicated geometrical conditions of the helical double strand DNA are strongly simplified, was essential in this regard. The role of transition metal centers in catalytic and enzymatic active sites is studied in two contributions. The biological catalytic activity of vanadium-containing enzymes could be characterized by chemical, biochemical and crystallographical data, but only computations are in a posi-
tion to get more information on the molecular mechanism of the reactions at the active site and on possible reaction intermediates. Organometallic molecular catalysts have unprecedented selectivity, and their efficiency can be compared with the astounding selectivity of enzymes. Hence it is an obvious goal to design such catalysts - in this case based on a chiral rhodium compound. An obvious problem, as in many such calculations, is the lack of force field parameters involving transition metals bonding. The parallel computing environment at the NIC was essential to use the newly developed Genetic Algorithm for the extensive geometrical and force constant optimization. A further contribution shows how experimental results on the study of liquid surfaces can be supplemented by simulation. Electron spectroscopy is able to give information on vertical diffusion within a small surface layer, while MD simulations can be used as ideal tool to study lateral movement and to gain details of inner surfaces. In this combination the ordering processes at the interface liquid/vacuum and the traffic between the bulk and the surface could be established. The last two contributions deal with theoretical developments of quantumchemical methods. The first develops multi-reference configuration interaction procedures, based on determinantal basis functions, and makes use of 128 nodes of parallel computing. Almost perfect scaling between 48 and 128 nodes is achieved. This new development allows to treat CI expansions up to 5x10^6 terms of which up to 5x10^6 are included in the variational wave-function. The code allows to treat medium-sized molecules with up to 100 electrons with very high accuracy, in ground and electronically excited states. This code is a major step forward to treat electron correlation in a proper way which allows to compute with high accuracy entire potential energy surfaces. The second theoretical advance shows how to calculate rovibronic energy levels for tetraatomic molecules. Such development would also not have been possible on a computer of work-station size. A straightforward treatment of such a problem would require several million basis functions to achieve a good description of the vibrational motion, a six-dimensional potential energy surface and the diagonalization of the hamiltonian matrix of the same order as the number of basis functions. Clever computational stepwise strategies led to a general code with impressive performance. It will be very valuable to guide spectroscopic assignments.

All examples demonstrate the power of computational chemistry and the essential role of the John von Neumann Institute for Computing.
The genetic information of life is encoded in the sequence of the bases of the DNA. In B-DNA, the dominant conformer of the DNA under physiological conditions, the base topology, i.e. the orientation and available space for the bases, is strictly limited by the helical structure, in which both strands wind around a common axis in opposite directions (Figure 1).

![Helical double strand in B-DNA](image)

Figure 1. Helical double strand in B-DNA

As a consequence, in B-DNA a purine-base always pairs with a pyrimidine-base. While guanine always pairs with cytosine adenine always pairs with thymine. Furthermore all pairs use the Watson-Crick-pairing-mode, as shown in Figure 2 and 3.

The structure and stability of such systems are determined by a variety of different interactions. Since the helical structure of DNA has been clarified by Watson and Crick, hydrogen bonds have been considered to represent the determinant factor for the stability of base paired systems and for the selectivity of the pairing. In the gas-phase, it has been indeed found experimentally and by calculations, that each hydrogen bond has a strength of 5-7 kcal/mol. In aqueous medium, however, a hydrogen bond contributes only 0.5-1.8 kcal/mol. The weakness of hydrogen bonds in water results because the...
nucleic bases form strong bonds to themselves and to water molecules. Van-der-Waals and dipole interaction between the planar base pairs lead to the so-called \(\pi\) - or base stacking interaction which is the second decisive interaction enabling the formation of duplexes\(^7\). Van-der-Waals interaction, that is caused by fluctuating dipoles, seems to play a central role for this stabilizing factor. In addition, permanent electrostatic attractions and hydrophobic effects contribute to the stacking energy\(^8\).\

While the hydrogen bonds are formed perpendicular to a single strand between both strands the stacking interaction is mainly oriented along a single strand and consequently stabilises the orientation of a single strand itself. While the strength of hydrogen bonds is strongly weakened in polar solvent the size of the stacking interaction is less influenced. Consequently a correct description of the strength of the base pairing and of the relative importance of the various contributions must include solvent effects. The weakness of the single interaction and the complicated interplay between the various interaction causes the difficulties in the theoretical description of such systems.

To differentiate between the various interaction governing the base pairing process, it is important to know how the geometrical structure of the backbone influences the base pairing itself and the secondary structure of DNA depicted in Figure 4. To obtain information about this topic, one can change the properties of the backbone by replacing functional groups of the sugar-rings. Alternatively one can completely replace the backbone as indicated in Figure 5.

The latter idea does not only open the way to a better understanding of base pairing, but can also be used to create new drugs, if these new DNA analogue substances form more stable double strands with DNA or RNA than with itself. If they are able to bind complementarily to DNA or RNA, this behaviour can be employed to bind to mutated sequences and consequently block the expression of defect genes. In the context of this so-called antisense-strategy, in 1991 a new biopolymer was described\(^15\)\(^-\)\(^18\). In these compounds the complete DNA-sugar-phosphate-backbone is replaced by a peptide-like system consisting of diaminoethylglycine-units (Figure 6).

The diaminoethylglycine-PNA (peptide nucleic acid) pairs well with RNA and is easy to synthesize. The pharmacological application has so far been hampered by the fact that no mechanism exists that transports alanyl-PNA into the nucleus\(^18\),\(^19\). Even though the pharmacological usability is limited so far, PNA's are excellent substances to study base
pairing, since backbone properties can be varied easily by using different amino acids. Even the secondary structure can be varied strongly. Examples are PNA’s being formed from alternating D- and L-amino acids\textsuperscript{20}. They build up double strands which are no longer helical, but possess linear geometrical structure (Figure 7). The distance between two neighbouring nucleic acids within the same strand is 3.6 Å, a value very similar to the favoured distance between stacked bases in the DNA (3.4 Å). Because of the rigidity of the peptidic backbone, the orientation of the nucleic acids is well defined so that the complicated geometrical conditions in the DNA are strongly simplified.

Consequently the alanyl-PNA (Figure 5 and 7) offers an excellent opportunity to differentiate between the three remaining interaction stabilizing a PNA-hexameric duplex, namely the hydrogen bonds between the strands, the van-der-Waals-interaction between neighbouring bases (stacking interaction) and the influence of the surrounding water molecules. In difference to B-DNA in alanyl-PNA systems purines also form base pairs with purines and besides the Watson-Crick pairing mode also the reverse Watson-Crick (Figure 8) and the Hoogsten pairing mode can be observed\textsuperscript{20} (Figure 9).

In the present project we want to obtain insight into the relative importance of the various interaction governing the double strand forming process of PNA. To achieve this goal we performed all sorts of computations starting from high level quantum chemical...
methods to low level force field ansätze. While the former are made to study the accuracy of the latter, in the force field computations the formation for a whole PNA hexamer in solution can be described. To obtain accurate descriptions high level methods will be used to fit new parameters for the force field. This is necessary, since very accurate parameters are only known for so-called canonical bases. They fail in the description of non-canonical basis such as xanthine (Figure 10).
In the present stage of the project we concentrate on xanthine-xanthine alanyl-PNA hexamers which were found to build an unexpectedly stable double strand. The stability of double strands is normally measured by the melting temperature. At the melting temperature the strands form a thermodynamical equilibrium consisting of 50% dissociated and 50% non-dissociated double strands. Changes in the concentration of dissociated vs. non-dissociated double strands can be monitored due to the strong difference between the absorption spectra of oriented double strands and single strands. Alanyl-PNA of nucleic bases which form three hydrogen-bonds normally possess a melting temperature of about 58°C while those bound by only two hydrogen bonds dissociate between 24°C and 32°C. The melting temperature of xanthine alanyl-PNA lies at 48°C, indicating a pairing mode with three hydrogen bonds. However, the normal diketo form of the xanthine molecule can only form two hydrogen bonds to another xanthine molecule opening the question to the reason of this unexpected high binding. To explain the high stability, it was proposed that tautomeric forms (Figure 11) participate in the base pairing. This enables the system to form pairing modes with three instead of two hydrogen bonds (Figure 12). The existence of this kind of pairing mode requires, that the energy of the third hydrogen bond overcompensates the tautomerization energy.

![Figure 11. Dienol-xanthine](image1)

![Figure 12. Xanthine-xanthine-base pairing involving dienol-xanthine](image2)

The exact pairing geometry cannot be determined by experimental methods, because such substances are synthesized only in very small amounts. Furthermore they are strongly diluted in aqueous solution. Consequently model calculation are important to investigate the base pairing mode of the xanthine-xanthine alanyl-PNA hexamer.

To investigate the role of tautomeric forms, in addition to the interaction discussed for normal double strand forming processes, the energy of the tautomerization has to be considered in the calculations. To see how well the different methods describe the tautomerization energy, we first performed calculations on uracil tautomers (Figure 13, 14 and 15).

![Figure 13](image3)

Table 1 contains the relative energies of selected tautomeric forms of uracil with respect to the most stable diketo form as a function of various theoretical approaches. It shows that the energy gap between 2-enol-4-keto-3H-uracil and the most stable conformer of the diketo form is about 11-13 kcal/mol. Table 1 also shows that, in comparison to the most accurate CCSD(T) approach, the density functional theory (B3LYP functional) overestimates the energy gap between the various tautomeric forms. One important example
is the dienol form. The energy gaps computed with the more elaborate MP2 approach are very similar to those obtained with the CCSD(T) approach.

<table>
<thead>
<tr>
<th>Tautomer</th>
<th>Basis set</th>
<th>B3LYP</th>
<th>MP2</th>
<th>CCSD(T)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-Enol-4-keto-3H</td>
<td>G-31G(d)</td>
<td>12.4</td>
<td>13.1</td>
<td>12.4</td>
</tr>
<tr>
<td></td>
<td>G-31++G(d,p)</td>
<td>11.1</td>
<td>10.5</td>
<td>10.2</td>
</tr>
<tr>
<td>Dienol</td>
<td>G-31G(d)</td>
<td>15.3</td>
<td>14.4</td>
<td>14.6</td>
</tr>
<tr>
<td></td>
<td>G-31++G(d,p)</td>
<td>12.8</td>
<td>10.9</td>
<td>10.8</td>
</tr>
</tbody>
</table>

Table 1. Computed relative energies of uracil tautomers. All energies are given with respect to diketo form. (All values are given in kcal/mol.)

Figure 16 shows the six possible twodentate base pairing modes for self pairing of diketo xanthine. The respective relative energies and dimerization energies of all base pairs are given in Table 2. The calculated base pairs show remarkable differences regarding the strength of hydrogen bonds. While the dimerization energies of the *reverse* Watson-Crick and Watson-Crick base pairs 1, 2 and 3 ($E_{\text{dim}} = -10.8, -10.4$ and -10.6 kcal/mol, respectively) lie in the lower range of the stabilization energies known for twodentate base pairs, the dimerization energy of the strongest bonded base pair N3-H/O2 - N3-H/O2 (6) is nearly twice as high ($E_{\text{dim}} = -22$ kcal/mol). In this base pair both hydrogen bonds involve the N3-H group. In the base pairs 4 and 5 which reveal also quite remarkable pairing energies ($E_{\text{dim}} = -17.1$ and -14.1 kcal/mol) only one N3-H bond participates in hydrogen bonding while the second donor is provided by the N1-H functionality. A comparison of the various base pairing modes shows that the N3-H group is essential for the strong hydrogen bond (ca. 5-6 kcal/mol). The nature of the carbonyl groups seem to be less important. As shown by the NBO analyses the unexpected strong hydrogen bonds result from an improved electron transfer from the oxygen lone pairs to the $\sigma^*$-orbital of the N3-H-bond. This can take place because the orbital energy of the $\sigma^*$-orbital (0.412 a.u.) of the N3-H bond is much lower than the orbital energy of the $\sigma^*$-orbital of the N1-H bond (0.442 a.u.). The computed interaction energies between the oxygen lone pairs and the $\sigma^*$-orbitals of both N-H bonds also nicely reflect the variations in the dimerization energies. We obtain about 19 kcal/mol if the N3-H bond is involved but only about 12 kcal/mol for the corresponding N1-H values. A similar explanation was previously discussed for the strength of the hydrogen bonds in ice.
Figure 16. All possibilities for twodentate xanthine-xanthine base pairs with the most stable diketo-xanthine tautomer.

Geometrical constraints on the bonding angles were discussed as possible reasons for variations in the strength of the hydrogen bonds in uracil dimers. Actually, also for xanthine the hydrogen bonds of the more stable pairing mode 6 possesses more idealized bond angles (178.5° for the N-H-O and 121.3° for the H-O-C hydrogen bond) than the less stable hydrogen bonds found in base pairs 1-3 (168.7 - 170.8° for the N-H-O and 132.2 to 134.2° for the H-C-O-hydrogen bond). However, if the bond angles of pairing mode 6 are distorted to those optimized for base pairs 1-3 the energy raises by only about 2 kcal/mol showing that geometrical constraints are not too much important. To investigate whether tautomeric forms can explain the unexpected high melting temperature of xanthine-xanthine pairing in alanyl-PNA hexamers next to the pairing modes of diketo-xanthine various pairs involving the tautomers 1-H-2-enol-6-keto-xanthine, 3-H-6-enol-2-
Table 2. Comparison of MP2/TZVPP-energies of xanthine-xanthine-base pairings. (All values in kcal/mol.)

Table 3. Comparison of MP2/TZVPP-energies of xanthine-xanthine-base pairings. (All values in kcal/mol.)

Keto-xanthine, and dienol-xanthine were investigated (Figure 17). Base pairs involving other tautomeric forms are energetically too unfavourable. As summarized in Table 3 the tridentate base pairing of diketo-xanthine with dienol-xanthine in the reverse Watson-Crick 7 and the Watson-Crick-pairing mode 8 have dimerization energies of -17.4 and -17.5 kcal/mol, respectively. The higher dimerization energies of 7 and 8 simply result from the additional hydrogen bond since the strength of the single hydrogen bonds of base pairs 7 and 8 are similar to those found in the twodentate base pairs 1-3. Since the tautomerization energy necessary to build up the appropriate isomer of the dienol form (4.6 kcal/mol) is smaller than the stabilization arising from the additional hydrogen bond (5-6 kcal/mol) both base pairs 7 and 8 were indeed predicted to be about 2 kcal/mol lower in energy than the twodentate diketo-xanthine complexes 1-3, but still higher than the base pairing modes 4-6. Base pairs that need two xanthine nucleobases in an enol tautomeric form are energetically unfavourable since the overall tautomerization energy is too large. The most stable examples are the base pairs with the 1-H-2-enol-6-keto-xanthine recognizing 3-H-2-enol-6-keto-xanthine (9) or 3-H-6-enol-2-keto-xanthine (10), respectively. They are predicted to be energetically disfavoured by 1.4 and 4.4 kcal/mol compared to the twodentate base pair 1 although they possess very high dimerization energies of -30.6 and -34.4 kcal/mol.
For the base pairs our computations predict the energetic order of pairing modes $6 < 5 < 4 < 8 \approx 7 < 3 \approx 1 \approx 2 < 10 < 9$. Besides the tridentate base pairs 7 and 8, also the twodentate pairs 4-6 offer explanations for high $T_m$ values found for the xanthine-xanthine double strands. The unusually high $T_m$ values cannot be explained by the twodentate base pairs 1-3 since the strength of their hydrogen bonds were found to be in the normal range of twodentate pairing modes.

In the first step of the present project we characterized the hydrogen bonding which represents one of major interaction governing the stability of the double strands formed by xanthine-xanthine alanyl PNA hexamers. We computed all possible xanthine-xanthine pairing modes involving the diketo tautomers and to investigate the importance of enol tautomers all tridentate xanthine-xanthine base pairs lying within the energetic range of the twodentate diketo tautomeric base pairs were also computed. All computations were performed on the MP2/TZVPP/B3LYP/6-31G** level of sophistication which was found to be accurate enough for a reliable prediction of the relative energies of the base pairs with and without tautomeric forms. While our model describes the hydrogen bonding between the bases with high accuracy it does not account for other effects determining the formation process of the double strands like influence of solvent, $\pi$-stacking, and backbone topology.

To include these interaction into our model we plan force field calculation which are able to describe the complete double strand including the surrounding water molecules. In a first step we will examine the accuracy of existing force fields. This is necessary since it is well known that the existing force fields are optimized for canonical bases but possess deficiency in the description of non-canonical bases. This is already seen in the tautomeric forms of xanthine which are artificially described lower in energy than the diketo form. For an adjustment of the force field to the present problem we will use the data computed up to now. In the final step we will use the optimized force field to simulate the dissociation.
process of the complete double strand including the solvent.
For our calculations we used approximately 5800 CPU-h’s on the CRAY T3E-256 and T3E-512 machines.

References

Molecules as Mechanical Machines. Genetic Algorithms in the Search of Appropriate Descriptions of Molecular Catalysts

Gottfried Huttner, Johan Friedrich, Axel Frick, Volker Schulz, Jörg Karas, Johannes Hunger, and Stefan Beyreuther

Anorganisch Chemisches Institut
Universität Heidelberg, 69120 Heidelberg, Germany
E-mail: g.huttner@indi.aci.uni–heidelberg.de

The development of selective organometallic molecular catalysts is one of the most promising areas of molecular chemistry and which presently develops at an unforeseen and still increasing pace. Organometallic molecular catalysts have a unprecedented selectivity which is only paralleled by the astounding selectivity of biological catalysts the enzymes. Organometallic catalysts make use of the same type of so called secondary interactions as enzymes to reach this selectivity. They make use of well designed surfaces which embed the catalytic centre and orient the reagent by secondary interactions. The orientation between the catalytically active centre and the substrate is pre–organised in such way that only one of many possible products will result. For the rational design of an organometallic catalyst the knowledge of the secondary interactions is a pre–requisite therefore. Since this type of interaction can not be reliably calculated by quantum mechanical methods for large molecules chemists have to use other types of models, the most promising approach being the force field model which describes a molecule in terms of a mechanical machine. The potentials which are necessary to describe the stiffness of the different types of springs and joints of such a model have to be inferred from experimental data. Force fields describing organic molecules have been developed and are highly successful in modelling and predicting the behaviour of organic compounds. On the other hand force field models for the description of organometallic compounds which would be based on as many experimental data as available are not generally known. The paper describes the development of an appropriate force field model for a class of rhodium compounds which are active in enantioselective catalysis by a novel approach. The parameters of all potentials involving contributions by the rhodium atom are optimised by Genetic Algorithms on the basis of as many experimentally determined structures as available. The efficiency of of this approach is demonstrated by the comparison of calculated and experimental data with respect to different shapes of the catalyst to their transformations into each other. Predicted and observed energies are found to agree within a few kJ/mol. These results recommend the approach chosen as a tool for the rational design of organometallic catalysts.

1 Introduction

Life is based on the interaction of molecules. The regulation of the intricate processes of life is based on the capability of molecules to recognise each other. How can a molecule, which has no intellectual capability, recognise at all? In the molecular regime, recognition is based on the interaction between molecular surfaces. Imagine a ball rolling over a plane and imagine, that the plane has some sticky area. The ball will roll and roll until it comes to that sticky place where it will be stuck. In the language of molecular recognition this would be interpreted as if the ball had recognised the sticky area, or, vice versa, as if the sticky area had recognised the ball. The two have recognised each other.

Molecules have very many different kinds of highly selective sticking areas. As everybody knows from daily life, there are different types of glues for different combinations
of materials. Some types of glues will be highly adhesive for plastic materials while others will stick to metals and again others will be good for bricks. The propensity of all these types of glues to stick especially well to certain types of materials is based on the interaction of the gluing material with the surface of the materials, which have to be stuck together.

Likewise with molecules: The surfaces of molecules are structured such that they will be attractive to certain types of molecular surfaces, indifferent to some other kinds and even repulsive to yet other types. Attraction, indifference and repulsion result from so-called secondary interactions. Such interactions are far weaker than chemical bonds. This means, they are easily made and broken again at normal temperatures and only if a few of them are made at a time, the total of such contacts will be strong enough to survive thermal motion. Thermal motion is a condition met in nature. The game of molecular recognition is solved in nature by making a highly developed use of these weak secondary forces.

Our life relies upon those well-optimised chemical interactions in biological systems. Modern daily life relies as well to a large extent on the availability of man made chemical products. Such products are made by transforming natural products such as crude oil, silica or lime stone, to name just a few, into products which, in most cases have no resemblance to the sources which they have been made from. This kind of “transubstantiation”, which appears almost mystical at the first glance, is brought about by whole series of well-designed chemical reactions. The individual chemical processes used in industry are tuned such as to give the highest possible selectivity. Selectivity means that in a well-designed chemical process the input chemicals will transform selectively into just the one desired output chemical. In the ideal solution to the problem of selectivity, there would be no waste products, whatsoever, and the output chemical could be used without the need of any further sort of purification as such.

Such an ideal process describes the maximum obtainable economic profit as the minimum environmental costs. The processes, which are used in producing bulk chemicals, even though being painstakingly optimised, are generally far from this goal.

The chemical processes in living nature, on the other hand, are generally very close to this goal. They tend to work with 100% selectivity, with a minimum of energy consumption and a minimum of waste products. The chemical systems working in living systems have been optimised by the evolutionary process over billions of years. The tricks played by these natural systems achieve extreme selectivity and efficiency and rely upon the highly developed use of secondary interactions. The molecules of life are structured such that they are able to recognise each other. Chemicals are transported to specific places in a living organism by virtue of them being recognised as just the very kind of molecule which is needed at a specific place. At this specific place, which is often an enzyme, i.e. a biological catalyst, the molecules are transformed with 100% selectivity into products, which the body has need of. No waste, no unnecessary expenditure of energy. Chemistry of life has solved the selectivity problem to an extent which man made chemistry is in many cases still far off. However, chemists know how to approach the degree of perfection by which the chemistry in living systems is characterised. They know that they have to make use of secondary interactions between molecules, to cope with the problem.

Chemists have found that organometallic catalysts present an elegant solution to many of the major selectivity problems. In such organometallic catalysts a metal centre is embedded in a cavity made up from innocent and unreactive constituents which are able to
form a well structured but, nevertheless, flexible cover on one side of the metal. The metal is the reactive centre in these catalysts. It sits there and waits for a substrate molecule to approach it and as soon as the substrate molecule is close enough, the metal will start to “digest” it. Digesting means it will either cut it into pieces at specific positions and release the fragments or it will stick together two molecules and then release the product. By this way the metal is free to start a new catalytic cycle grabbing and digesting as soon as the products of this process are released.

Such a catalyst may work like a kind of sewing machine making long chain molecules from small substrate molecules in sewing one piece to the other again and again. High tech polymers are made by just this approach routinely, nowadays. In other types of metal mediated reactions which are brought about by the organometallic catalysts new functionalities are added to the substrate molecules at specific places in a highly specific orientation. The specificity of such organometallic catalysts relies upon the specificity of the secondary interactions between the scaffolding which builds up the periphery of the metal — chemists call this “the ligands” — and the substrate molecule. Only substrate molecules which fit well into the cavity formed by the scaffolding will have the chance to approach the metal close enough such as to undergo the reaction with it.

One may imagine the channel leading from the outside of the organometallic catalyst to the metal in its interior as a highly structured surface with sticky points there and there and repulsive points at other places. This highly structured channel will orient the substrate molecule such that there is only one way for it to approach the metal. This is the reason for the extremely high selectivity which may be achieved by organometallic catalysis. The development of such catalysts, the improvement and the development of rules which allow for a rational design of specific catalysts are topics which are at the centre of state of the art molecular chemistry. This statement is underlined by the fact that the 2001–Nobel Price in chemistry has been given to three chemists, all of them being intimately involved in the design and optimisation of organometallic catalysts.

The problem with designing specific catalysts has still, by and large, to be solved by trial and error methods. While it is often known how a specific metal activates a substrate molecule and how it brings about the appropriate transformation it is not so well known how the secondary interactions in the periphery of the metal will influence the kind of approach of the substrate molecule to the metal. The individual steps of the transformation occurring at the metal may, nowadays, be quite reliably modelled by quantum mechanical calculations. Secondary forces, however, are not easily computed by quantum mechanical approximations even by the most modern techniques. Chemists have to rely on other types of models if they wish to rationalise the effect of secondary interactions within these models. One of these approaches is based upon the idea that a molecule with all its different building blocks and linkages between them might be described as a kind of mechanical machine with balls and joints and gears quite like a mechanical machine in the visible world. In order to describe the response of such a mechanical machine on a stimulus it is necessary to have a precise plan of its architecture and a precise knowledge of the stiffness of the joints and of all the springs acting within it.

Transferring this idea to the description of the behaviour of the molecules is straightforward in so far as the architecture of molecules is well known to a very high degree of accuracy from structure analyses. Bonds between atoms have to be replaced by springs of a certain stiffness, or, to put it into other words, a bond between two atoms is simu-
lated by a potential curve centred at the average distance of this kind of bond, where the potential function is described by only a few parameters. In many cases the harmonic potential, characterised by just one single force constant, will do. The angles, subtended by two bonds which radiate from one and the same atom, the so called valence angles, may be treated by the same approach as may the torsion around the central bonds of a four atom arrangement, the so called torsion angle. The forces between two atoms which are not bonded to each other, may be modelled by Lenard–Jones–type potentials made up of attractive and repulsive potential terms.

The problem in describing molecules as Pico molecular machines results from the fact that the force constants describing the stiffness of all the flexible parts of a molecule are not known from the beginning. Nevertheless, sets of potential functions with the appropriate force constants have been developed which produce a consistent description of the properties of organic molecules within the force field approach. This type of “molecular mechanics” model has been developed into a reliable tool for the prediction of properties of organic molecules and is now routinely used in every day organic chemistry.

Modelling of organometallic compounds, on the other hand, is impeded by the fact that no reliable potentials are available for describing the interaction between a metal and its immediate neighbourhood, its ligands. Organometallic compounds which are active in catalysis will generally make bonds to carbon atoms, hydrogen atoms, nitrogen atoms and so on, but none of the relevant potentials is known a priori.

There is yet another problem in the application of molecular mechanics methods to organometallic molecules: Different from the constituents of organic molecules (carbon atoms, hydrogen atoms, nitrogen and oxygen atoms) for which only a very limited set of co–ordination types (linear, triangular, tetrahedral) is found in nature, metals show many different types of co–ordination. It is far more difficult, therefore, to standardise these potentials, since every individual type of atom in a specific type of surrounding will have to be described by a specific set of potential functions.

One way around this problem is to restrict the model to a specific class of compounds all of which having the same co–ordination number and type and to try to develop appropriate force field parameters for this specific type of compound. If the type of compound chosen is a reactive one and if sufficient experimental data are available for the structures and reactivities of this set of compounds one may try to develop a force field description of this set of molecules and then try to correlate the results of molecular mechanics calculations with observed properties. This approach has been taken in the project which is described here in very short terms.

2 The Chemical Problem

If an object in three–dimensional space consists of a minimum of four different sub–objects and if these sub–objects may be ranked according to one or the other of their properties, this object defines a sense of rotation in \( \mathbb{R}^3 \).

For this phenomenon the term chirality is used amongst chemists. Molecules are three–dimensional objects and they contain chemically different groups in different positions. If, as it is often the case, their symmetry is low, they define a definitive sense of rotation as soon as different groups are ranked.

This means, that such chiral molecules exist as pairs with the only difference between
the members of a pair lying in their right–handedness or left–handedness or, as chemists say, in their chirality. Many of the molecules which are active in biological systems are chiral molecules. Pharmaceuticals, which have to interact with the molecules of life, will cause different biological responses depending on their chirality. There are many examples of drugs where it is known that only one type of chirality causes a beneficial response while the other one, the enantiomeric form – as chemists say- may either have no detectable response at all or, even worse, a detrimental response. Lack of appreciation of this fact may cause severe problems, one such example being thalidomide where one enantiomere is extremely helpful while the other one is extremely teratogenic. The lack of knowledge of this fact has caused a lot of suffering. Another example where one enantiomere is biologically active while the other one is inactive, is the so called L–dopa but which is one of the major curatives in Parkinson’s disease. Only one enantiomere, as chemists say, the L–form, L–dopa, is active, while the other enantiomere, D–dopa, is inactive. It is clear, therefore, that for production of pharmaceutical compounds it is essential to make use of selective synthetic processes which selectively produce only the desired enantiomeric form of a compound.

The solution to this kind of problem came about only in the last decades. Organometallic catalysts of pre–designed chirality are the key to this solution. The work discussed in this paper has its roots in just this field. During the preparative synthetic work of the group it had been found that certain chiral ligands are easily prepared in just the desired enantiomeric form and that these ligands may engage in bonding to a rhodium centre. As it was known that rhodium compounds are active catalysts in the production of compounds like L–dopa the catalytic properties of such rhodium compounds have been analysed. When the novel chiral rhodium compounds were used as catalysts, the following observation was made: Starting from achiral unsaturated substrates, saturation — i.e. addition of two hydrogen atoms to the unsaturated substrate — occurs with a selectivity of up to 91% with respect to the formation of just one enantiomere of the product⁴.

The metal–ligand part consists of the rhodium centre and the chiral ligand and the co–ligand part consists of an eight–membered hydro–carbon cycle — during catalysis the metal–ligand part persists while the place occupied by the co–ligand (Fig. 1) will hold the reaction species. Within the metal ligand part the rhodium atom, the two phosphorus atoms and the three carbon atoms of the chain make up a six–membered cycle. This cycle is chiral since at the central carbon atom symmetry is broken by the presence of one OH–substituent and one H–substituent at this atom (Fig. 1). The two phosphorus atoms contain two further substituents each. These substituents are again different, one substituent group consisting of so called mesityl residues (i. e. phenyl residues substituted by three methyl groups Me) in a symmetrical way, the other substituent group consisting of phenyl entities (Fig. 1).

Within the part containing the co–ligand consists of an eight–membered carbon hydrogen cycle which is bonded to the rhodium centre by the interaction between the rhodium atom and its double bonds (Fig. 1). This eight–membered cycle is just kind of a protective group. During the catalytic process itself this cycle is replaced by substrate molecules and hydrogen. The catalytic species which does the real work is thus different from the stable and fully characterised compound and this is the reason for the convention followed by chemists to call such a compound a pre–catalyst as it is a molecule from which the active catalyst will be formed during the catalytic process itself.
Figure 1. Chemical formula of the catalytically active chiral rhodium compound. The three–dimensional architecture and the transition between geometrically different forms of this compound were predicted by force field methods in quantitative agreement with experimental data. The “pre–catalyst” molecule consists of two parts — a metal ligand part containing the rhodium centre, the two phosphorus atoms and all the groups linked to them. This ligand part stays intact during the catalytic cycle. The other part of the compound contains an eight–membered hydro–carbon cycle as a protective group which is replaced by the reactants during catalysis. The six–membered cycle formed by the ligand and the metal is capable of existing in different forms. The forms shown (λ–twist, δ–twist, bottom) are both chiral and are the low–energy “conformations” of this cycle.

Figure 2. Projection of the two low–energy structures of the catalytic compound which differ in the conformation of the metal–ligand cycle (λ–twist form, δ–twist form).

The interesting part of the molecule is the six–membered cycle (Fig. 1, bottom). This six–membered cycle exists in two enantiomeric forms, as structure analyses and NMR–analyses clearly demonstrate. The two forms differ in the type of twist of their scaffolding and are chiral, therefore. Were it not for the chiral substitution of the central carbon atom (HO, H, CH$_2$PMes$_2$, CH$_2$PPh$_2$) the two forms — which chemists call λ– and δ–forms — would of course be of equal energy. The given sense of chirality at the central carbon atom which has been pre–determined by the synthesis of the ligand will create an energy difference between these two forms. Imagine a pair of hands of two partners. Both right hands fit nicely together as do both left hands. But there is a kind of misfit between the left hand of one partner and the right hand of the other one or vice versa. The fit and
the forces acting between these two chiral objects is different for these two types of pairs.
The energy difference between the two different twist forms of the six–membered cycle —
caused by the chirality around the central carbon atom of the chain — is still small enough
to be overcome by the thermal energy at ambient temperatures ($\Delta E \approx 4 \text{kJ/mol}$) and is also
small enough to allow for the co–crystallisation of both forms in one and the same crystal
(Fig. 2).

Even if the energy difference between the enantiomeric species is as low as 4 kJ/mol
the selectivity in imprinting the chirality of the catalyst molecule to the substrate upon re-
action with it is up to 91%. Any understanding of how this high selectivity comes about
must be based on an understanding of the flexibility of the six–membered cycle and the
motions of the phosphorus substituents connected with it. To this end a detailed exper-
imental NMR–Study was performed by which the structure of the compound in solution
and the inter–conversion of the $\lambda$– and $\delta$–forms of the chelate cycle with all it’s geometrical
and energetical implications has been extracted in a quantitative way$^{10}$. These experimental
results form a sound basis for evaluating the results of any computational approach. In
order to solve, at least, part of the chemical problem, namely to rationalise the structures
and the flexibility of the pre–catalyst, a force field model for this set of compounds was
developed$^6$.

3 The Computational Problem

A force field model tries to describe a molecule in the sense of a mechanical machine.
While the force field — i.e. the potentials to describe the properties correctly — for the
organic part of the pre–catalyst is well known and sophistically evaluated, that part of the
force field which describes the interaction of the metal with the ligands, is unknown from
the beginning.

The necessary condition for any sensible force field description of this part is that this
description has to reproduce the observed structures of compounds of the class to which the
pre–catalyst belongs. Even if this condition is not necessarily sufficient for the prediction
of energetic differences between different geometric forms of one an the same molecule a
good agreement may be found that it might well be so. An approach was chosen therefore,
which consists in optimising the force constants of all those potential functions to which
the metal is an immediate contributor. To make this approach work it is necessary to
determine as many structures within a given family of compounds such as to have enough
variance in the relevant data and a broad enough data base to make such a refinement
process appropriate.

In a series of papers$^5$ referring to the statistical basis of this approach as well as to
the application of this approach to organometallic molecules, it has been shown that the
results obtained this way are well in accord with the observations, both with respect to the
prediction of the shape of molecules and to the prediction of energy differences between
different forms of molecules.

The basic mathematical problem opposing this kind of approach is the lack of any op-
timisation method which could lead to an optimal solution with certainty within a finite
number of steps. While “global optimisation” cannot be brought about by any mathemati-
cal method in its true sense with certainty, optimisation by Genetic Algorithms is a highly
efficient approach with this kind of optimisation problem. In fact the approximation of
Genetic Algorithms as well as the merits of Neural Network Analysis in this field had first been demonstrated in the series of papers cited. The optimisation problem consists in finding the optimal values of all the force constants which imply a contribution by the metal for as many compounds of a given class as possible at a time.

The evaluation criterion has to be based on some measure of the distance between the computed structure and the observed structure in n–dimensional space. By using Genetic Algorithms as the optimisation tool this evaluation has to be done for any compound in the data set for any parameter set within a population over all generations over and over again.

This is a heavy load even for nowadays computing machines and is only feasible in a parallel computing environment. The great advantage of the Genetic Algorithm in this respect is that it is naturally parallel. This means that the evaluation of fitness, which is the really time consuming part of the procedure since it calls for a full force field optimisation in each step — may be performed at different processors at a time. Collecting the evaluation results and setting up a next generation for the Genetic Algorithm needs practically no time.

In order to perform the necessary computations, a force field program written in "C" by Tan was appropriately modified and embedded in different shells which allow its use in parallel computing environments. The Genetic Algorithm program package was taken from "PGAPack". Parallelisation was achieved by making use of the "MPI" library.

By using this type of approach, force field parameters for the description of all those interactions involving the rhodium atom in compounds of the type of the pre–catalyst were optimised on the basis of 11 crystallographically determined structures of compounds of this class. Based on these parameters a complete search in the conformational space of the pre–catalyst was performed.

Figure 3. Definition of the conformational space of the catalytically active compound. After optimising the force field model on the basis of 11 compounds of this general type a full grid search in the eight–dimensional conformational space was defined by $\phi_1 - \phi_5$ and $\tau_1 - \tau_3$ performed.

The conformational space was defined by the rotational positions of the aryl groups ($\phi_1 - \phi_4$, Fig. 3), a set of torsion angles within the six–membered cycle ($\tau_1 - \tau_3$, Fig. 3) and the rotational position of the OH–group ($\phi_5$, Fig. 3). A complete search within this eight–dimensional conformational space is again time consuming and is greatly speeded.
up by the use of parallel computers. Sorting the 10000s of results, such that a kind of adiabatic hyper-surface may be constructed, is an organisational problem which has been solved by writing the appropriate code.

Figure 4. Interconversion of $\lambda$– and $\delta$–forms of the catalytically active compound is accompanied by a strictly coupled rotation of the aryl groups at $\phi_3/\phi_4$. The projection of the hyper-surface onto the $\phi_3/\phi_4$ co-ordinate plane reveals that two energetically slightly different low-energy pathways exist. This coupled prediction of a rotation is in full agreement with the experimental findings.

The result of this analysis is shown in figures 4 and 5. Figure 4 shows a projection of the relevant hyper-surface onto the $\phi_3/\phi_4$ co-ordinate plane. $\phi_3$ and $\phi_4$ refer to the rotation of the mesityl substituents. It is seen (Fig. 4) that the two mesityl groups rotate in a strictly coupled manner such that if $\phi_3$ rotates in a clockwise sense, the rotation of $\phi_4$ is counter-clockwise. The two rotational pathways: $\phi_3$ clockwise $\phi_4$ counter-clockwise, (right/left in Fig. 4 label D) and the opposite sense rotation: $\phi_3$ counter-clockwise, $\phi_4$ clockwise (left/right in Fig. 4, label L) are of slightly different energy (Fig. 4). This finding is in complete agreement with the NMR observations$^{10}$.

By the complete analysis of the eight-dimensional conformational space of the pre-catalyst it was found that the $\delta$–form of the cycle is 3.1 kJ/mol more stable than the $\lambda$–form (Fig. 5) in almost numerical agreement with the experimental value of 3.4 kJ/mol for this energy difference. The least energy pathway of this type of transformation was as well analysed by force field methods and by experimental NMR techniques. An activation energy of around 69 kJ/mol was calculated — again in excellent agreement with the experimental value of 64.4 kJ/mol (Fig. 5).

4 Conclusion

Highly reliable and predictive force field models describing the behaviour of organometallic compounds may be developed on the basis of the relevant structural data of as many
The force constants describing the interactions with the metal as derived by their optimisation using genetic algorithms together with the known force field parameters of the organic parts of such molecules are able, when fed in to the appropriate force field program, to reproduce the static as well as the conformational behaviour of such molecules in excellent agreement between experimental and calculated data.

With respect to the documented high reliability such specific force fields are an efficient tool for the detailed analysis of catalytic processes mediated by organometallic catalysts.

5 Acknowledgment

The authors wish to thank the German Science Foundation and the “Fonds der Chemischen Industrie” for substantial funding and the grant for computing time by the “Interdisciplinary Centre for Scientific Computing (IWR)”, Heidelberg and the “John von Neumann Institute for Computing (NIC)”, Jülich. All force field developments were performed on a Parsytec GC Power-Plus-192 at the IWR in Heidelberg and the conformational space analysis on a CRAY T3E at the NIC in Jülich using 5000 h of CPU time.

References


Efficient Calculation of Rovibrational Energy Levels of General Tetratomic Molecules

Mirjana Mladenović

Institut für Physikalische Chemie
Universität Göttingen
Tammannstr. 6, 37077 Göttingen, Germany
E-mail: mmladen@gwdg.de

Computational methods for the determination of bound rovibrational energy levels of tetratomic molecules are presented. The methods are designed for orthogonal internal coordinates in the body-fixed formulation. The discrete variable representation and sequential inclusion of the internal degrees of freedom are employed in combination with the successive diagonalization-truncation technique. This procedure allows the construction of the final full-dimensional Hamiltonian matrix of relatively modest size, which is diagonalized with the help of conventional dense matrix algorithms to give both rovibrational energies and wave functions. The characterization of the molecular energy levels calculated by the methods described here is done in an automatic fashion owing to the availability of the wave functions.

1 Introduction

Quantum mechanical investigations of the spectral and dynamical behavior of chemical systems are important for attaining a deeper theoretical understanding of chemistry at the molecular level. High accuracy quantum mechanical calculations over a wide range of molecular energies afford useful insights e.g. into the nature of anharmonicity and intermode coupling effects. Since all of the information that can be known about a quantum state are contained in the wave function, the determination of the eigenvectors (wave functions) simultaneously with the determination of the eigenvalues (energies) is essential for finding a reliable state/process characterization by the theoretical means. The decoding of the information carried by the wave function is a problem of great importance in physical chemistry.

Four-particle systems are a challenging computational task since several million basis functions may easily be required to achieve a good description of the internal vibrational motion. For improperly designed basis, the computation can easily suffer from basis set explosion. The rate determining steps in large-scale calculations for tetratomic molecules are the evaluation of a six-dimensional potential energy function and the diagonalization of the Hamiltonian matrix. While the kinetic energy matrix elements are readily evaluated analytically, the integrals over the six-dimensional potential energy are in general solved with help of numerical quadrature rules. The matrix diagonalization is performed either by means of a standard (Householder) algorithm in conjunction with numerical schemes for the construction of a very compact basis\(^1\)\(^-\)\(^5\) or by employing a Lanczos algorithm in approaches favoring uncontracted primary basis set functions\(^6\)\(^-\)\(^8\). The eigenvalue calculation by means of methods based on a Lanczos algorithm requires no storage of the Hamiltonian matrix. The storage of Lanczos vectors is, however, necessary for the wave function determination.
Highly efficient theoretical methods for the accurate quantum mechanical calculation of a large number of the bound rovibrational energy levels (eigenvalues and eigenvectors) of general tetratomic molecules are developed by the author. The methods, termed DVR(+R)+FBR, DVR(+Rd)+FBR, DVR(6), and DVR(3)+DGB, employ the discrete variable representation of several vibrational coordinates in combination with the sequential diagonalization-truncation scheme and involve no dynamical approximation. Their unique feature is a very detailed eigenfunction analysis based on vibrationally averaged geometries and several adiabatic projection schemes for the automatic energy level characterization, which are carefully designed with the purpose of assigning approximate quantum numbers to exact quantum states. The energy level characterization approach based on the quantum state correlation scheme, which is recently developed for the DVR(6), is the first method treating this topic in a quantitative fashion.

The methods developed make use of the kinetic energy operator for orthogonal internal vectors in the body-fixed formulation (Sec. 2). The computational methodology is based on the discrete variable representation (Sec. 3). For the angular degrees of freedom, two basis sets are considered (Sec. 4) and their application for DVR based approaches is briefly described (Sec. 5).

2 Kinetic Energy Operator for Tetratomic Molecules

In the spherical polar parametrization, the internal geometry of a rotating tetratomic molecule described by three internal vectors $\mathbf{R}, \mathbf{d}_1, \mathbf{d}_2$ is given in terms of the three radial coordinates $R, d_1, d_2$, the two bending angles $\theta_1, \theta_2$, and the dihedral angle $\chi$, as shown in Fig. 1. For an entirely orthogonal set of internal vectors, such as the (diatom-diatom)
vectors of Fig. 1(a), the body-fixed expression of the kinetic energy operator $\hat{T}$ attains the maximally separable form compactly written as \(^3\)

$$
\hat{T} = \hat{T}_{\text{vib}} + \hat{T}_{\text{rot}} + \hat{T}_{\text{rv}},
$$

(1)

where

$$
2 \hat{T}_{\text{vib}} = 2 \hat{T}_{\text{str}}(R, d_1, d_2) + f(R, d_2) d_2^2 
- f(R, d_1) \left[ \frac{\hbar^2}{\mu R^2} \left( \frac{\partial^2}{\partial \theta_1^2} + \cot \theta_1 \frac{\partial}{\partial \theta_1} \right) - \frac{1}{\sin^2 \theta_1} \left( \hat{J}_z - \hat{i}_{2z} \right)^2 \right],
$$

(2)

$$
2 \hat{T}_{\text{rot}} = \frac{1}{\mu_R R^2} \left[ \hat{J}^2 - 2 (\hat{J}_z - \hat{i}_{2z})^2 \right],
$$

(3)

$$
\hat{T}_{\text{rv}} = \frac{1}{\mu_R R^2} \left[ i \hbar \left( \hat{J}_y - \hat{i}_{2y} \right) \frac{\partial}{\partial \theta_1} + \cot \theta_1 \left( \hat{J}_x - \hat{i}_{2x} \right) \left( \hat{J}_z - \hat{i}_{2z} \right) - \hat{J}_x \hat{i}_{2x} - \hat{J}_y \hat{i}_{2y} - \hat{J}_z \hat{i}_{2z} \right].
$$

(4)

In the latter expressions, $\hat{T}_{\text{str}}$ is the kinetic energy of the stretching vibrations and $\mu_R, \mu_{d_1}, \mu_{d_2}$ are the reduced masses associated with the radial coordinates $R, d_1, d_2$. The inverse of the reduced mass associated with the bending vibration is

$$
f(R, d_i) = \frac{1}{\mu_R R^2} + \frac{1}{\mu_{d_i} d_i^2}.
$$

(5)

The vibrational angular momentum operator describing the orbiting of the vector $d_2$ in the body-fixed $(R, d_1)$ frame of Fig. 1(b) is denoted by $I_2$ and $l_2 = l_2(\theta_2, \chi)$. The body-fixed projections of the total angular momentum operator $\hat{J}$ are $\hat{J}_x, \hat{J}_y, \hat{J}_z$. The kinetic energy operator of Eq. (1) contains no radial-angular derivative cross terms. This feature is important for the construction of suitable potential adapted radial basis functions.

The compact expression of $\hat{T}$ from Eq. (1) is applicable to any set of three translationally and rotationally invariant orthogonal vectors used to describe the internal geometry of a tetratomic molecule. The reduced masses $\mu_R, \mu_{d_1}, \mu_{d_2}$ are the only vector specific parameters in Eq. (1). As shown before, $\mu_R, \mu_{d_1}, \mu_{d_2}$ are given by simple analytical forms and can easily be calculated for a chosen set of internal vectors. The uniqueness of $\hat{T}$ allows the application of a single computer program to different classes of four-particle systems. This feature is a distinct practical advantage important in variational energy level calculations.

### 3 Discrete Variable Representation

In the one-dimensional case, the Hamiltonian matrix elements $H_{ij}$ in the orthogonal basis $\{ \phi_i \}$ are given by

$$
H_{ij} = \int \phi_i \hat{H} \phi_j \ dx = T_{ij} + V_{ij} = T_{ij} + \sum_z T_{iz} V(x_z) T_{jz}.
$$

(6)

Here, the integrals over the potential energy $V$ are solved numerically by employing the integration nodes $x_z$ and the integration weights $w_z$, such that $T_{iz}$ equals to $\sqrt{w_z} \phi_i(x_z)$. The representation of the Hamiltonian matrix in the $\phi$-basis of the finite size $N$ is known as the $N$-dimensional finite basis representation (FBR). For the numerical integration of Eq. (6) performed by means of the appropriate $N$-point quadrature rule, the $N$-dimensional
FBR is related by the $N \times N$ unitary FBR-DVR transformation matrix $T = \{ T_{iz} \}$ to the discrete variable representation (DVR) as follows

$$(DVR) \mathcal{H}_{z_1 z_2} = \sum_{i,j} T_{iz}^{(FBR)} T_{ij}^{(DVR)} + \delta_{z_1 z_2} (DVR) V(x_z).$$

In the DVR, the potential energy matrix is diagonal with respect to the DVR points $\{ x_z \}$.

As seen in Eq. (7), the integrals over $V$ in the DVR become a number $V(x_z)$ giving a local value of $V$ at the DVR point $x_z$. This is a distinct advantage of the DVR over the FBR counterpart important from the computational viewpoint. Since the DVR allows an efficient distribution of actual numerical work for the evaluation of $V$, DVR based methods are often preferable approaches for solving the Schrödinger equation due to their flexibility and easy implementation. Theoretical methods involving the discrete variable representation are originally introduced by Light$^9,10$. The DVR has evolved into a standard approach for the quantum mechanical calculation of energy spectra of triatomic molecules$^9,11,12$.

The integration nodes and weights of Gauss quadrature rules are calculated by diagonalizing the Jacobi matrix constructed from the three-term recurrence relation$^13$. The dual orthogonality of $N$ DVR points is insured by the Christoffel-Darboux identity$^14$.

4 Angular Basis for Tetratomic Molecules

We consider two types of angular basis functions for tetratomic molecules. The parity-adapted angular basis functions, termed Basis 1, are of a non-direct product type and given for the total angular momentum $J$ by$^3$

$$A^{Jp}_{Kkj} = N_{Kk} \beta_j^{k-K} \left[ Y_j^k |J, K, M\rangle + (-1)^{J+K+k+p} Y_l^{-k} |J, -K, M\rangle \right]$$

$$B^{Jp}_{Kkj} = N_{Kk} \beta_j^{k+K} \left[ Y_j^k |J, -K, M\rangle + (-1)^{J+K+k+p} Y_l^{-k} |J, K, M\rangle \right]$$

where $Y_j^k = Y_j^k(\theta_2, \chi)$ are spherical harmonics, $\beta_j^{k-}$ normalized associated Legendre functions in $\cos \theta_1$, and $|J, K, M\rangle$ symmetric top eigenfunctions. The quantum numbers of the projection of $J$ onto the $Z$-axis of the space-fixed and onto the $z$-axis of the body-fixed frame are $M$ and $K$, respectively. The quantum number associated with $\hat{l}_z$ is $k$. The basis functions of Eq. (8) allow easy analytical evaluation of the kinetic energy matrix elements and guarantee their finite size in spite of singular terms in Eq. (2) that are divergent as $\theta_i$ goes either to 0 or $\pi$.

Total angular momentum $J$, its space-fixed projection $M$, and parity $p$ are strictly conserved quantum numbers for eigenstates of a tetratomic molecule, i.e. all Hamiltonian matrix elements are diagonal in $J$, $M$, and $p$. For the basis of Eq. (8), the kinetic energy matrix elements are also diagonal in $j$ and $l$ and nonzero for $\Delta K = \pm 1$ and $\Delta k = \pm 1$, such that the kinetic energy matrix is sparse after the angular integration$^3$. The potential energy matrix in Basis 1 factors into two parity-adapted blocks. The matrix element of $V$
between e.g. the two $B_{Kji}^{Jp}$ functions reads in the FBR as

$$
V_{k_2j_2l_2}^{k_1j_1l_1}(R, d_1, d_2) = N_{Kk_1} N_{Kk_2} (-1)^{k_1+k_2} (1/\pi) \times \int_{-1}^{1} \int_{-1}^{1} P_{l_1}^k (x_2) P_{l_2}^{k+K} (x_1) d x_1 d x_2 \\
\times \int_0^{2\pi} \cos(k_2 - k_1) \chi d \chi V(R, d_1, d_2, \theta_1, \theta_2, \chi),
$$

where $x_i = \cos \theta_i$. In the latter equation, the associated Legendre functions of different order are coupled due to the dependence of the potential energy on the torsional coordinate $\chi$ and the design of the DVR in connection with Basis 1 is not straightforward.

The parity-adapted angular basis of a direct product type, called Basis 2, is written for $J = 0$ as

$$
P_{n_1}^{(\alpha_1, \beta_1)} (\cos \theta_1) P_{n_2}^{(\alpha_2, \beta_2)} (\cos \theta_2) C_k (\chi),
$$

where $P_{n_i}^{(\alpha_i, \beta_i)}$ stands for orthonormalized Jacobi functions, whereas $C_k$ denotes Chebyshev polynomials $T_k$ of the first kind and Chebyshev functions $U_k$ of the second kind used respectively for even-parity and odd-parity calculation. Since the order $(\alpha_i, \beta_i)$ of the Jacobi functions is decoupled from $k$ in Eq. (10), Basis 2 is readily adapted for the DVR.

### 5 Computational Strategy

Several computational strategies are developed for tetratomic molecules. The DVR(+R)+FBR and DVR(+Rd)+FBR approaches employ Basis 1, whereas the DVR(6) and DVR(3)+DGB make use of Basis 2. All four methods employ a contraction scheme resulting from several diagonalization/truncation steps to construct a very compact and efficient basis set, in which the final matrix is of modest size and can be diagonalized by standard routines which, in addition to energies, provide wave functions in a simple fashion. Wave functions are indispensable for the calculation of various quantities, such as vibrationally averaged geometries, whose values are instructive for making quick preliminary quantum number assignments to the calculated states.

The fact that all off-diagonal matrix elements of the kinetic energy operator from Eq. (1) are attributed a single factor $1/2 \mu R^2$ from the radial space is utilized in the DVR(+R)+FBR and DVR(+Rd)+FBR in conjunction with a pointwise representation of the radial coordinate $R$. The DVR(+R)+FBR and DVR(+Rd)+FBR methods are favorable for the computation of rovibrational ($J \neq 0$) energies of any tetratomic molecule and for the vibrational calculation for linear tetratomic molecules since Basis 1 avoids the possible singularities in the kinetic energy matrix. The DVR(6) and DVR(3)+DGB methodologies are advantageous for nonlinear tetratomic molecules since the discretization of the entire three-dimensional angular space allows a very efficient distribution of numerical work involving the evaluation of the six-dimensional potential energy function.

The computational methods described here are employed for the calculation of (ro)vibrational energies of several tetratomic molecules, like HO\textsubscript{2}C, HF\textsubscript{2}, H\textsubscript{2}O\textsubscript{2}, H\textsubscript{2}C\textsubscript{2}, and H\textsubscript{2}CO. The high efficiency and flexibility of our approaches have allowed the determination of various quantities, such as spectroscopic constants, dissociation energies, and the
number of bound states, as well as to study characteristic energy level patterns in various parts of the energy spectrum.\[^2\text{-5}\].

5.1 Computational Strategy for Basis 1

The DVR(+R)+FBR method makes use of combined discrete variable and finite basis representations of the angular coordinates \((\theta_1, \theta_2, \chi)\), a discretized Jacobi distance \(R\), and eigenfunctions for \((d_1, d_2)\) in conjunction with several diagonalization/truncation steps. The pointwise representation of the bending angles \(\theta_1, \theta_2\) and the FBR for the torsional coordinate \(\chi\) are introduced as the primary description in the (DVR+FBR) of the three-dimensional angular space \((\theta_1, \theta_2, \chi)\).

In the (DVR+FBR) of the two-dimensional space \((\theta, \chi)\), the FBR-DVR transformation matrix \(T\) for \(\theta\) is given by the following direct sum

\[
T = \sum_m \oplus T^m.
\]

The elements of the submatrix \(T^m\) are

\[
T^m_{l_1z_1} = \sqrt{\omega_{z_1m}} P_{l_1z_1}^{m}(z_1m),
\]

where \(z_1m = \cos \theta_1\) and \(\omega_{z_1m}\) are the \(i\)-th DVR point and weight for Gauss-Legendre quadrature of order \(m\). The matrix \(T\) is orthogonal and block diagonal in the (DVR+FBR).

The similarity transformation by \(T\) changes the finite basis representation in the functional space spanned by \(Y_m^l\) to the combined (DVR+FBR) specified by \((m, z_1m)\). The (DVR+FBR) matrix element of \(V = V(\theta, \chi)\) becomes

\[
\langle z_1m_1, l_1m_1 | V | z_2m_2, l_2m_2 \rangle = \sum_{l_1, l_2} T^m_{l_1z_1} \langle l_1m_1 | V | l_2m_2 \rangle T^m_{l_2z_2}
\]

\[
\approx V^{m_1, m_2}(z_1m_1) R^{z_1m_1}_{z_2m_2}
\]

when the \(\theta\)-integration in the FBR is performed by Gauss-Legendre quadrature of order \(m\) selected such that \(m = \min(m_1, m_2)\). The auxiliary matrix \(R\) is defined by

\[
R^{z_1m_1}_{z_2m_2} = \sum_{l_2=m_2} T^m_{l_2z_2} \sqrt{\omega_{z_1m_1}} P_{l_2}^{m_2}(z_1m_1).
\]

In Eq. (13), the integral over \(\chi\)

\[
V^{m_1, m_2}(z_1m_1) \sim \int_0^{2\pi} V(z_1m_1, \varphi) e^{i(m_2-m_1)\varphi} d\varphi
\]

can be solved by means of Gauss-Chebyshev quadrature for each of the \((m_1, m_2)\) pairs at the DVR \(\theta\)-point \(z_1m_1\).

The potential energy DVR matrix element of Eq. (13) is diagonal in \(z\) only for \(m_1 = m_2\). One may, however, note that a single evaluation of the potential energy at the point \(z_1m_1\) is required in Eq. (13) for the calculation of all matrix elements \(\langle z_1m_1, l_1m_1 | V | z_2m_2, l_2m_2 \rangle\) with variable \(z_2m_2\). In actual calculation, the matrices \(T\) and \(R\) can be precomputed and retained for later use.

In the case of tetratomic molecules, the matrix element of the potential energy in the DVR comprised of the \(\theta_1\) grid points \(a_{i,k}\) and \(\theta_2\) grid points \(z_{j,k}\) can be calculated from

\[
90
\]
the corresponding FBR result of Eq. (9) by employing the transformation matrices $T_1$ and $T_2$ defined in analogy to Eq. (11). For $J = 0$, we obtain

$$V_{k_2z_2a_2k_2}^{k_1z_1a_1d_1} = N_{K_{k_1}}N_{K_{k_2}}(-1)^{k_1+k_2} \delta_{a_1a_2} R_{z_2k_2}^{k_2} R_{a_2k_2}^{a_2k_1} 2/M \times \sum_g \cos(k_2 - k_1) \chi_g \sum_{z_1,x_2} C_{x_1,x_2}^a C_{x_1,x_2}^a V(R_\alpha, x_1, x_2, a_1k_1, z_1k_1, \chi_\alpha).$$ (16)

In the latter equation, the $\chi$-integration is done by means of the $M$-point Gauss-Chebyshev quadrature involving the nodes $\chi_g$. The expansion coefficients of the two-dimensional $(d_1, d_2)$ eigenfunction in the discrete basis $x_1, x_2$ are denoted by $C_{x_1,x_2}^d$ in Eq. (16).

The discrete points $R_\alpha, x_1, x_2$ used in Eq. (16) are obtained by diagonalizing the appropriate coordinate in an eigenfunction basis precomputed for a chosen reference potential. The matrix elements of $1/\mu_i d_i^2$ occurring in the reduced mass for the bending vibration, Eq. (5), are precomputed for the 2D $(d_1, d_2)$ eigenfunctions and saved for later use.

The grid points $(x_1, x_2)$ are employed as a primary basis for $(d_1, d_2)$ in the DVR(+R)+FBR. The potential energy matrix element given by

$$V_{k_2z_2a_2x_1x_2}^{k_1z_1a_1x_1x_2} = N_{K_{k_1}}N_{K_{k_2}}(-1)^{k_1+k_2} \delta_{a_1a_2} \delta_{x_1x_1} \delta_{x_2x_2} R_{z_2}^{k_2} R_{a_2}^{a_2k_1} 2/M \times \sum_g \cos(k_2 - k_1) \chi_g V(R_\alpha, x_1, x_2, a_1k_1, z_1k_1, \chi_\alpha)$$ (17)

for the DVR(+R)+FBR are simpler than the corresponding expression from Eq. (16) for the DVR(+R)+FBR due to the elimination of the similarity transformation by the matrix $C$.

The performance of the DVR(+R)+FBR algorithm appears competitive with approaches based on the Lanczos method. Our calculation2,3 for the vibrational levels of (HF)$_2$ and HOCO took respectively 15 CPU min and 9 CPU h on a CRAY T90 (single processor). Wu et al.7 obtained the lowest 40 bound state energies and the associated wave functions for (HF)$_2$ in 10.5 CPU min on 126 processors of a CRAY T3E. The calculation by Lehoucq et al.6 gave only the level energies of HOCO in 10.7 CPU h on 54 nodes of an IBM SP. The potential energy surface due to Quack and Suhm15 is used for (HF)$_2$ and due to Schatz, Fitzcharls, and Harding16 for HOCO.

### 5.2 Computational Strategy for Basis 2

The DVR(6) and DVR(3)+DGB methods make use of the discrete variable representation of the three angular coordinates in combination with Basis 2. The angular degrees of freedom are taken sequentially into account. After adding a new angular coordinate, the higher-dimensional Hamiltonian matrix is calculated and diagonalized. Truncation of the basis follows each diagonalization. The procedure of the DVR(6) and DVR(3)+DGB is schematically shown in Fig. 2. The grid points in $\theta_1, \theta_2, \chi$ are denoted by $a, z, g$, respectively.

The three radial degrees of freedom are treated simultaneously for each of the angular triples $(a, z, g)$ in Step 1. The eigenvalue problems for the fixed-angle three-mode Hamiltonian for the stretching vibrations

$$T_{str}(R, d_1, d_2) + V(R, d_1, d_2, \theta_1, \theta_2, \chi)$$ (18)
are solved by employing 3D radial DVRs in the DVR(6) approach and real three-dimensional distributed Gaussian functions in the DVR(3)+DGB method. After solving the 3D radial eigenvalue problems, the primary radial basis is reduced by retaining only those 3D eigenvectors $|k\rangle$ of energy smaller than a preselected energy cutoff parameter $E_{a_{\text{max}}}^a$.

In Step 2, the bending coordinate $\theta_1$ is coupled to the radial degrees of freedom. The 4D Hamiltonian matrices diagonal with respect to both $z$ and $g$ are computed for the basis $|k\rangle|a\rangle$. Having solved the 4D eigenvalue problems, the 4D basis is truncated by retaining only 4D eigenvectors $|j\rangle$ associated with the eigenvalues less than a predefined energy cutoff $E_{a_{\text{max}}}^a$. In Step 3, the 5D eigenvalue problems comprising the second bending coordinate $\theta_2$ are formulated in the basis composed of $|j\rangle|z\rangle$. The number of the 5D eigenvectors $|i\rangle$ saved after the 5D diagonalizations is controlled by a preselected parameter $E_{a_{\text{max}}}^g$. The final 6D Hamiltonian matrix for all terms included is computed in the basis of $|i\rangle|g\rangle$ and diagonalized to give level energies and wave functions with the help of conventional algorithms for symmetric matrices.

The angular DVR in conjunction with sequential diagonalization/truncation procedure allows the development of schemes for automatic quantum number assignments for triatomic$^{11,12}$ and tetratomic$^4,5$ molecules. The strategy of the DVR(6) and DVR(3)+DGB methods readily permits adiabatic torsion approximation since the five-mode (stretch+bend) eigenenergies $i$, calculated at chosen torsional DVR points $\{g\}$ in Step 3 of Fig. 2, form effective (adiabatic) potentials $^{adi}V^i$ for the torsion. The effective potentials $^{adi}V^i$ are constructed by employing the quantum state character correlation scheme$^4$ and used to obtain torsional levels adiabatically separated from the stretch+bend motion. The determination of the adiabatic (zero-order) eigenvectors does not involve any new matrix element beyond those already evaluated. Automatic quantum number assignments to the calculated levels are obtained by locating the dominant zero-order contribution in the adiabatic expansions of the accurate full-dimensional wave functions. Adiabatic expansions of the numerically exact wave functions are obtained by simple matrix multi-

Figure 2. Computational strategy of the DVR(6) and DVR(3)+DGB methods.
The adiabatic projection methods are used to study the intramolecular dynamics of H$_2$O$_2$ and HOCO, as well as for the investigation of local-mode vs normal-mode phenomena in the excited stretching states of H$_2$O$_2$ and H$_2$CO.

The integration and energy cutoff parameters used for the calculation of the vibrational energy levels of H$_2$CO, H$_2$O$_2$, and HOCO by means of the DVR(6) method are summarized in Table 1. The primary basis comprising 1 000 000-5 000 000 functions is reduced by $E_a^\text{max}$, $E_z^\text{max}$, and $E_g^\text{max}$ to the size $N_{\text{final}}$ of 4 000-5 000 for the final full-dimensional vibrational Hamiltonian matrix in Table 1. The CPU times given in Table 1 refer to the calculation of eigenenergies with eigenvectors and to the eigenfunction analysis for a large number of the calculated states. The potential energy surface of Carter, Handy, and De- maison$^1$ is employed for H$_2$CO, of Koput, Carter, and Handy$^{17}$ for H$_2$O$_2$, and of Schatz, Fitzcharls, and Harding$^{16}$ for HOCO.

Acknowledgments

The author is grateful to Prof. P. Botschwina for his support and to the Deutsche Forschungsgemeinschaft (SFB 357). Access to the Cray-T90 vector computer of the John von Neumann - Institut für Computing (NIC) at Forschungszentrum Jülich through grant HGU03 is kindly acknowledged.

References


<table>
<thead>
<tr>
<th></th>
<th>H$_2$CO</th>
<th>H$_2$O$_2$</th>
<th>HOCO</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(K, N_1, N_2)$</td>
<td>(8, 12, 12)</td>
<td>(14, 12, 12)</td>
<td>(14, 15, 35)</td>
</tr>
<tr>
<td>$[M, M_1, M_2]$</td>
<td>[10,10,10]</td>
<td>[12,9,9]</td>
<td>[14, 8, 6]</td>
</tr>
<tr>
<td>$E_a^\text{max}$ (cm$^{-1}$)</td>
<td>23000</td>
<td>20000</td>
<td>3000</td>
</tr>
<tr>
<td>$E_z^\text{max}$ (cm$^{-1}$)</td>
<td>21500</td>
<td>20000</td>
<td>2000</td>
</tr>
<tr>
<td>$E_g^\text{max}$ (cm$^{-1}$)</td>
<td>20000</td>
<td>16000</td>
<td>-1000</td>
</tr>
<tr>
<td>$N_{\text{prim}}$</td>
<td>1 152 000</td>
<td>1 959 552</td>
<td>4 939 200</td>
</tr>
<tr>
<td>$N_{\text{final}}$</td>
<td>4 311</td>
<td>4 840</td>
<td>4 090</td>
</tr>
<tr>
<td>CPU (h)</td>
<td>2.0</td>
<td>3.8</td>
<td>7.9</td>
</tr>
</tbody>
</table>

Table 1. Numerical parameters employed for the DVR(6) calculation of the vibrational energy levels of H$_2$CO, H$_2$O$_2$, and HOCO. The number of the angular DVR $(\chi, \theta_2, \theta_1)$ points is denoted by $(K, N_1, N_2)$, whereas $[M, M_1, M_2]$ stands for the number of the radial DVR points in $[R, d_1, d_2]$. The energy cutoff parameters $E_a^\text{max}$, $E_z^\text{max}$, and $E_g^\text{max}$ are employed to truncate the primary basis set at different stages of the DVR(6) calculation. The size of the primary basis set and of the final 6D matrix are denoted by $N_{\text{prim}}$ and $N_{\text{final}}$, respectively. The CPU times are given for a CRAY T90 (single processor).
Computersimulation of Free Liquid Surfaces

Jens Oberbrodhage\textsuperscript{1} and Harald Morgner\textsuperscript{2}

Institute of Physical and Theoretical Chemistry
Department of Chemistry and Mineralogy
University Leipzig
Linnéstraße 2
04103 Leipzig, Germany
\textsuperscript{1} E-mail: oberbrod@rz.uni-leipzig.de
\textsuperscript{2} E-mail: hmorgner@rz.uni-leipzig.de

Structure and dynamics on the microscopic level are important for the understanding of reactions taking place in liquid surfaces and liquid / liquid interfaces. We use different experimental techniques which are briefly described: besides conventional methods for studying macroscopic properties like surface tension we use a spectroscopic approach to gain information on a molecular level.

After outlining this spectroscopic technique experimental results are introduced. They have motivated to the computer simulation studies described afterwards.

In the present article we focus on binary liquid mixtures, one of them being miscible (formamide and benzyl alcohol) the other one exhibiting a miscibility gap (formamide and hexane).

In both cases we examine whether experimental results can be supplemented by the simulation. A special interest deals with the question how much orientational and dynamical properties of one liquid are affected by the presence of a second species.

While many computer simulations examine bulk properties we present a special strategy to create a liquid lamella with two interfaces liquid / vacuum. This gives the opportunity to study molecular properties for different surroundings, namely the liquid bulk and the surface exposed to vacuum.

1 Introduction

A large part of research done in our group focusses on the characterization of liquid surfaces based on experimental techniques. The composition of the surface region, molecular orientation and the kinetics of adsorption processes are of interest.

Different experimental techniques are used to monitor these properties: a macroscopic method to study changes in the liquid surface due to surfactants is the static or dynamic surface tension (tensiometry). Using concepts developed in thermodynamics like the Gibbs adsorption equation such measurements lead to properties like the surface excess of the different components in a mixture.

Since these experiments do not trace features on a molecular scale a special interest in our group is the application of spectroscopic methods to elucidate system parameters on a molecular level.

One technique being an established tool to characterize solid surfaces due to its high surface sensitivity is electron spectroscopy: illumination of the sample leads to electron emission, which originate from a layer being only several Å up to a few nm thick. Analysis of the energy and angular distribution of these photoelectrons gives a handle on target composition, binding properties and structural information. Variation of photon energy and
electron emission angle allows recording of concentration depth profiles. In the early 70th
electron spectroscopy was used for studying liquid surfaces for the first time\(^1\) (X-ray Pho-
toelectron Spectroscopy XPS, ESCA). In 1986 MIES (Metastable Induced Electron Spec-
troscopy) was introduced as perfect surface sensitive type of electron spectroscopy\(^2\): the
projectile beam consists of excited helium atoms having an electronic energy of \(\sim 20eV\).
The thermal kinetic energy of \(\sim 50mV\) prevents the helium atoms from penetrating into
the target material. MIES is chemically sensitive, traces molecular orientation in the out-
ermost surface layer and allows quantitative analysis of particle ratios.

Electron spectroscopy averages over a certain time and sample volume. With our equip-
ment adsorption processes as function of the age of the liquid surface can be studied. The
regime accessible goes from \(\sim 2-25ms\) or from a few hundred \(ms\) up to a few \(s\), depend-
ing on the type of target arrangement choosen. This allows to monitor diffusion processes
perpendicular to the liquid surface. A lateral movement is not traced by the experiment but
can be evaluated from computer simulation. In case of an inner surface for polar / unpolar
systems with miscibility gap the high surface sensitivity of electron spectroscopy prevents
a study due to the limited mean free path of \(e^-\). In this case MD simulation is an ideal tool
to gain details of inner surfaces.

In the following section we first describe experimental studies that gave rise to our
simulations. Afterwards results of the computer simulations are outlined.

2 Motivation for the Simulation Due to Experimental Results

2.1 The Binary Liquid System Formamide / Benzyl Alcohol

The system formamide (FA) and benzyl alcohol (BA), which is miscible for molar fractions
ranging from 0 to 1, was investigated both with electron spectroscopy and with tensiometry.
Both methods reveal BA to be the surface active component, its surface tension being
\(39\text{ mN/m}\) compared to \(52\text{ mN/m}\) in case of FA.

With MIES which is sensitive for the molecular orientation it could be seen that both
species show preferential orientation in the liquid surface: The FA lies more or less flat
in the time average whereas BA tends to expose the benzene ring to vacuum. The surface
active BA starts to form a layer on top of the formamide when its concentration is increased.
It is assumed that this layer shows holes for a certain ratio BA/FA in which FA is embedded
with an orientation determined by the dipole moment of the benzyl alcohol.

The aim of our simulation study is to elucidate this ordering process at the interface
liquid / vacuum and trace differences between bulk and surface properties.

2.2 The Binary Liquid System Formamide / Hexane

The general interest in 2-phase-systems polar / unpolar with a miscibility gap with regard
to industrial application is expressed in many studies with water as polar component and a
second unpolar liquid.

In our experimental studies we have replaced water by formamide in order to bene-
fit from its lower vapor pressure. In two types of experiments we have investigated the
influence of a phase transfer catalyst (tetrabutylammoniumiodide = TBAI) onto the inter-
face formamide / alkanes. In both studies we find that the alkanes can form a closed thin
(mono)layer on the liquid formamide. Addition of TBAI destroys the closed alkane layer and leads to accumulation of both, TBAI and alkane, below the surface.

In addition to the limitation due to its time averaging the high surface sensitivity of electron spectroscopy prevents us from studying the interface polar / unpolar in usual 2-phase-systems which are macroscopically large. The only experimental approach using electron spectroscopy uses a very thin layer of alkane with a thickness not higher than the mean free path of the electrons which amounts to a few nm. Here MD-simulation is an appropriate method to overcome these experimental obstacle in order to model 2-phase-systems in general.

### 3 Special Features for Simulating Liquid Surfaces

A high percentage of computer simulation carried out deals with bulk properties. The use of three dimensional boundary conditions allows to study in principle macroscopically extended systems by looking at a simulation box of typically a few 100Å edge size.

In case of our simulation of liquid surfaces we follow a different strategy (fig. 1): we use a simulation box which has no cubic shape but is elongated in Z-direction (the later surface normal). A liquid slab is then centered in Z and periodically continued in X- and Y-direction. The Z-size is large enough to avoid interaction between the actual simulation box and its copy below and above.
The slab geometry has the advantage that two interfaces liquid / vacuum are created. This can either be used to improve statistics or one can imbed some surfactant in one surface in order to compare the pure surface with the enriched one. This gives the possibility of studying molecular properties in three different surroundings: the bulk liquid and two different interfaces liquid / vacuum.

On the other hand the extend in Z must not be too high in order to keep the number of vectors calculated in the Ewald summation as small as possible.

4 Simulation Results

4.1 Introduction

This section deals with results from our MD simulation. Parameters of the different systems are listed in tables 1, 2.

4.2 Binary System Formamide and Benzyl Alcohol

<table>
<thead>
<tr>
<th>molecular parameters</th>
<th>name</th>
<th>formular</th>
<th>sites</th>
<th>number</th>
</tr>
</thead>
<tbody>
<tr>
<td>species 1</td>
<td>benzyl alcohol</td>
<td>C₆H₅CH₂OH</td>
<td>9</td>
<td>420</td>
</tr>
<tr>
<td>species 2</td>
<td>formamide</td>
<td>HCONH₂</td>
<td>5</td>
<td>80</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>simulation parameters</th>
<th>program</th>
<th>computer</th>
<th>knots</th>
<th>Box X,Y,Z [Å]</th>
<th>timestep [fs]</th>
</tr>
</thead>
<tbody>
<tr>
<td>DL-POLY</td>
<td>CRAY T3E Jülich</td>
<td>32</td>
<td>44, 44, 120</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>local workstation</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>run parameters</th>
<th>simulated time [ps]</th>
<th>temperature [K]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1385</td>
<td></td>
<td>300</td>
</tr>
</tbody>
</table>

Table 1. MD-simulation benzyl alcohol + formamide, parameters.

We carried out simulations for both liquids and for the mixture\(^4,5,8\). This offers the opportunity to explore how much molecular features are influenced by the second component.

The experimental studies of BA reveal a strong preferential ordering of the benzyl alcohol in the surface region as well as an exceptionally large surface potential of several hundred \(mV\). According to experimental studies the surface molecules in the liquid BA orient in such a way that the benzene rings point towards the vacuum and the polar group is directed towards the bulk of the liquid.

The simulation confirms the experimental results concerning the orientation of surface molecules. The surface potential resulting from the simulation is \(350mV\). Thus sign and magnitude are in accordance with the experimental work.
An examination of the structure of hydrogen bonds by means of bond angle distribution and bond state distribution reveals that the hydrogen bond is nearly unchanged when switching from bulk to surface.

The orientation of surface molecules causes a specific dynamic behaviour of individual components of the molecule. We find a speed up of the reorientation of certain axes when going from bulk to surface which depends on the separation of these axes from the polar group. This originates from the fact that benzene ring of surface molecules experiences a less dense region than axes near the polar group or the polar group itself.

In addition the average local environment of a polar group for surface molecules is very similar to that of bulk molecules.

The simulation of a NVE-ensemble of pure formamide shows that on the average the molecular plane is oriented perpendicular to the liquid surface with N and O pointing towards the liquid bulk and the CH-group exposed to vacuum. At first glance this seems to be a contradiction to the experimental finding where the plane is mainly parallel to the surface (see sec. 2.1). This puzzle can be solved by keeping in mind that MD averages the molecular orientation over a broad distribution. In the experiment it is taken into account how long the molecules remain in a certain orientation. Since there is a higher probability to trace the molecule in the turning points of a rotation the orientation parallel is overemphasized in the spectroscopic data.
Furthermore it could be seen that the liquid structure is dominated by a H-bonded network reaching from bulk into the surface region. Close to the surface smaller parts of H-bonded molecules separate from the main cluster. Studying dynamic properties shows that movement around certain axes of formamide is slower in the surface region originating from the preferential orientation of the molecules.

The binary system consisting of 420 BA and 80 FA molecules shows that features of the pure component are not much influenced by the second species: The liquid surface consists preferentially of BA, benzene rings are exposed to vacuum, the hydroxy group points into the bulk. It is no rigid arrangement and the ordering does not continue into the bulk. FA molecules are found mainly in the neighbourhood of hydroxy groups below the first molecular layer. The close distance ordering results from H-bonds. Density profiles are shown in fig. 2.

4.3 Formamide and Hexane

<table>
<thead>
<tr>
<th>molecular parameter</th>
<th>name</th>
<th>formular</th>
<th>sites</th>
<th>number</th>
</tr>
</thead>
<tbody>
<tr>
<td>species 1</td>
<td>formamide</td>
<td>HCONH₂</td>
<td>5</td>
<td>432</td>
</tr>
<tr>
<td>species 2</td>
<td>hexane</td>
<td>CH₃(CH₂)₄CH₃</td>
<td>6</td>
<td>128</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>simulation parameters</th>
<th>program</th>
<th>computer</th>
<th>knots</th>
<th>Box X,Y,Z [Å]</th>
<th>timestep [fs]</th>
</tr>
</thead>
<tbody>
<tr>
<td>DL-POLY</td>
<td>local</td>
<td>workstation</td>
<td>32</td>
<td>38.5, 38.5, 120</td>
<td>0.75</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>run parameter</th>
<th>simulated time [ps]</th>
<th>temperature [K]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>3040</td>
<td>290</td>
</tr>
</tbody>
</table>

Table 2. MD-simulation formamide and hexane, parameters.

The binary system formamide and hexane (HX) is a typical 2-phase system showing macroscopic phase separation. We are interested in properties of the polar / unpolar interface and study what type of molecular features is influenced by the second component.

The simulation reveals that a sharp phase separation does not exist on an atomic level. Both liquids penetrate each other over a distance of ∼ 10 Å (fig. 3). The FA bulk molecules do not show any preferential orientation but in both interfaces (FA / vacuum and FA / HX) the known tilting of the C-N-backbone is found. This feature is not influenced by the unpolar liquid.

For HX the situation is somewhat different: analysis of the number of trans-configurations per molecule shows that the major part of the unpolar molecules is elongated in the outer surface (HX / vacuum). In the inner interface (HX / FA) the presence of FA leads to an increased amount of HX in coiled conformation (fig. 4).
5 General Remarks and Outlook

One general restriction of MD simulations even with nowadays computer power is the limited number of interaction sites and the small time interval accessible.

Experimentally we study segregation processes taking place in \( \text{msec} - \text{sec} \) depending on the respective surfactant. In contrast classical MD methods cover a time regime of several \( 100\text{ps} \) up to a few \( \text{ns} \). Since this means that diffusion itself can not be simulated we choose a different strategy: A particle ratio \( \text{surfactant} : \text{solvent} \) was taken from the experiment and reproduced in the liquid surface simulated. Afterwards the simulation was used to analyze properties like orientation processes which take place at a much shorter time scale.

In order to extend the simulation time window up to experimental values we plan to use the so called \textit{targeted dynamics} method\(^6\): movement or conformational changes of a molecule are accelerated by applying an additional constraint force along a given reaction coordinate (for example: distance from an aim point). We have already tested this approach for the adsorption of thiol molecules at a gold surface\(^7\) and for the movement of a single formamide molecule through a slab of benzyl alcohol\(^8\).

Tracing the total energy of the respective molecule under constrains gives a handle on preferred orientations and positions without the demand of simulating long enough until these situations are reached randomly.

We plan to use this \textit{targeted dynamics} algorithm for the study of concentration depth profiles of different salt solutions and the behaviour of more complex surfactants like phos-
Figure 4. Length distribution of the alkane for different surroundings. The major part of the alkane molecules exhibits an elongated conformation in the interface HX / vacuum. Presence of formamide in the interface HX / FA leads to a higher amount of coiled hexane.

pholipids (lecithin for example).

Acknowledgments

The simulations presented were mainly carried out at the CRAY T3E in the HLRZ Jülich. Part of the work was supported by a grant from the Deutsche Forschungs Gemeinschaft (DFG).

References

Transition Metal Centers in Biological Matrices:
Structure and Function of Vanadate in Vanadium
Haloperoxidases and as Phosphate Analog

Winfried Plass
Fachbereich Chemie und Biologie, Anorganische Chemie II
Universität Siegen, 57068 Siegen, Germany
E-mail: plass@chemie.uni-siegen.de

This research project is intended to elucidate the structure and function of vanadium haloperoxidases as well as the causes and consequences of the chemical analogy between vanadate and phosphate. In this context quantum chemical calculations mostly based on density functional theory are performed on relevant model systems. Along this line an important question is concerned with the role of the protein environment via supramolecular interactions. Although particularly important for the cases considered here, such a supramolecular interplay between metal centers and protein environment is a general feature for metalloenzymes.

1 Introduction

The interest in vanadium chemistry from a biological and pharmacological perspective has exploded over the past 20 years. This is mainly based on the discoveries of the role of vanadium as an insulin mimic, and the presence of vanadium in certain haloperoxidases and nitrogenases. A key point for the understanding of how vanadium actually works in biological systems is given by the chemical analogy between vanadates and phosphates. One of the still open questions in this context is whether vanadate generally acts as a stable transition state analog for phosphate blocking specific receptor sites or whether there are cases for which the vanadate substitution actually leads to new active sites with a completely different reactivity with consequences for the pharmacological application and the toxicity of vanadium.

2 Vanadium in Biological Systems

The biological function of vanadium is well established. Of particular importance is the ability of vanadium to influence phosphate-metabolizing systems and the fact that vanadium is an inherent part of enzymatic active sites. For both types of vanadoenzymes known today, the vanadium-containing haloperoxidases (see figure 1) and the vanadium nitrogenases from the nitrogen fixing bacteria Azotobacter, functional analogs are found in nature which are either more widely spread or more efficient, e.g. the heme-containing haloperoxidases and the conventional nitrogenases with molybdenum cofactor, respectively. This immediately leads to the question of how this enzyme systems did evolve and in particular whether the vanadium-containing enzymes known today are retained functional analogs, which simply could sustain evolutionary forces. New insight concerning these questions may be gained on the basis of newly found similarities for vanadate and phosphate in biological systems.
Figure 1. Structure of the chloroperoxidase of the fungus Curvularia inaequalis determined by X-ray crystallography.

The wide spread physiological effects of vanadium are mainly attributed to the similarity between its anionic form, vanadate(V), and phosphate. But there are also important differences between these two anions. At physiological pH values monovanadate is found as doubly protonated $[\text{VO}_2\text{(OH)}_2]^-$ species, whereas phosphate occurs in the monoprotonated form $\text{HPO}_4^{2-}$. This is also important for possible mechanisms of the transport systems for these two anions. In addition vanadium is easily reduced under physiological conditions to yield cationic species. The third difference is given by the pronounced ability of vanadium to adopt higher coordination numbers. The higher coordinative flexibility of vanadium can deliberately be used for the structural characterization of phosphate metabolizing enzymes.

### 3 Phosphate-Vanadate-Analogy

Recently the crystal structures of several stable enzyme aggregates of phosphatases with vanadate as transition state analog have been reported. An interesting example are the protein tyrosine phosphatases, which are involved in signal transduction mechanisms for controlling and regulating intracellular processes (e.g. the insulin receptor system) — in this context it is worth noting that vanadate complexes show insulin-mimetic effects. In these aggregates the vanadium is trigonal bipyramidal and linked to the protein with a single axial bound cysteine residue, whereas the oxygen atoms of the vanadate moiety are involved in a hydrogen bonding network. A similar structure is found for the active site of rat prostatic acid phosphatase with the complexed transition state analog vanadate (figure 2). In this case the vanadium is linked to the protein through an axial bound histidine residue.
Figure 2. Structure of the active site of the rat prostatic acid phosphatase with complexed vanadate. Hydrogen bonds are shown as broken lines.

Striking similarities are observed for the only structurally characterized vanadoenzymes known today, the vanadium haloperoxidases - e.g. the chloroperoxidase of the fungus *Curvularia inaequalis* (see figure 3)\textsuperscript{11–13}. As in the case of the rat prostatic acid phosphatase the vanadate is directly linked to the protein only through the axial bound histidine residue and is embedded in the protein via an extensive hydrogen bonding network.

Vanadium-containing haloperoxidases (V-HPO) are enzymes catalyzing the two electron oxidation of a halide (X\textsuperscript{−}) to the corresponding hypohalous acid according to Eq. (1)\textsuperscript{14}.

\[
\begin{align*}
H_2O_2 + X^- + H^+ & \rightarrow H_2O + HOX \\
\end{align*}
\]

HOX may further react with a broad range of nucleophilic acceptors to form a diversity of halogenated compounds. These haloperoxidases are named after the most electronegative halide they are able to oxidize, and thus a vanadium chloroperoxidase (V-CPO) is able to oxidize chloride, bromide and iodide.

Based on the recently published crystal structures of the native and apo forms as well as the peroxide- and azide-bound derivatives of this chloroperoxidase\textsuperscript{11–13} several interesting questions still remain and even new once have to be addressed: 1) What is the electronic structure of the active site vanadate moiety, i.e. mono- or dioxo vanadium(V) species (VO\textsuperscript{3+} or VO\textsuperscript{2+})? 2) How is the peroxy group or the chloride ion bound to the active site? 3) What is the influence of the protein environment upon the structure of the active site and the mechanism? 4) Does the apo-protein of this chloroperoxidase exhibit
phosphatase activity as expected based on its structural similarities to phosphatases?

Concerning the last question an interesting result has recently been published by Wever et al., which were able to show that the apo-protein of the chloroperoxidase isolated from the fungus *Curvularia inaequalis* indeed exhibits phosphatase activity\(^\text{15}\). Although it is obvious from the kinetic data that the active site of this V-CPO is not optimized for phosphatase activity, it is nevertheless clear that within the same supramolecular environment it is possible to catalyse two very different reactions such as those of haloperoxidases and phosphatases. Consequently this structural motive should also be observed for the active sites of other V-HPOs and phosphatases. Moreover, this structural similarity should give rise to genetic relationships as well. According to the sequence alignment of relevant enzymes nearly all amino acid residues coordinating vanadate in V-CPO isolated from *Curvularia inaequalis* (Lys353, Arg360, Ser402, Gly403, His404, Arg490 and His496 see figure 3) are conserved in the V-HPOs sequenced thus far as well as within three families of acid phosphatases (see figure 4)\(^\text{15}\). Within this active site architecture the fully conserved histidine residues play an important role. His496 is the residue covalently linking the vanadate to the protein, whereas His404 is essential for catalysis and is proposed to function as an acid-base group. Together with the observed phosphatase activity of the apo-protein of the V-CPO this indicates a very similar architecture of the active sites of these V-HPOs and acid phosphatases.

The dendrogram depicted in figure 4 is based on the sequence alignment of 14 phosphatases and V-HPOs. These results indicate that these enzymes have divergently evolved from a common ancestor. In combination with the observed stability of V-CPO from
Curvularia inaequalis, which in particular resist a high concentration of their substrate \((\text{H}_2\text{O}_2)\) and their product (HOX) that would readily inactivate the heme-containing HPOs, this contradicts the introductory hypothesis that V-HPOs are retained enzymatic systems that simply could sustain evolutionary forces. This is further evidenced by the optimized functionality of V-CPO from Curvularia inaequalis, that is given by its recently proposed putative role as source for the starting material for the production of a chemical weapon, the hydroxyl radical (see Eq. 2), which in turn can be used by the fungus to damage the protective lignocellulose of plant tissues allowing the penetration of the plant cell for nutrient acquisition\(^1\).

\[
\text{HOCl} + (\text{O}_2)^{•−} \rightarrow (\text{OH})^{•} + \text{Cl}^{−} + \text{O}_2
\]  

It is remarkable, however, that for another class of acid phosphatases, that is exemplified by the rat prostatic acid phosphatase (see figure 2), no sequence similarity is found with the enzymes given in the dendrogram of figure 4. Nevertheless the observed structural similarity of the active site, obvious from figures 2 and 3, indicates that a convergent evolution seems to have occurred with respect to these enzyme systems.

Besides the interesting evolutionary aspects of these enzymes, there are also important implications for both research in the acid phosphatase and in the V-HPO field. A first remarkable example for this — in particular from an anthropocentric point of view — is the structure of mammalian glucose-6-phosphatase (G-6-Pase). This enzyme catalyzes the last step in both gluconeogenesis and glycogenolysis and as such it is the key enzyme in glucose homeostasis. G-6-Pase deficiency is the cause of glycogen storage disease.
Figure 5. Membrane topology models for G-6-Pase. Left: Current six transmembrane-helix topology model. Right: Newly proposed nine transmembrane-helix topology model. Putative G-6-Pase active site residues are depicted as closed ovals.

type 1 (von Gierke disease) and is characterized by severe clinical manifestations such as hypoglycemia, kidney enlargement or growth retardation, which usually cause an early death of the patients.

The recently published model for the membrane topology of human G-6-Pase\textsuperscript{17} is not consistent with the results outlined above and had to be modified accordingly as depicted in figure \textsuperscript{5}\textsuperscript{18}. The active site residues of G-6-Pase identified by sequence alignment are all situated on the same side of the membrane and are located within the helices II, III, IV and V. In this new model the histidine residue His\textsuperscript{176} is the nucleophile forming the phosphohistidine enzyme-substrate intermediate. The phosphate moiety is positioned by interaction of the negatively charged oxygens with the positively charged amino acids Lys\textsuperscript{76}, Arg\textsuperscript{83} and Arg\textsuperscript{170}. In analogy to the active site of V-CPO the residues Ser\textsuperscript{117} and Gly\textsuperscript{118} may also donate hydrogen bonds and the histidine group His\textsuperscript{119} may provide the proton needed to liberate the glucose moiety.

As shown by this first example, the common architecture of the active sites of the vanadium-containing peroxidases and the aligned acid phosphatases has important implications, that even reach into seemingly independent fields of research.

4 Mechanism of Vanadium Haloperoxidases

The current understanding of the catalytic activity of V-HPOs based on chemical, biochemical and crystallographic data is summarized in figure 6. It raises new questions. What controls the specific turnover activity, which can vary over 1000-fold depending on the source of the enzyme? What defines the basis of the halide specificity in V-BPO and V-CPO?
To answer these questions a more detailed knowledge on the molecular mechanism of the reactions at the active site is needed. In particular information on possible reaction intermediates is not available at present. Therefore theoretical investigations on suitable model systems including the supramolecular environment of the protein matrix are part of the present project.

This leads to an additional point of interest related to the influence of the protein environment on the particular reactivity of the active site. This is particularly important since the apo-protein of the chloroperoxidase isolated from the fungus Curvularia inaequalis exhibits phosphatase activity. This clearly indicates that within the same supramolecular environment it is possible to catalyse two very different reactions such as those of haloperoxidases and phosphatases. A proposed mechanism for the phosphatase
activity of the V-HPO protein environment is depicted in figure 7.

5 Molecular Model Systems

For a detailed understanding of the reactivity of the biological systems described above a comparison with chemical models is important. This is in particular the case when it comes to the question of the influence of the protein environment on the reactivity. We have synthesized a series of such model systems which exhibit hydrogen bonding at the vanadate center related to the observed supramolecular environment of the native V-HPO proteins (see figures 8).19, 20 The hydrogen bonding network for one of the model systems as observed in the crystal structure is depicted in figure 9.

These model systems in addition also are used to calibrate the methodology of our theoretical investigations. This is important, since the description of hydrogen bonds within theoretical models is a difficult task.
6 Computational Methods

The theoretical investigations related to this project are based on density functional (DFT) calculations. The choice of both the appropriate density functional and the employed basis sets is verified by calculations on model systems ranging from simple vanadate moieties to molecular assemblies like the ones depicted in figure 8. The mechanistic aspects for the protein related systems are addressed by QM/MM methods in which the major part of the protein environment is treated on a molecular mechanics (MM) basis and the key part of the active site by quantum mechanics (QM). For medium sized systems representing the active site structure including the important part of the supramolecular protein environment also molecular dynamics simulations based on the Car-Parrinello method are part of the project. All calculations have been performed at the computer center (ZAM) of the Forschungszentrum Jülich utilizing two massively-parallel computers (CRAY T3E-600 and CRAY T3E-1200) and a vector-parallel computer (CRAY T90). The computer system used for a particular application depends on the efficiency of the implemented software and the size of the molecular assembly to be calculated. For the QM/MM calculations and the molecular dynamics simulations the CRAY massively-parallel computers turned out to be essential tools with an average job size of 128 nodes with about 2 hours of CPU time.

Acknowledgments

This project is supported by a grant of computer time of the John von Neumann Institute for Computing (NIC) and a research grant of the Deutsche Forschungsgemeinschaft.

References

Massively Parallel Individually Selecting Configuration Interaction: A Progress Report

P. Stampfuß and W. Wenzel

Institut für Nanotechnologie, Forschungszentrum Karlsruhe
Postfach 3640, 76021 Karlsruhe, Germany
E-mail: wolfgang.wenzel@int.fzk.de

We report on the progress of our scalable implementation of the configuration-selecting multi-reference configuration interaction method for massively parallel architectures with distributed memory. With this code, calculations with Hilbert spaces containing more than $10^{12}$ configurations are now routinely feasible on up to 128 processors of the IBM-SP2 and up to 256 processors on the CRAY-T3E. We briefly explain the key ingredients of the parallel implementation and report the results of two illustrative applications, regarding the fast quantum dynamics near the conical intersection of NO$_2$ and the investigation of reaction pathways of members of the endyne family, respectively.

1 Introduction

In the development of quantum chemical methods for complex molecules, a consensus has emerged that two important effects must be taken into account in a balanced and accurate fashion in order to arrive at quantitatively correct results. First, dynamical correlations, i.e. the mutual influence two electrons exercise on each other when they pass at close distance, must be accounted for. Secondly, one must be able to accommodate the multi-reference nature of the electronic states in many complex molecules. This effect is particularly important in many transition metal compounds where the close proximity of d-energy levels generates a large number of important multiplets. It is also relevant when one wants to describe an entire potential energy surface, where bond-breaking or bond-rearrangements can occur. An adequate treatment of multi-reference effects is mandatory for the quantitative treatment of electronically excited states. Meeting the needs of both these requirements simultaneously is responsible for the high cost of accurate quantum chemical calculations.

For many years the multi-reference configuration interaction method (MRCI) has been one of the benchmark tools for highly accurate calculations of the electronic structure of atoms and molecules$^{1-3}$. Due to its high computational cost, however, applications of the MRCI method remain constrained to relatively small systems. For this reason the configuration-selective version of the MRCI-method (MRD-CI), introduced by Buenker and Peyerimhoff$^{4-6}$, has arguably become one of its most widely used versions. Even within this approximation, the cost of MRCI calculations remains rather high$^{7-12}$. In order to further extend the applicability of the method, it is thus desirable to employ the most powerful computational architectures available for such calculations. Here we report on the progress of the first massively parallel, residue-driven implementation of the MRD-CI method for distributed memory architectures.

In this manuscript we focus on the details of the implementation of the method and provide timings for benchmark applications that demonstrate the scalability of the method for up to 128 nodes for Hilbert spaces of dimension up to $5 \times 10^9$ of which up to $5 \times 10^6$
elements were selected for the variational wavefunction\textsuperscript{13}. The determinant based code we report here was developed in an object oriented implementation using C++ as the implementation language.

2 Method and Implementation

2.1 Configuration Interaction Methods

The electronic structure of a many-electron system is described by a many-electron wavefunction that can be represented as a weighted average of many-electron configurations. The set of permissible configurations is called the Hilbert space of the system. In each configuration (see Figure (1) (a)), each electron occupies one of many possible orbitals (that are chose from a finite basis). According to the rules of quantum mechanics each orbital can be occupied by at most one electron. In the presence of electrostatic interactions in molecules, the electrons may hop from one orbital pair to another, according to a Hamilton operator

\[ H = \sum_{ij} h_{ij} c_i^\dagger c_j + \sum_{ijkl} h_{ijkl} c_i^\dagger c_j^\dagger c_k c_l, \]

where the symbol $c_i$ represents the operation to remove an electron in orbital $i$ from a given configuration and $c_i^\dagger$ creates an electron in this orbital if none previously existed. $h_{ijkl}$ represents the hopping amplitude for this process (see Figure (1)(b)). The energy of a many
electron wavefunction is computed by applying the above operator to the wavefunction, thus creating another weighted average of configurations (with changed coefficients) and to sum the pointwise products of the coefficients of identical configurations in the original and the transformed wavefunction. The individual configurations are thus interacting via the Hamilton operator, methods that explicitely compute these interactions belong to the family of configuration interaction (CI) methods.

In all CI methods the coefficients of the weighted average that represents the wavefunction are iteratively adjusted to minimize the energy of the resulting state. The variational theorem of quantum mechanics assures us that the many-body wavefunction with the lowest energy is that of the desired ground state of the system. To obtain accurate results, one would like to choose a basis that is as large as possible to describe the many-body wavefunctions. However, if all electrons are permitted to occupy all orbitals — a method known as full CI — the number of configurations grows as \( N!/[n_e!(N - n_e)!] \) with the number of orbitals \( N \) and the number of electrons \( n_e \). For benzene, as a small example \((n_e = 42)\), in a basis of just 100 orbitals, this would result in an astronomical number of \( 10^{28} \) configurations. Fortunately, the many body wavefunction is often dominated by only a few configurations, the so called reference configurations. Applying the Hamilton operator to those, generates a new set of excited configurations that are the most important in the description of the system. In the multi-reference configuration interaction method (MRCI) only the reference and the excited configurations are included in the Hilbert space. Their number grows as \( N_{\text{ref}} \times n_e^2 \times N^2 \) and the number of possible transitions generated by equation (1) grows as \( N_{\text{ref}} \times n_e^2 \times N^4 \), i.e. with the sixth power of the number of electrons (we assume that the number of orbitals is also proportional to the number of electrons). Returning the example of benzene, the calculation can still include \( 10^9 \) configurations, with \( O(10^{11}) \) matrix elements.

Instead of explicitely considering all possible configurations in the MRCI Hilbert space one can reduce the computational effort further by dynamically selecting only those configurations that are important in the actual calculation. In this method, called MRD-CI\(^5\), only the most important configurations of the interacting space of a given set of primary configurations are chosen for the variational wavefunction, while the energy contributions of the remaining configurations are estimated on the basis of second-order Rayleigh-Schrödinger perturbation theory\(^{14,9,5} \). Since the variationally treated subspace of the problem consists of only a fraction of the overall Hilbert space, the determination of eigenstates in the truncated space requires far less computational effort.

Even in this approximation the computational requirements remain large even for relatively small molecules. It is therefore sensible to exploit the most powerful computational resources available to perform these calculations, i.e. massively parallel computers with distributed memory. Unfortunately efforts of chemists and physicists to use these machines for methods of wavefunction-based theoretical chemistry, have so far been met by-and-large with difficulties. The reasons for these difficulties are often the lack of intuitive algorithms which coordinate the calculation on a large number of processors while keeping the communication efforts under control. For MRCI based applications, both the data for the wavefunction and the Hamiltonian coupling constants \( h_{ijkl} \) are so large that they cannot be stored on a single node. Hence both pieces of data must be moved around during the calculation, which must be coordinated on all nodes of the machine in such a way that no idle periods occur on any node of the machine. To generically address the problem for
In order to compute the matrix elements of the Hamilton operator between two configurations we exploit an enumeration scheme in which each matrix element between two configurations is associated with the subset of orbitals that occur in both the target and the source configuration. This unique subset of orbitals is called the transition residue mediating the matrix element and serves as a sorting criterion to facilitate the matrix element evaluation on distributed memory architectures. For a given many-body state, we consider a tree of all possible transition residues as illustrated in Figure (2). For each such residue we build a list of residue-entries, composed of the orbital-pairs (or orbital for a single-particle residue) which combine with the residue to yield a selected configuration and a pointer to that configuration. While the number of transition residues is comparatively small, the overall number of residue-entries grows rapidly (as $N_{selected} n_e^2$) with the number of configurations $N_{selected}$ and the number of electrons $n_e$. Once the residue

![Figure 2. Schematic representation of the two-particle residue-tree.](image)

For each element of the configuration-list (A) all possible two-particle residues are constructed. In the configuration illustrated in (B) each box represents one occupied orbital, the shaded region corresponds to the residue and the two white boxes to the orbital pair. The $(n_e - 2)$-electron residue configuration is looked up in the residue-tree (C), where an element (D) is added that encodes the orbitals that where removed, information regarding the permutation required and the index of the original configuration in the configuration list. Solid arrows in the figure indicate logical relationships, dotted arrows indicate pointers incorporated in the data structure. The residue-list, along with all elements must be rebuilt once after each expansion loop, the effort to do so is proportional to product of $n_e$ with the number of configurations. The number of matrix elements encoded in a single element of the residue-tree is proportional to the square of the number of entries of type (D).

the unstructured wavefunction in MRD-CI, we have developed a transition residue driven matrix element evaluation scheme that accomplishes this task in an explicitly scalable fashion, independent of the number of nodes, the number of electrons and the number of orbitals. In the following we will briefly outline the key ideas of this approach, report the scaling behavior in benchmark calculation and summarize two illustrative application of the program.

### 2.2 Parallel Implementation

In order to compute the matrix elements of the Hamilton operator between two configurations we exploit an enumeration scheme in which each matrix element between two configurations is associated with the subset of orbitals that occur in both the target and the source configuration. This unique subset of orbitals is called the transition residue mediating the matrix element and serves as a sorting criterion to facilitate the matrix element evaluation on distributed memory architectures. For a given many-body state, we consider a tree of all possible transition residues as illustrated in Figure (2). For each such residue we build a list of residue-entries, composed of the orbital-pairs (or orbital for a single-particle residue) which combine with the residue to yield a selected configuration and a pointer to that configuration. While the number of transition residues is comparatively small, the overall number of residue-entries grows rapidly (as $N_{selected} n_e^2$) with the number of configurations $N_{selected}$ and the number of electrons $n_e$. Once the residue
Figure 3. Schematic representation of the computation of two-particle matrix-elements in the expansion step using the residue-tree. For a given configuration (A) we form all two-particle residues, which are looked up in the residue tree. In the configuration illustrated in (A) each box represents one occupied orbital, the shaded region corresponds to the transition residue and the two white boxes to the orbital pair. The \(v_e\) electron residue configuration is looked up in the residue-tree (B). Each orbital pair (C) associated with the residue encodes a matrix element with an element of the configuration list (D). The orbital indices of the required integral are encoded in the orbital pairs in (C), the coefficient of the source configuration is looked up directly in (D). Only one lookup operation is required to compute all matrix elements associated with the given transition residue and only the subset of matrix elements that lead to selected source-configurations are constructed.

In order to demonstrate the scalability of the implementation we have conducted benchmark calculations for typical applications of the program, concerning the evaluation of the importance of the triple and quadrupole excitations for the potential energy surfaces of the oxygen molecule \((10^9 \text{ configurations, } 1.8 \times 10^6 \text{ selected})\) and the excited states for benzene \((1.3 \times 10^9 \text{ configurations, } 1.6 \times 10^6 \text{ selected})\), respectively (for details see\(^\text{13}\)).

Figure (4) shows the total computational effort (excluding the time to read the integral file) of the aforementioned scaling runs as a function of the number of nodes. In these plots, the computational effort for all logic-steps are subsumed in one category, the expansion loop and the iteration loop constitute the other main components of the program. For benzene we find almost perfect scaling from 48 to 128 nodes, the overall speedup factor from 64 to 128 nodes is 1.86. For the benchmark calculation of \(\text{O}_2\) a more pronounced increase in the overall computational effort is observed in going from 32 to 128 nodes, in particular in with the last doubling from 64 to 128 processors.

3 Applications

The algorithm described above has been used for a number of applications, e.g. the photochemistry of \(\text{NO}_2\), the computation of electron affinities of oxides, the spectroscopy of transition-metal dihalides, the elucidation of the electronic structure of benzofuroxane and
Figure 4. Total CPU time in (sec) for the fully converged calculation of the ground state of the two benchmark calculations described in the text as a function of the number of nodes. A straight line indicates perfect scaling of the computational effort with the number of nodes. The shaded areas in the bars, from top to bottom, indicate the contributions of the matrix element evaluation, logic and the selection loop, respectively.

recently the investigation of single transition metal centers in biomolecules, e.g. azurin. Here we illustrate its usefulness with two examples, regarding the conical intersection of NO$_2$ and the reaction mechanisms of a family of potential anti-cancer drugs.

3.1 Conical Intersection in NO$_2$

Vibronic interactions between different electronic potential energy surfaces are a generic feature of polyatomic molecules. Such interactions are strong for near-degenerate or degenerate electronic states, and the spectroscopy and dynamics of a molecule are strongly
Figure 5. Two dimensional cut through the three-dimensional adiabatic potential energy surface of the \( \tilde{X}^2A_1 \) and \( \tilde{B}^2B_1 \) electronic states of NO\(_2\) in the vicinity of the conical intersection computed using the quasi-diabatic representation of the underlying potential energy surfaces discussed in the text. The mean distance was \( r = 1.25 \) Å. The two horizontal axes measure the angle and the asymmetric stretch respectively. The vertical axes shows the energy in millihartree with respect to an offset of \( E_0 = -204.0 \) a.u. influenced by the associated non-adiabatic couplings\(^{15-18}\). A typical scenario in polyatomic systems is the occurrence of conical intersections of potential energy surfaces, where two electronic states become degenerate along a hypersurface of the PES. In dealing with such a system theoretically one needs to go beyond the Born-Oppenheimer approximation and solve complex coupled differential equations in order to monitor the nuclear motion simultaneously on more than one electronic state. Although this concept is well known in the literature, the actual construction of diabatic electronic states for a polyatomic system is still a highly difficult task and has only partly been achieved to date\(^{19}\).

NO\(_2\) is a well known triatomic molecule which has an outstandingly complex spectroscopy. Most of the investigations on this system to date are concerned with understanding the highly dense spectral lines in its optical spectra\(^{20-25}\). Adiabatic three-dimensional potential energy surfaces were computed using the configuration-selecting multi-reference configuration-interaction method\(^5\) with the cc-pVTZ basis set\(^{26}\). The calculations were carried out in \( C_s \) symmetry using state-averaged approximate natural orbitals generated with MR-BWPT\(^{27,28}\). A two-dimensional section through the computed adiabatic PES in this representation is shown in Figure (5), its features were found to be in excellent agreement with the available experimental data.

Employing these coupled \( \tilde{X}^2A_1/\tilde{A}^2B_2 \) diabatic potential energy surfaces Mahaptra and Köppel calculated the photodetachment spectrum of NO\(_2^-\) for a transition to the \( \tilde{A}^2B_2 \) electronic state of NO\(_2\) with the aid of a time-dependent wave packet propagation scheme.
The details of their investigation are beyond the scope of this report (see\textsuperscript{29}) so that only the most important points are summarized here. The $\tilde{A}^2B_2$ photoelectron spectrum originates right at the onset of the nonadiabatic regime (the $\tilde{X}^2A_1/\tilde{A}^2B_2$ cusp occurs $\sim 0.08$ eV above to the minimum of the $\tilde{A}^2B_2$ electronic state) and it is highly sensitive to the strength of the nonadiabatic interaction. The photoelectron spectrum obtained for the uncoupled $\tilde{A}^2B_2$ electronic state reveals progression along the bending vibrational mode of NO$_2$; the peaks are $\sim 0.1$ eV apart, which is about a quantum of bending vibration of NO$_2$ in the $\tilde{A}^2B_2$ state. The $\tilde{A}^2B_2$ photoelectron spectrum obtained including the coupling to the $\tilde{X}^2A_1$ state reveals main progression of lines along the bending vibrational mode of NO$_2$ and in addition clustering around each main line is an effect of nonadiabatic interactions on this photoelectron band. Due to these interactions the high energy vibrational states of $A_1$ symmetry mix with the low-lying vibrational states of $B_2$ symmetry and the resulting spectral lines are clumped into groups of non-overlapping resonances. To compare this coupled state spectrum with the experimental recording of Weaver \textit{et al.}\textsuperscript{23} we convoluted it by a Lorentzian function with FWHM of 28 meV. The resulting convoluted spectrum compares well with the experimental one.

### 3.2 Enediyne

The enediyne family of molecules has been investigated for some time because these molecules can undergo a cyclization reaction to a para-dihydrobenzene derivative that is capable of lysing cellular DNA and hence cause cell death\textsuperscript{30}. This property of the molecules,
if selectively activated in diseased or cancerous cells offers the possibility to use members of this family as cell-specific drugs against such diseases. The ultimate goals is to design a compound that in cancerous cells will spontaneously undergo cyclization at physiological conditions and thus kill diseased cells. In order to predict the effectiveness of specific compounds it is important to understand the electronic structure of educt and product of the cyclization reaction as well as the height of its barrier. Possible applications of derived compounds aside the quantum chemical description of the cyclization reaction proves difficult and interesting because of the strong change in the electronic structure of the molecule during the reaction.

We have therefore undertaken a set of MRD-CI benchmark calculation into the Bergman cyclization of the simplest member of the enediyne family. We have performed a systematic investigation of these compounds using cc-pVDZ, cc-pVTZ and cc-pVQZ basis sets (the largest ever considered for these molecules). As illustrated in Figure (6) we have been successful in establishing that MRD-CI reproduces the CCSD(T) results. The transition state (using the CCSD geometry) is predicted 2 kcal/mol less than by CCSD(T), the product lies 7.2 / 5.8 kcal/mol above the educt depending on whether a multi-reference Davidson correction is applied in the MRD-CI calculation or not. This brackets the values of the CCSD(T) calculation.

4 Summary

In recent years we have developed an explicitly scalable implementation of an established quantum chemical method that permits the investigation of complex chemical processes with high accuracy. Presently this program permits the routine treatment of molecules containing about 100 correlated electrons at benchmark accuracy and thus allow the application of this technique to new areas of scientific investigation, such as the analysis of enzymatic reaction that are catalyzed by transition metal centers. The first set of calculations, aiming to elucidate the enzymatic reaction mechanism of isopenicillin-N-synthase with the help of a new inhibitor-protein complex are presently under way in our group.

Acknowledgments: Part of this work was supported by DFG Grants KEI-164/11-2, WE 1863/10-1 and WE 1863/11-3. We acknowledge the use of computational resources at the NIC and the RZ Karlsruhe.

References

Modeling Si/B/N/(C) Ceramic Materials

Marcus Gastreich¹, Silke Reinhardt¹, Markus Doerr¹, and Christel M. Marian¹,²

¹ Fraunhofer Institut Algorithmen und Wissenschaftliches Rechnen (SCAI)
    Schloss Birlinghoven, 53754 St. Augustin, Germany
    E-mail: {marcus.gastreich, silke.reinhardt, markus.doerr, christel.marian}@scai.fhg.de

² Institut für Theoretische Chemie, Heinrich-Heine-Universität
    Universitätsstr. 1, 40225 Düsseldorf, Germany
    E-mail: marian@theochem.uni-duesseldorf.de

1 Introduction

Boron nitride and silicon nitride exhibit a diversity of industrial applications.¹–³ A combination of both, i.e., a material in the system Si/B/N promised to have a higher temperature resistance and to generally exhibit a lower tendency to crystallize. In fact, an amorphous solid with the desired properties was synthesized from molecular precursors in the 1990ies by Martin Jansen and coworkers.⁴ Since that time, a variety of precursor derived Si/B/N/C ceramics have been developed.⁵–¹⁰ Some of the carbon containing polyborosilazanes can even be spun and drawn to fibers. These ceramics are supposed to have great potential as working materials.¹¹–¹³

For amorphous compounds as the above ones, structural data is extremely hard to extract solely from experiment. To aid the structural analysis, we have pursued three complementary approaches:

1. Quantum chemical investigation of the mechanisms and energetics of initial polymerization reactions
2. Computer simulation of the amorphous solid state and comparison of bulk properties
3. Ab initio calculation of local properties and set-up of quantitative structure-property relations (QSPR)

In the following, we shall give a brief survey over our contributions to the structural modeling of amorphous Si/B/N/(C) materials and of their genesis.

2 Reactions of Molecular Precursors

[(Trichlorosilyl)amino-dichloro]borane (TADB) and [(trichlorosilyl)dichloroboryl]ethane (TSDE) are two prominent molecular precursors for SiBNC ceramics (see Fig. 1). Both carry a silicon trichloride and a boron dichloride functional group. Typically, these molecular precursors are concatenated by liquid amines under cleavage of B-Cl or Si-Cl bonds (aminolysis) to yield larger molecular and polymeric units. Cross-linkage of the latter to form a three-dimensional SiBN(C) network is believed to occur at high temperatures, when the polymers are pyrolyzed (“burnt”). TADB and TSDE are multifunctional molecules.
Thus, besides N-attack at Si or B, many different reactions may occur leading to differently linked polymers and eventually to different microscopic distributions of elements in the solid.

What are the mechanisms and heat of reactions for an aminolysis of an Si-Cl and a B–Cl bond? Can the central nitrogen in the Si–N–B unit act as a Lewis base and cause a cross-link or branching of the molecular backbone? What is the probability of further aminolysis reactions with the solvent compared to a linkage of the monomeric units? Does the latter occur statistically via boron and silicon or is there a preference for one or the other side? Can the central Si–N–B and Si–C–B units break up upon aminolysis of the Si–Cl and B–Cl bonds? How do reaction conditions (solvent, temperature) influence the polymerization? In the following, we will try to give an answer to a few questions arising in this context.

2.1 BCl$_3$ and NH$_3$

To get a first idea of the reaction mechanism during the polymerization process, we began with the less complicated case of the ammonolysis of BCl$_3$ studied by means of quantum chemical methods.\textsuperscript{14, 15}

In the gas phase, i.e., for BCl$_3$ reacting with a single NH$_3$ molecule, the reaction coordinate is intuitively clear. Furthermore, entropic contributions are supposed to play a minor role. Therefore, time-independent methods can be employed to determine the course of the reaction. Minima and saddle points were located on the potential energy surface by roughly scanning along the constrained minimum energy path followed by BFGS searches.\textsuperscript{14} Energies were computed at the level of second-order Møller-Plesset perturbation theory (MP2) using the Turbomole program package.\textsuperscript{16} (Density functional theory (DFT) yields similar total reaction energies and barrier heights as MP2, but severely underestimates adduct formation energies.) In the course of the reaction, firstly a very stable adduct H$_3$N·BCl$_3$ with four-fold coordinated N and B atoms is formed. The subsequent elimination of HCl is endothermic with respect to the adduct and proceeds via a high-lying transition state and a hydrogen-bonded complex. Overall, the first and second ammonolyses of BCl$_3$ are exothermic whereas the substitution of the last chlorine atom for an amino group is slightly endothermic in the gas phase (not shown).
When $\text{BCl}_3$ is dropped into liquid ammonia, more than one $\text{NH}_3$ molecule will participate in the reactions. Furthermore, ammonia is a polar solvent and thus preferentially stabilizes ionic or highly polar states. We have studied the mechanism and the energetic course of the substitution of a first chlorine atom of BCl$_3$ in the presence of eleven NH$_3$ molecules under periodic boundary conditions (PBC). In this case, the reaction path and mechanism cannot be foreseen. To obtain an unbiased picture, we allowed the system to evolve in time without constraints by means of Car-Parrinello molecular dynamics (CPMD). All CPMD simulations were performed using the massively parallelized Car-Parrinello program, version 3.0h, as installed on the Cray T3E at the NIC. The solvent is found to have a strong influence on both, the energetic course and the mechanism of the reaction (see Fig. 2). In the presence of ten additional ammonia molecules, the first ammonolysis releases 219 kJ·mol$^{-1}$ compared to 50 kJ·mol$^{-1}$ in the gas phase. Furthermore, every single step is exothermic: Formation of a boron trichloride ammonia adduct [Cl$_3$B·NH$_3$], proton transfer yielding a salt-like intermediate $[\text{NH}_4^+][\text{Cl}_3\text{BH}_2^-]$, and elimination of a chloride ion leading to the products BCl$_2$NH$_2$ and [NH$_4^+$]Cl$^-$. All the intermediates and products are stabilized through hydrogen bonds with solvent ammonia molecules. Contrary to the gas phase reaction, the adduct formation proceeds via a transition state. The existence of such a (small) barrier is easily explained: The nitrogen lone-pair of the adduct-forming NH$_3$ molecule has to be resolved from the H-bonded network before it can attack the Lewis acid BCl$_3$. Once this barrier has been overcome, the reaction completes within picoseconds.
2.2 TADB and NH₃

As a next case, we studied reactions between the molecular precursor TADB and NH₃. Let us firstly concentrate on the gas phase results. As expected, the main primary reaction is a substitution of a chlorine atom for an amino group at the boron center, the major product being (trichlorosilyl)(aminochloroboryl)amine (TACBA). If equal amounts of TADB and NH₃ are present, the further course of reaction is dominated by the oligomerization of TACBA to form very stable trimers with fourfold coordinated B and N. Elongation and branching of the Si–N–B backbone occur preferably after a single ammonolysis of a B–Cl or Si–Cl bond of the molecular precursor because the Lewis acidity of boron and silicon diminishes with an increasing number of amino substituents. Probable side reactions are processes during which the Si–N–B unit of the precursor opens up. Among the latter is the dissociation of the H₃N·BCl₂NHSiCl₃ adduct to yield BCl₂NH₂ and SiCl₃NH₂ and the attack of the central amino group of TACBA at the boron center of another TADB in which bis[(trichlorosilyl)amino]chloroborane (TACB) and BCl₃ are formed. Both reactions lead to a segregation of boron-rich and silicon-rich regions. It is for these reasons that TADB is not suited as a molecular precursor for the chemical vapor deposition of homogeneous borosilazane ceramic coatings.

The transition states of all the above mentioned substitution reactions are more polar than the reactants and will thus be stabilized if the polymerization is conducted in liquid ammonia. HCl eliminations are particularly favored by additional ammonia molecules: not only are their reaction barriers lowered markedly, their heats of reaction increase significantly due to the formation of NH₄Cl. For this reason, they are considerably enhanced over dissociation reactions and oligomerizations. In solution, the substitution of a chlorine atom for an amino group at the boron side of TADB remains the major primary process. The side reaction (i.e., the H₃N·BCl₂NHSiCl₃ dissociation into BCl₂NH₂ and SiCl₃NH₂) is suppressed by an excess of ammonia. An important process in the continuing stage is a second ammonolysis at the boron side of TACBA. Finally, there is an indication for a preferred linkage of the monomeric TACBAs via their boron sides. Apart from the sequence just depicted, borazine rings are found to be particularly stable and should therefore be a frequent motif in the covalent network. Si–N–Si links are assumed to be formed at a later stage when the polymer is pyrolized. The preferred elongation of the polymer backbone via boron and the formation of borazine derivates probably leads to a non-uniform microscopic distribution of boron and silicon in the final ceramic. Indeed, recent ¹¹B-{²⁹Si}-REDOR NMR experiments support these assumptions.

2.3 TSDE and Methylamine

Very recently, we started to look at the initial reactions of the molecular precursors TSDE (see Fig. 1) and methylamine which lead to an SiBNC ceramic with very high temperature resistance (2100°C). We present here only preliminary results of these quantum chemical investigations. The mechanisms of the aminolysis reactions between NH₂CH₃ and TSDE in the gas phase closely resemble those of the TADB ammonolysis. However, comparing the energetic courses, marked differences are observed: The formation of adducts and products during the TSDE aminolysis yields much more energy and the reaction barriers are considerably lower. These trends can be explained by the fact that TSDE contains a saturated ethyl central unit whereas the Lewis acidity of boron and silicon in TADB is
reduced due the partial double bond character of the Si–N–B unit. The transition state for the HCl elimination at the boron side of TSDE is located even below the reactants. A possible side reaction is again the destruction of the original monomer unit.

![Diagram of reaction energies and structures](image)

**Figure 3.** Simulation of the aminolysis of [(trichlorosilyl)dichloroboryl]ethane by liquid NH$_2$CH$_3$.

Solvent effects have been taken into account by treating the reaction of TSDE with 15 NH$_2$CH$_3$ in a periodic box. So far, however, limited simulation times have prevented the observation of a spontaneous reaction during the CPMD modeling. Instead, we have calculated a constrained minimum energy path for the aminolysis of TSDE, assuming a reaction mechanism similar to the one observed in the model system BCl$_3$ + 10 NH$_3$ (see Section 2.1). In an excess of NH$_2$CH$_3$, the overall reaction releases 176 kJ·mol$^{-1}$ and is thus considerably more exothermic than in the gas phase (51 kJ·mol$^{-1}$ without zero-point vibrational energy corrections). Furthermore, reaction barriers are of the order of 20 kJ·mol$^{-1}$ and should easily be overcome. The contrary is true for the B–C-cleavage of the adduct into BCl$_2$NHCH$_3$ and SiCl$_3$CH$_2$CH$_3$. This reaction is much less exothermic than in the gas phase (22 kJ·mol$^{-1}$ vs. 116 kJ·mol$^{-1}$). Also the transition state is destabilized as it requires the migration of a proton or hydrogen atom from the amino nitrogen to the ethyl carbon. It can therefore be safely assumed that this side reaction is suppressed in solution. Presently, we are investigating the most important entropic effects on the aminolysis reaction by thermodynamic integration along the constrained minimum energy path (blue moon ensemble$^{23}$).
3 Parametrization and Validation of Classical Force Fields

Structural models for solids are either isolated clusters or periodically repeated units. In the case of amorphous materials, one has to compensate for the faulty assumption of periodicity by large unit cells. Otherwise, the amorphous character would not be grasped at all. On the other hand, for cluster models, long-range order is not an issue, but small aggregates easily suffer from surface effects. Thus, unit cells and clusters favorably comprise thousands of atoms. Despite the progress of ab initio molecular dynamics methods in this field,24 ab initio studies of Si$_3$B$_3$N$_7$ employing unit cells of this size are presently out of reach. Also, systematic global optimization procedures are not feasible which is why the phase space is sampled by classical molecular dynamics (MD) or Monte Carlo (MC) simulations. The latter require a predetermined expression for the potential energy, that is: a cost function. When the SFB 408 was founded, analytical interaction potentials for ternary Si/B/N compounds were not available in the literature. One of the initial stages of such computer modelling was thus the parameterization of an appropriate classical force field.

![Figure 4. Periodic (left) and cluster (right) model of amorphous Si$_3$B$_3$N$_7$ (red: silicon, green: boron, blue: nitrogen). Coordination polyhedra are indicated for silicon and boron only.](image)

Boron, nitrogen, and silicon are typical network building elements; bonding in solid boron nitride, silicon nitride, and borosilazane compounds exhibits partially ionic and partially covalent character. Ionic bonding is dominated by long-range Coulomb attraction and short-range Pauli repulsion. Both are isotropic in space, i.e., they do not depend on the position coordinates of two particles but only on their distance. Covalent bonding, on the other hand, involves directional forces. In a classical force field, these are typically represented by three-body interaction terms. The presence of three-body potentials in general improves force constants and other second-order properties of covalent compounds (vibrational properties, phonons, bulk moduli etc.). A complication arises when the coordination number varies for a pair of elements during simulations because three-body terms
are biased towards a certain bond angle and concomitantly towards a particular atomic coordination. It is thus extremely difficult to parameterize an interaction potential including three-body terms which simultaneously reproduces structures and second-order properties of differently coordinated modifications of a substance.

For solid state simulation, there occurs another principal difference to classical (harmonic) force fields: In the molecular modeling of organic and biochemical compounds, typically different parameter sets are utilized for different coordinations. This procedure implies that covalent bonds are neither broken nor built in the course of the simulation. The situation is different for ceramics: these materials are synthesized at high temperatures, and bond breaking and bond formation therefore have to be part of the modeling. Further, in amorphous networks such as borosilazanes, irregularities may occur: NMR spectra of amorphous Si$_3$B$_3$N$_7$ indicated that the prevailing boron and nitrogen coordination is trigonal and that silicon is mostly tetrahedrally surrounded in this material, but under- and over-coordinations should not be ruled out in general.

Three different potentials have been parametrized for the modeling of Si/B/N/(H) compounds, labeled by $Q$, $TB$, and $WAT + H$. The potential $Q$ is based on a charged model for Si/B/N compounds. The assignment of effective charges is necessary for the determination of dipoles and associated properties, such as infra-red intensities and possible splittings between longitudinal and transversal optical phonon modes (LO/TO splittings). On the other hand, long-ranged Coulomb interactions require infinite summation techniques which add considerable complexity to a computer simulation and slow it down. $TB$ and $WAT + H$ refrain from introducing explicit charges. Instead, Coulomb repulsions between Si–Si, B–B, N–N, and B–Si atomic pairs are mimicked by exponentially damped $1/r$-type potentials. By construction, these are short-ranged and avoid Ewald summation procedures. The potential energy functions acting between Si–N and B–N are expressed as Morse potentials. In order to obtain the proper layer structures of hexagonal and rhombohedral BN we included nitrogen-nitrogen dispersion interactions. The potential $WAT + H$ is particular in the sense that it additionally contains energy functions describing N–H, B–H, and Si–H bonds. For the parametrization of these energy functions, geometrical structures and harmonic vibrational frequencies of molecular amines, boranes, silanes, borazines, etc. were computed by means of density functional theory.

The parameters of the energy expressions are fitted to reproduce properties of molecules and crystals in a training set which ideally is sufficiently broad to yield transferable interaction potentials. All fitting has been performed using the GULP program. We fit at constant-pressure and zero-temperature conditions using a weighted least squares procedure. Observables are either physical properties such as binding energies, vibrational frequencies etc. or structural data. Wherever applicable — i.e., in regions in which a structural relaxation results in a positive definite Hessian — the relax fitting mechanism has been employed; in this fitting mode crystalline or molecular structures are optimized at every step, and coordinates are taken as observables instead of forces. Detailed information on molecular and solid state reference data and on our selection concerning the separation of the data set utilized for the training or the test of the force-field parameters may be found in the respective publications.

To get an impression of the quality of the potentials, let us give some examples for validation studies. All potentials reproduce the structures with very low root mean square deviation (RMSD) values; for example, the $Q$-model considering 279 observables leads
Figure 5. Parametrization and validation of a classical potential for the modeling of amorphous solids (left) and its application in continuous random network (CRN), molecular dynamics (MD), Monte Carlo (MC), and reverse Monte Carlo (RMC) procedures (right).

to an RMSD of 0.011879 Å. If one loosens the constraint of conserving symmetry of the (crystalline) test set, the value amounts to 0.01269 Å — whilst having to take into account 663 observables. The other potentials behave similarly, with limited loss of accuracy (not shown) due to the neglect of explicit charges. Besides structures, second-order properties usually are good targets for validation purposes. Such properties are important because they are related to the acceleration of particles in simulations. Thus, especially vibrational data and bulk moduli represent often-seen test cases. Being the potential with the highest evaluation time, the charged particle potential $Q$ exhibits the most-transferable description of second-order properties.

An energy function is tested in a tougher way when applied to the description of dynamic properties. The standard validation scenario is the comparison of radial distribution functions from experiment with a theoretical one obtained after an MD simulation or other optimization techniques. We paradigmatically show this in Figure 6 for a 979-atomic cluster model of Si$_3$B$_3$N$_7$ giving reference to some of the available experimental sources in the SFB 408. It has to be taken into account that different elements scatter differently and as a function of the radiation used; this is why the intensity for some of the peaks differs between the transmission electron microscopy (TEM) results and the ones from X-ray experiments. In our case, we treated the theoretical data to be comparable to X-ray experiments. It can be seen that the distance region to around 3.5 Å is well reproduced by all potentials. The early signals are sharp and at distances where one expects them (typical B–N bond around 1.44 Å, typical Si–N bond around 1.72 Å). There are no artifacts demonstrating the robustness of the potentials employed. Higher distances are less accurately calculated which is mainly due to the fact that the model is relatively small (radius $\approx 10$ Å) — here, we do not aim at establishing a definite structural model for Si$_3$B$_3$N$_7$. It is beyond the scope of this article to also document the overlapping signals in inverse space which is often done to document the validity of the cost function in terms of long-
range forces. All in all, today, we have three (four if hydrogen is taken into account) cost functions at hand; the application purposes determine the amount of tradeoff which has to be made on the side of either time or accuracy.

4 Parametrization of NMR Chemical Shifts and Reverse Monte Carlo Refinement

Another valuable for the exploration of the structure of amorphous solids are NMR experiments. We could demonstrate that especially the N chemical shift is very sensitive to changes in the local environment around this atom.\textsuperscript{31} NMR chemical shifts can nowadays be calculated with sufficient accuracy at the \textit{ab initio} level of theory. These calculations are limited however to systems with less than about 100 atoms. Even compounds of that size required the facilities at NIC. In addition, such methods are not directly applicable to structural models for amorphous solids (which usually contain several thousands of atoms). To be of use, we have parametrized the chemical shift $\delta$ as a function of the local structure. We trained the fit with \textit{ab initio} calculations on smaller clusters containing 40-70 atoms. We started out with the binary system BN; all clusters were hydrogen-saturated cut-outs of the crystals hex- and cubic BN. Clusters were enlarged to convergence of $\delta$ for central N atoms so that we were able to predict the formerly unknown $\delta({}^{15}\text{N})$ for hexagonal boron nitride. The calculated chemical shifts agree within a few ppm with an experiment that was conducted half a year later by Jeschke \textit{et al} in the framework of the SFB 408.\textsuperscript{32} Calculations on Si$_3$N$_4$, the other binary subsystem of Si$_3$B$_3$N$_7$, show that the N chemical shift in
this system should be parametrizable in a similar way.

A widely used approach for the refinement of structural models for amorphous solids is the Reverse Monte Carlo method. Basic idea of this method is to modify a structural model randomly until the properties calculated from this model agree as good as possible with experimental data. Experimental data is usually taken from diffraction experiments (X-ray, neutron, or electron). As such data represent 1D-projections of a 3D structure, there emerge ambiguities w.r.t. the generated structure model. To overcome this shortcoming, it is desirable to take as much information into account as possible. A standard procedure is to limit the solution space to energetically accessible structures — evaluated by the $T_B$-potential in our case. An obvious improvement then is to further augment the cost function by NMR information. As a first test case, we are currently applying this scheme to amorphous BN.

**Acknowledgments**

We would like to dedicate this article to the memory of Attila Çağlar, a colleague from the SFB 408, who died much too early at the age of 33. We have very much appreciated him as one of the first researchers to employ our potentials in his work on BN-nanotubes and are grateful having experienced his open mindedness. He will be deeply missed.

This work has been funded by the Deutsche Forschungsgemeinschaft in the framework of Sonderforschungsbereich 408 “Anorganische Festkörper ohne Translationssymmerie”.

Figure 7. Generation of a structural model in a Reverse Monte Carlo procedure.
We thank the John von Neumann Institute for Computing in Jülich for the allocation of computing time and technical support — especially Drs. Johannes Grotendorst and Norbert Attig.

References

Elementary Particle Physics

G. Münster

Institut für Theoretische Physik, Universität Münster
Wilhelm-Klemm-Str. 9, 48149 Münster, Germany

The two classical branches of elementary particle physics are the experimental and the theoretical one. Experimental particle physics started more than a hundred years ago with the discovery of the electron. The constituents of the atomic nuclei, protons and neutrons followed, but more and more new and exotic elementary particles were discovered in the course of time. The analysis of their behaviour revealed three different kinds of fundamental forces between the particles: first the electromagnetic interactions, which include the electric and magnetic phenomena, and are most important in everyday life, secondly the weak interactions, which are responsible for certain types of radioactivity and nuclear reactions, and thirdly the strong nuclear interactions, which bind the constituents of atomic nuclei together. In addition, there is the familiar gravitational force, which, however, does not play an important role in the physics of elementary particles.

The first complete theory of elementary particles was Quantum Electrodynamics, QED, the theory of electrons of photons. It describes the electromagnetic interactions of charged particles in a way consistent with both quantum theory and the special theory of relativity. It originates in the early work of Jordan, Pauli, Heisenberg, Dirac and others and was brought to maturity in the 1940s by Feynman, Schwinger, Dyson and Tomonaga. In QED the electromagnetic interactions are mediated by photons, the particles of light. The duality of matter and waves finds its full expression in QED: the quantized electromagnetic field can equivalently be considered as an ensemble of photons.

QED was not only a complete but also a perfect theory. It led to predictions, whose experimental examinations belong to the tests of physics with the highest precision. The necessary calculations do not rely on electronic computers, they are based on series expansions. The physical quantities of interest are expanded in powers of the electromagnetic coupling, the finestructure constant \( \alpha = \frac{e^2}{4\pi\epsilon_0c} \), whose numerical value is approximately \( 1/137.036 \). It is the smallness of this number which allows such precise predictions to be made in QED.

The situation is different for the case of the strong interactions, where the coupling strength is by no means small. Methods analogous to the series expansions of QED soon came to an end, and much effort was spent to find suitable calculation schemes in the theory of strong interactions. In the early 1970s an ingenious proposal for a theory of the strong interactions was made, namely Quantum Chromodynamics, QCD. It describes the constituents of all nuclear matter, the quarks, and their mutual interactions via the exchange of so-called gluons. However, it was difficult to arrive at concrete numerical predictions from it.

A breakthrough was initiated by Ken Wilson’s formulation of QCD on a space-time lattice. The replacement of the space-time continuum by a four-dimensional lattice allowed the application of a great variety of new methods, some of them coming from statistical mechanics of lattice systems. One of these methods was introduced to the field of elementary
particle physics by Creutz, Jacobs, Rebbi and by Wilson himself, namely the numerical simulation of theories on a lattice by means of Monte Carlo algorithms. It turned out to become one of the most powerful tools for obtaining predictions from QCD and other models of elementary particle physics.

The simulation of a model is different from the traditional theoretical approaches. It is more of a numerical experiment, performed on a computer, with unknown outcome. The result is afflicted with systematic and statistical errors, like in real experiments. Because of its intermediate character, we can call it a computer experiment.

With increasing computer power, the results from numerical simulations of QCD and other models have become better and better. The major field of application within elementary particle physics is still QCD. The articles of Lippert, Orth and Schilling from Wuppertal, and of Sommer and Wittig for the ALPHA collaboration give an introduction to the basic ingredients of QCD and its Monte Carlo simulation. They discuss how problems like the calculation of the spectrum of bound states and the determination of fundamental parameters of QCD can be solved by performing computer experiments. Other quantities of great interest for the phenomenology of particles include the structure functions and form factors of nucleons, which are discussed in the article of Schierholz for the QCDSF collaboration.

In the early history of the universe and in different astrophysical objects like neutron stars, matter is under extreme conditions, i.e. at very high temperatures or densities. For the theoretical study of nuclear matter in such situations Monte Carlo simulations are an indispensable tool. Research in this field is described in the articles by Karsch et al. from Bielefeld, and by Montvay, Hands et al. from DESY, Hamburg, and Swansea.

Although numerical simulations of models of particle physics have been performed for more than 20 years now, there is still progress in the development of algorithms and lattice formulations of theories. The article of Jansen et al. discusses obstacles and their removal for the realization of an important symmetry, chiral symmetry of quarks, on the lattice.

QCD is part of the so-called Standard Model of elementary particle physics. Despite the great success of this model, it has some shortcomings, and physicists try to go beyond the Standard Model. The most popular attempts in this direction involve supersymmetry, an extended concept of symmetry. The study of supersymmetric theories with the help of numerical simulations is the subject of the article of Montvay, Münster et al. from the DESY-Münster-Roma collaboration.

Finally, the article of Jansen shows, how sources of systematic errors in computer experiments, namely the finite lattice spacing and the finite volume of the lattice, can be dealt with and can even be exploited to gain physical information.
Quarks in a Box

Thomas Lippert, Boris Orth, and Klaus Schilling
Department of Physics
University of Wuppertal, 42097 Wuppertal, Germany
E-mail: \{ lippert, orth, schilling \}@theorie.physik.uni-wuppertal.de

The determination of the masses of the light quarks requires solving the hadronic binding problem of quantum chromodynamics (QCD), the fundamental theory describing the strong interactions between the quarks. Due to asymptotic freedom, the high-momentum sector of QCD can be treated by perturbative methods, however, on the low energy hadronic scale the running coupling of QCD becomes large. As a consequence, perturbative methods are not adequate for the hadronic binding problem, the scenario of which is dominated by the QCD vacuum state. The only known way to solve rather than model the problem is through ab-initio numerical simulations by high-end computers on a finite discrete space-time lattice. In this paper, which is intended for a non-expert audience, we give an illustrated introduction to the hadronic binding problem and its treatment by computer simulation on Tera-scale computers. We focus on the finite-size approach as pursued by the GRAL project to Go to more Realistic And Lighter quark masses, the main obstacle to current simulations.

1 The Hadronic Binding Problem

Half a century of experimental research in elementary particle physics has revealed a plethora of information on the properties of sub-nuclear matter. Symmetry, the primary building principle of twentieth century physics, allowed Gell-Mann and Ne’eman to arrange the complicated empirical patterns of the particle states in terms of higher representations of a special unitary group—the so-called flavor SU(3), SU(3)_f. At this time, the three unknown entities associated with the fundamental representation of SU(3)_f, according to Gell-Mann named up (u), down (d) and strange (s) quarks, were considered as purely mathematical objects. According to the first naive quark model, nucleons like the proton and the neutron are composed of three constituent quarks as depicted in Fig. 1a, while mesons are built up from a quark joined by its anti-quark, see Fig. 1b. At the time the simple quark model was quite an un-orthodox proposal as quarks carry only a fraction of the electron’s charge\(^1\).

It then took physicists nearly ten years to formulate a dynamical description of the interaction which binds quarks into hadrons. The key step was to generalize the principle of local gauge invariance, based on the abelian group U(1), that has been so successful for the description of electrodynamics, to a non-abelian form, the so-called color SU(3), SU(3)_c: In addition to the electric charge as carried by the electron, quarks come with three new charges conventionally labeled by colors. And the electromagnetic force between electrons, mediated by the uncharged photon, is generalized to the strong force between the quarks, resulting from the exchange of gluons basically coming in eight different color charge combinations. The number eight is reflected in the dimension of the adjoint representation of SU(3)_c. Fig. 2 shows a cartoon of a gluon exchanging the colors between two of the three constituent quarks within the proton.

Actually, the picture in Fig. 2 is not at all the entire story as it does not capture the
phenomenon of quantum fluctuations. In fact, it was one of the great triumphs of relativistic quantum field theory to provide a very accurate explanation for the tiny Lamb-shift, which is—according to quantum electrodynamics (QED)—due to quantum fluctuations that lift the degeneracy of the $2s_{1/2}$ -- $2p_{1/2}$ levels of hydrogen contrary to the previous predictions of Dirac’s equation. In QED, a photon, for instance, can virtually fluctuate into an electron-positron pair that exists for a short time in accordance with Heisenberg’s uncertainty principle $\Delta E \Delta T \approx \hbar$, as schematically depicted as Feynman diagram in Fig. 3a.

In QCD one would expect quantum fluctuations to be all the more important since due to the strong coupling, $\alpha_s$, which is by about two orders of magnitude larger than the electromagnetic coupling $\alpha = e^2/4\pi$, the probability for pair creation is much larger. Therefore, the proton is not just made up of constituent quarks and gluons but is a “soup” of quarks, anti-quarks and gluons. From the outside the proton looks like a particle carrying an effective quark number of 3, an electric charge 1 and color charge 0. Theoretically, it often is a good approximation to distinguish between “valence” quarks and a sea of virtual quark-anti-quark pairs and gluons, in order to explicitly express the appearance of a net quark number. An artist’s impression of the sea of quarks, anti-quarks and gluons is given in Fig. 4.

All our experimental evidence shows that quarks cannot be isolated but always form bound hadronic states with color charge zero, a phenomenon denoted as confinement of quarks. Confinement prevents us from direct experiments with quarks. In particular, we are not able to determine the masses of quarks as for instance the masses of isolated electrons.

Why then bother about the masses of quarks, if such objects cannot be isolated anyway? Well, the clue to further progress in elementary particle physics is an understanding of the violation of quark flavor symmetry in weak interactions, as for instance evident in the decay $K \to \pi\pi$. Within the Glashow-Salam-Weinberg (GSW) theory, weak interactions between quarks are readily formulated as the exchange of the massive gauge bosons $W^+$, $W^-$ and $Z$. A quantitative understanding of flavor dynamics therefore requires knowl-
edge of the properties of the quarks. In fact their masses are important input parameters of the current standard model (SM) of elementary particles, a combination of QCD and GSW. Accessing the quark masses thus implies solving the problem of how quarks, the fundamental objects, bind into hadrons, the "elementary" particles of the pre-quark-era.

Let us discuss the binding problem as a Gedanken-experiment by considering the sim-
ple case of the hydrogen atom. Suppose that electrons occur only confined within atoms. How could we still determine the mass \( m_e \) of such elusive electrons? The answer is to solve Schrödinger’s equation, which predicts the energy levels

\[
E_n(m_e) = -\frac{1}{2}m_e c^2 \frac{\alpha^2}{n^2}
\]

(1)

and charge radii

\[
r_n(m_e) = \frac{n^2 \hbar}{\alpha m_e c}.
\]

(2)

From the measurements of the transition energy \( E_2 - E_1 \) and Bohr’s radius one can determine the mass of the “confined” electron using the ratio \( \frac{(E_2 - E_1)(m_e)}{r_n(m_e)} \).

Quite analogously, if in QCD the theoretician could determine hadron masses \( M_{h_i}(m_{q_j}) \) as functions of the quark masses \( m_{q_j} \) (the indices \( i, j \) numbering hadrons and quarks, respectively), it would be possible to infer the quark masses from ratios of the experimental values of \( M_{h_i} \). This finally brings us to the hadronic binding problem, which is considerably more complex than the atomic one. While the electron carries a mass of 511 keV and the hydrogen binding energy is of the order of 10 eV only, we are faced here with very heavy bound states (with a nucleon of 931 MeV) composed of very light quarks (\( u \) and \( d \) carry masses of 1 to 10 MeV). Thus, contrary to atomic scale physics, the femtoworld of hadrons is a relativistic world *par excellence*, where in accordance with Einstein’s equation \( E = mc^2 \) the nucleon mass mostly consists of binding energy.

Given this situation, the obvious way to tackle the binding problem of the strong interaction is to start out from the ubiquitous sea of strongly interacting virtual gluons and quark-antiquark pairs as the basic entity. The state of lowest energy of this medium is denoted as the QCD “vacuum” state. Knowledge about the vacuum state would enable us to determine the propagation of quarks through the quark-gluon sea and subsequently to determine the properties of hadrons composed of quarks interacting through the medium. Note that this scenario is quite different from the above-mentioned situation of the Lamb-shift: there, the quantum fluctuations are computed as a perturbation of eigenstates of the hydrogen atom, which are a simple solution of Dirac’s equation. On the other hand, in QCD, hadrons themselves can be considered as small distortions of the vacuum. The determination of the QCD vacuum, however, as might already be clear from Fig. 4a, is an extremely complicated computational task.

It is well known that the propagation of electrons and photons through the vacuum of QED can be treated to high accuracy by low-order power expansion in \( \alpha \), as \( \alpha \ll 1 \). At the hadronic energy scale, however, the running coupling of QCD becomes large \( \mathcal{O}(1) \), and consequently the QCD vacuum state turns out to be a non-perturbative problem.
Figure 4. (a) Artist’s snapshot of the fluctuating quark-antiquark-gluon sea. (b) Same situation on the lattice. The quark fields are restricted to the lattice sites, the gluonic fields are exchanged only along the links of the lattice.
Currently, the only known way to solve the hadronic binding problem beyond modeling is through ab-initio numerical simulations on high-end computers using a finite discrete space-time lattice. The lattice essentially serves two purposes, namely (a) to provide a suitable regularization scheme for the quantum field theory and (b) to make the problem computationally tractable by shifting it from real-space into a finite box.

(a) The need for regularization is well known in the context of the perturbative analysis of quantum field theories. In that case infinities arise in the computation of higher-order Feynman diagrams which have to be compensated for by adding infinite counter terms. The integration process itself must of course avoid infinities. The regularization prescription renders the integrals finite by means of a regularization parameter (corresponding to a high-momentum cut-off), which is subsequently taken to infinity. Analogously we cannot implement a non-countable, infinite number of space-time points on a computer. A lattice with regularization parameter \( a \), the lattice spacing, renders the space-time points countable, and by Fourier transform we recognize that the inverse lattice spacing \( a^{-1} \) is just a cut-off at large momentum. Of course one has to send the cut-off to zero, \( a \to 0 \), performing a sequence of simulations, in order to be able to extrapolate eventually to continuum results—the so-called continuum limit.

(b) It is evident that we cannot handle an infinite number of lattice sites on a computer, even if it is countable. This fact is a fundamental problem for interactions with infinite range like, for instance, gravity. Fortunately, strong interactions are short ranged, i.e., they fall off rapidly enough outside the proton, a feature already described by the early effective Yukawa theory of pion exchange between nucleons. Thus, we are allowed to restrict our lattice laboratory to a finite box containing the hadron to be investigated. A finite box size \( L \) essentially serves as a low-momentum cut-off. Below, we shall demonstrate how to take advantage from the finite size \( L \) of the lattice by provoking finite-size-effects.

The analogous picture to Fig. 4a for the lattice world is presented in Fig. 4b. The quark fields are restricted to the sites of the lattice, gluons are restricted to the links.

We describe quantum fluctuations of this lattice system by considering the superposition of all its conceivable wave mechanical time evolutions (“paths”). In analogy to the famous Huygens principle of light propagation, the underlying wave mechanics is characterized by a phase factor, \( \exp\left(\frac{-iS}{\hbar}\right) \), where the functional \( S[\bar{\psi}, \psi, A] \) is nothing but the QCD action. This leads to the “path integral” which is a functional integral in the configuration space \( [\bar{\psi}, \psi, A] \):

\[
Z = \int D\bar{\psi}_x D\psi_x e^{-\frac{iS}{\hbar}[\bar{\psi}, \psi, A]}.
\]  

(3)

Actual lattice calculations use the Euclidean form of quantum field theory, which renders the path integral a real-valued partition function, as known from statistical mechanics. The Euclidean form of Eq. (3) is achieved by the analytic continuation of the time variable \( t \to -i\tau \). The ensuing effect is a transformation of the Minkowski metric into a Euclidean metric, while a positive definite Boltzmann weight \( \exp(-\beta S) \) emerges. This form of the path-integral is well suited for statistical evaluation,

\[
Z = \int \left( \prod_{n, \mu} [dU_{\mu}(n)][d\bar{\psi}_n][d\psi_n] \right) e^{-\beta S_{[U]} - S_{[\bar{\psi}, \psi, U]}}.
\]  

(4)
$S_g$ is the action due to the gluon fields, $S_f$ is the part of the action due to the quarks. The gluonic field is described by SU(3) matrices $U_{\mu}(n)$ extending from lattice point $n = (n_1, n_2, n_3, n_4)$ to $n = (n_1, n_2, n_3, n_4) + e_{\mu}$, $\psi_n$ denotes the quark fields located at the sites $n$. As the latter are Grassmann variables, and $S_f$ is a bilinear form, the fermionic variables can be integrated out, leaving us with

$$Z = \int \prod_{n, \mu} [dU_{\mu}(n)] \det(M[U]) e^{-\beta S_g}.$$  \hspace{1cm} (5)

The Euclidean path-integral Eq. (5) is evaluated by means of stochastic Monte Carlo algorithms. It suffices to generate a representative ensemble of fluctuating vacuum gauge fields $\{U_i\}$, $i = 1, \ldots, N$ which incorporate the effects of virtual fermion-anti-fermion pairs via the determinant of $M$, the lattice Dirac matrix. The ensemble of such vacuum states contains the complete physical information as pictured in Fig. 4. Given the ensemble of say $N$ vacuum configurations, any observable, like operators which describe valence quarks propagating in the vacuum or even objects built from sea quarks loops, can be “measured” as ensemble average together with their statistical error,

$$\langle O \rangle = \frac{1}{N} \sum_{i=1}^{N} O_i[U_i] \quad \text{and} \quad \sigma^2_{O} = \frac{2}{N} \left( \frac{1}{N} \sum_{i=1}^{N} |O_i[U_i]|^2 - \langle O \rangle^2 \right).$$  \hspace{1cm} (6)

Let us for the moment set $\det(M)$ equal to 1 in Eq. (5). In this case, the vacuum state is of purely gluonic nature. Lattice QCD based on such a kind of vacuum states is called “quenched approximation” because sea quarks are decoupled. Often this approximation describes the physics situation accurately enough. The advantage is that such a quenched simulation is substantially cheaper than the full one, because one can make use of a simple Markov transition by applying stochastic modifications link-wise to the variables $U_{\mu}(n) \to U'_{\mu}(n)$ with a local Metropolis decision.

However, using Metropolis for full QCD, the Metropolis decision

$$P(U \to U') = \min \left[ 1, \exp\left( -\Delta S_g \frac{\det(M[U'])}{\det(M[U])} \right) \right]$$  \hspace{1cm} (7)

requires the computation of the fermionic determinant for each link $U_{\mu}(n)$ separately, which is prohibitively expensive. A better way to proceed is to compute a global update of the links as achieved by the ingenious hybrid Monte Carlo algorithm\(^3\), the standard method for simulations of full QCD. Nevertheless realistic full QCD simulations require computers of the class of hundreds of teraflops, as we shall see next.

Let us illustrate these statements by presenting a picture of a typical lattice QCD experiment in Fig. 5, the determination of the flux tube between two heavy quarks that are being drawn apart. The underlying gauge configurations have been generated in the quenched approximation. Subsequently, the potential was computed on these configurations by means of a simple operator, the Wilson loop. The color encodes the result for the action density. Note the fluctuations due to the stochastic method. As the problem has a cylindric symmetry, it can be plotted in two dimensions. The sequence of sheets corresponds to increasing quark-antiquark separations.

The energy between the quarks grows linearly with their distance—confinement at work! As we did not include sea-quark loops, one expects the flux tube not to break up into two mesons.
3 The Multi-Scale Problem of QCD

Our previous considerations already have shown how lattice gauge theory merges quantum field theory with the techniques of statistical physics. The approach to the continuum limit \((a \to 0)\) is carried out on a sequence of lattices with decreasing spacings \(a\). In order to vary \(a\), we have to tune the parameter \(\beta \propto \frac{1}{g^2}\) suitably—with \(g\) being the bare strong coupling constant. For \(a \to 0\) lattice physics would become insensitive to the spacing of the lattice. Such behavior typically occurs near a second-order phase transition, where the correlation length diverges. In practice, one never can achieve \(a = 0\). Hence, results from different finite lattice spacings are extrapolated to \(a = 0\).

In lattice QCD, a correlation length is associated to an inverse physical mass, \(\xi = 1/m\). Our box size \(L\) must fulfill \(\xi \ll L\), so that \(\xi\) can be accommodated in the finite box. The lightest particle of the strong spectrum with the largest correlation length is the \(\pi\) meson. Therefore, \(L\) must be larger than \(\xi_\pi\), because otherwise finite size effects will spoil the results.

On the other hand, the lattice spacing \(a\) has to be chosen small enough to resolve the objects to be investigated. However, decreasing the lattice spacing decreases the finite box size unless we compensate for this effect by increasing the number of lattice sites. The crucial point is that we are limited in \(L/a\) by the available computer power. Presently, we can only choose between too coarse a lattice spacing \(a\) or too large a pion mass. We see
that lattice QCD is a typical multi-scale problem.

Future realistic QCD simulations with dynamical fermions would require to operate beyond the $\rho \rightarrow \pi \pi$ decay threshold and closer to the continuum limit than achieved so far. In order to estimate the simulation efforts we have determined the costs of full QCD simulations from the results of our large scale simulation projects SESAM and T\(\chi\)L. In these lattice experiments we have used the hybrid Monte Carlo algorithm with two degenerate flavors of Wilson fermions.

SESAM/T\(\chi\)L has generated 10 ensembles of full QCD vacuum configurations with \(O(5000)\) HMC trajectories each, at \(\beta = 5.6\) and 5.5 in the region \(0.57 < \frac{m_\pi}{m_\rho} < 0.85\). The lattice sizes are \(16^3 \times 32\) (SESAM) and \(24^3 \times 40\) (T\(\chi\)L), corresponding to physical sizes of \(1.372(36)\) fm (SESAM) and of \(1.902(34)\) fm (T\(\chi\)L) after chiral extrapolation. Running primarily on APE100 systems of the NIC at DESY/Zeuthen, DFG/Bielefeld and INFN/Rome, the total costs of the simulations sum up to about 0.06 Tflops-yrs. The configurations which have been generated on APE100 systems are stored at the ZAM/FZ-Jülich storage facilities for detailed physics evaluation on the Cray T3E systems.

In table 3 we try to convey to the reader an idea about the characteristics of full QCD lattice simulations by presenting some key quantities from SESAM/T\(\chi\)L. \(\frac{m_\pi}{m_\rho}\) is the value of the \(\pi\) mass divided by the \(\rho\) mass as obtained in the simulation (the experimental value being 0.1724). The third row shows the so-called integrated autocorrelation time \(\tau_{\text{int}}\) that determines the actual statistical significance of a Markov chain based simulation, see Eq. (6). These numbers tell us that we have about 200 stochastically independent configurations per ensemble.

<table>
<thead>
<tr>
<th>(\beta)</th>
<th>(m_\pi/m_\rho)</th>
<th>(T_{\text{equi}})</th>
<th>(\tau_{\text{int}})</th>
<th>(m_\pi/m_\rho)</th>
<th>(T_{\text{equi}})</th>
<th>(\tau_{\text{int}})</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.6</td>
<td>0.83</td>
<td>5200</td>
<td>19(4)</td>
<td>0.81</td>
<td>5400</td>
<td>25(6)</td>
</tr>
<tr>
<td></td>
<td>0.76</td>
<td>5250</td>
<td>33(4)</td>
<td>0.68</td>
<td>4950</td>
<td>36(4)</td>
</tr>
<tr>
<td>5.5</td>
<td>0.85</td>
<td>3500</td>
<td>19(2)</td>
<td>0.80</td>
<td>4000</td>
<td>24(3)</td>
</tr>
<tr>
<td></td>
<td>0.75</td>
<td>5000</td>
<td>38(2)</td>
<td>0.68</td>
<td>5000</td>
<td>47(3)</td>
</tr>
</tbody>
</table>

Table 1. Some characteristic quantities from SESAM/T\(\chi\)L simulations.

From our determination of autocorrelation times and the CPU time measurements we are in the position to estimate the costs of full QCD simulations. Suppose we aim for an accuracy comparable to state-of-the-art quenched simulations. The standard has been set by the CP-PACS group at Tsukuba/Japan. They have carried out quenched simulations\(^4\) in 1997 on the 600 Gflops CP-PACS/Hitachi SR2001 parallel system built at Tsukuba university. They achieved finite \(a\) results for light hadrons with errors < 1% and subsequently could extrapolate to continuum results with errors between 1 and 3%. Using their setting, we find the upper bounds to the CPU time (see table 3) needed to carry out an analogous simulation with \(n_f = 2\) Wilson fermions. We conclude that such a full QCD simulation will be a task for a 100 Teraflops system.
Table 2. Extrapolation of costs for the determination of light hadron masses with 2 flavors of Wilson fermions in analogy to the quenched setting of the CP-PACS group.

<table>
<thead>
<tr>
<th>$a$ (fm)</th>
<th>$\frac{m}{m_\pi}$</th>
<th>0.75</th>
<th>0.70</th>
<th>0.60</th>
<th>0.50</th>
<th>0.40</th>
<th>$\sum$ [Tflops-yrs]</th>
<th>$z_2$</th>
<th>$L/a$</th>
<th># of confs</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.102</td>
<td>0.46</td>
<td>0.79</td>
<td>0.21</td>
<td>0.58</td>
<td>18</td>
<td>25(8)</td>
<td>4.3</td>
<td>32</td>
<td>800</td>
<td></td>
</tr>
<tr>
<td>0.076</td>
<td>1.2</td>
<td>1.8</td>
<td>3.5</td>
<td>7.0</td>
<td>15</td>
<td>28(9)</td>
<td>2.8</td>
<td>40</td>
<td>600</td>
<td></td>
</tr>
<tr>
<td>0.064</td>
<td>3.7</td>
<td>6.7</td>
<td>12</td>
<td>23</td>
<td>50</td>
<td>95(35)</td>
<td>2.8 (!)</td>
<td>48</td>
<td>400</td>
<td></td>
</tr>
<tr>
<td>0.050</td>
<td>17</td>
<td>28</td>
<td>60</td>
<td>120</td>
<td>260</td>
<td>485(150)</td>
<td>2.8 (!)</td>
<td>64</td>
<td>200</td>
<td></td>
</tr>
</tbody>
</table>

4 Going Realistic And Light–The GRAL Project

What ways are there at present to extend the range of full QCD simulations towards smaller quark masses? A crucial parameter for the costs of QCD simulations is the number of lattice sites, $(L/a)^4$. In previous simulations $L/a$ has in general been chosen such that finite-size-effects were largely suppressed at given quark mass.

The idea of our new project called GRAL (Going Realistic And Light) is to reverse the strategy and to perform a comprehensive study of finite-size-effects. Deliberately accepting finite-size-effects allows one in principle to extend full QCD simulations into the regime of lighter quark masses. For given bare coupling and quark mass, QCD vacua will be generated on a sequence of box sizes that will allow for an extrapolation in $L/a$. For such a finite-size-scaling analysis we expect a formula like that of Fukugita, Parisi et al.\(^5\) to be applicable: $m(L) = m_\infty + cL^{-\nu}$.

Figure 6. Infinite volume extrapolation.

Before embarking on a large-scale production, however, this program requires an extensive scanning of parameter space (spanned by the bare coupling $\beta$, the bare quark mass and the number of lattice sites) in order to locate the physically interesting and computationally feasible operating region.
To this end we specify, for a start, \( L/a = 16 \) as the maximal spatial extent of a series of lattices employed. Furthermore we require the simulation on this lattice to yield meson masses with \( \frac{m_\pi}{m_\rho} = 0.4 \) (i.e. below the decay threshold at 0.5) and a finite-size-parameter of \( \xi_\pi / L = 0.2 \). This particular value for \( \xi_\pi / L \) has been chosen in view of a recent SESAM/T\( \chi \)L quark mass analysis\(^6\), where 0.2 was found to mark the onset of significant finite-size-effects in the hadronic spectrum.

In summary, the envisaged properties of our target point on a \( 16^3 \times 32 \)-lattice are as follows:

<table>
<thead>
<tr>
<th>( \frac{m_\pi}{m_\rho} )</th>
<th>( \xi_\pi )</th>
<th>( L ) [fm]</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.4</td>
<td>0.2</td>
<td>3.2</td>
</tr>
<tr>
<td></td>
<td>0.31</td>
<td>0.78</td>
</tr>
<tr>
<td></td>
<td>0.116</td>
<td>1.86</td>
</tr>
</tbody>
</table>

The lattice spacing \( a \) has been estimated on the basis of the SESAM and T\( \chi \)L data at \( \beta = 5.6 \) and 5.5. As can be seen from the table, the price for sticking to \( L/a = 16 \) is a fairly coarse lattice. According to our scaling cost extrapolations, we expect a total cost of approximately 0.2 TFlop/s \cdot h to produce 100 statistically independent gauge field configurations at the target point. For more details see Ref. 7.

With shrinking box size local fluctuations in the vacuum field configurations will eventually destabilize the global updating process of the Hybrid Monte Carlo Algorithm. The crucial issue regarding the feasibility of our strategy is the question of when this breakdown occurs. Thus, our first task will be to establish a suitable “window” of lattice volumes were our sampling algorithm remains stable.

5 Summary

With this contribution we aimed at illustrating the lattice approach to the determination of hadronic properties for a non-expert audience. We argued that the hadronic binding is a highly relativistic problem, where quarks with small masses form nucleons which are about a factor of 100 heavier. We have emphasized the important role of the QCD vacuum fields. Such fields can be generated by lattice gauge simulations. Given an ensemble of vacuum gauge fields we can compute the propagation of quarks that combine to hadrons through the fluctuating quark-gluon sea. We have estimated the huge costs of future simulations of the full theory, assuming accuracies similar to those of state-of-the-art quenched simulations. These considerations demonstrate that new ideas are required to approach the regime of small realistic quark masses. One promising idea is pursued by our GRAL project, which means to Go Realistic And Light using finite-size techniques. In this manner we will attempt to go beyond the previous milestones of SESAM/T\( \chi \)L.

Acknowledgments

We acknowledge the essential support of the SESAM/T\( \chi \)L projects and the current GRAL project by the John von Neumann Institute for Computing (former HLRZ) with about 150 TFlops-hrs compute time on APE100 computers at DESY/Zeuthen and about 50 TFlops-hrs on the Cray T90 and T3E systems at Jülich for post-processing of our field configurations. The availability of multi-Tbyte storage space at ZAM/FZ-Jülich played an essential rôle
in our Europe-wide simulation project. We heartily thank the members of these computer centers for their friendly support.

We thank G. Bali and C. Schlichter for providing the quark-action image and the members of SESAM/TχL, in particular W. Schroers, for important contributions and discussions.

References


Fundamental Parameters of QCD

ALPHA Collaboration, Rainer Sommer\textsuperscript{1} and Hartmut Wittig\textsuperscript{2}

\textsuperscript{1} DESY
Platanenallee 6, 15738 Zeuthen, Germany
E-mail: rainer.sommer@desy.de

\textsuperscript{2} DESY
Notkestrasse 85, 22603 Hamburg, Germany
E-mail: hartmut.wittig@desy.de

The theory of strong interactions, QCD, is described in terms of a few parameters, namely the strong coupling constant $\alpha_s$ and the quark masses. We show how these parameters can be determined reliably using computer simulations of QCD on a space-time lattice, and by employing a finite-size scaling method, which allows to trace the energy dependence of $\alpha_s$ and quark masses over several orders of magnitude. We also discuss methods designed to reduce the effects of finite lattice spacing and address the issue of computer resources required.

1 The Standard Model of Particle Physics

In the last few decades, particle physicists have explored the fundamental forces down to distance scales of $\approx 10^{-18}$ m. It was found that the experimental observations are described to very high accuracy by a theory which is known as the Standard Model of particle physics. During the 1990s in particular, the predictions of this theoretical framework have been put to very stringent tests in accelerator experiments across the world. Perhaps one of the most impressive examples of its predictive power, the line shape of the $Z$-resonance in $e^+ e^-$ scattering, is shown in Fig. 1.

The Standard Model describes the interactions of the fundamental constituents of matter through electromagnetic, weak and strong forces in terms of three different quantum gauge theories. The success of the Standard Model is not only a consequence of the mathematical simplicity of its basic equations, but also because the forces they describe are relatively weak at the typical energy transfers in current experiments of about $10 - 100$ GeV.\textsuperscript{a} The strengths of the interactions are characterized by so-called coupling constants. When the forces are weak, the predictions of the theory can be worked out in terms of an expansion in powers of these coupling constants, a procedure known as perturbation theory. For instance, in Quantum Electrodynamics (QED), the quantum gauge theory describing the interactions between electrons and photons, the coupling constant is the well-known fine structure constant $\alpha \approx 1/137$. Its smallness guarantees that only a few terms in the power series are sufficient in order to predict physical quantities with high precision.

The gauge theory for the strong force is called Quantum Chromodynamics (QCD), in which quarks and gluons assume the rôles of the electrons and photons of QED. Quarks are the constituents of the more familiar protons and neutrons. The coupling constant of

\textsuperscript{a}In particle physics it is customary to use “natural units” where the speed of light, $c$ and Planck’s constant, $\hbar$ are set to one and energies as well as masses are given in GeV. As an orientation note that $m_{\text{proton}} \approx 1$ GeV, where 1 GeV $= 1.602 \cdot 10^{-7}$ J.

149
Figure 1. The line-shape of the Z-resonance in $e^+e^-$ scattering as a function of the center of mass energy. The theoretical prediction is shown under the assumption that nature contains $N_\nu$ different neutrinos.

QCD, $\alpha_s$, then characterizes the strength of the interaction between quarks and gluons in a similar manner as the fine structure constant does in QED. One important property of all coupling “constants” in the Standard Model is that they depend on the energy transfer $\mu$ in the interaction process. In this sense they are not really constant, and one usually refers to them as couplings that “run” with the energy scale. For instance, at $\mu \approx 100$ GeV the strong coupling constant has been determined as $\alpha_s = 0.12$. Although this is much larger than the fine structure constant of QED, the method of perturbation theory still works well. However, if the energy scale $\mu$ is decreased from 100 GeV it is found that the value of $\alpha_s$ increases. In fact, at $\mu \approx 1$ GeV it becomes so large that perturbation theory cannot be relied upon any more. It is then obvious that particle theorists require a tool which is able to deal with large values of $\alpha_s$. In other words, what is needed is a non-perturbative method to work out the predictions of QCD in this situation.

As mentioned before, the simple and elegant theory of QCD is formulated in terms of quarks and gluons. Evidence for their existence has been accumulated in scattering experiments at high energies. Yet what is observed in experiments at low energies, say, $\mu \lesssim 1$ GeV are protons, neutrons, $\pi$-mesons and many other particles, all known as hadrons. In fact, a striking property of QCD is “confinement”, which means that quarks and gluons cannot be produced in experiments. Intuitively this property of QCD can be understood in terms of the strong growth of forces between quarks as their separation in-
creases. Thus, the only directly observable particles are the bound states of quarks and gluons, i.e. the hadrons.

Particle physicists are then faced with the task of connecting the theoretically rather simple regime of QCD at high energies with the properties of protons, $\pi$-mesons and other hadrons observed at low energies. This task is made all the more difficult since analytic methods such as perturbation theory fail completely to describe the world of hadrons. This has led to the development of numerical techniques, namely computer simulations of QCD formulated on a discrete lattice of space-time points. This method allows for a non-perturbative treatment of the theory in the low-energy regime.

The basic reason why the two regimes can be connected is the fact that QCD contains only the fundamental gauge coupling $\alpha_s$ and the masses of the quarks as free parameters. All observables, such as the mass of the proton, can in principle be predicted in terms of these quantities. The ALPHA Collaboration has embarked on a project to connect low and high energies in practice, by means of extensive computer simulations of lattice QCD. Following the original idea of $^1$, new methods have been developed and applied, which have yielded rather precise results. It is then possible to turn the tables: starting from the accurate experimental information on the properties of hadrons, the fundamental parameters of QCD can be determined numerically.

In particular, the high-energy behaviour of the strong coupling constant $\alpha_s$ is given by

$$\alpha_s^{-1} \propto \ln \left( \frac{\mu}{\Lambda} \right),$$

(1)

This implies that the energy dependence of $\alpha_s$ is specified completely in terms of a single parameter, $\Lambda$, measured in energy units. It is only natural to take $\Lambda$ as a basic parameter of QCD instead of the energy-dependent $\alpha_s$. The ratio of $\Lambda$ to the proton mass is computable in lattice simulations – but not in perturbation theory. Thus, one of the main goals of the ALPHA Collaboration is the precise determination of $\Lambda$.

With QCD being one of the pillars of the Standard Model, it is clear that the precise knowledge of its parameters such as $\Lambda$ is important for the ongoing quest for generalizations of the Standard Model. Such more complete theories are needed to describe the early stages of the universe and are also expected to be relevant at energies which will be accessible at future particle colliders.

## 2 Lattice QCD

In the mathematical formulation of QCD, the basic quantities are quark and gluon fields, which are functions of the space-time coordinates $x_\mu$, with $x_0$ identified with time. The classical field equations, which describe their dynamics, are differential equations – generalizations of Maxwell’s equations for electromagnetism. To allow for a numerical treatment, it is then natural to discretize the differential
equations with a discretization length $a$, termed the lattice spacing. This procedure turns differential operators into finite difference operators and the fields are defined only at the points of a space-time lattice, illustrated in Fig. 2.

Quantization is achieved by Feynman’s path integral representation. It involves integrations over all degrees of freedom weighted with the exponential of the classical action. Let $\Omega$ denote an observable, represented e.g. by a combination of quark and antiquark fields. Its expectation value, $\langle \Omega \rangle$, is defined as

$$\langle \Omega \rangle = \frac{1}{Z} \int D[U] D[\psi, \bar{\psi}] \Omega e^{-S_G[U]-S_F[U, \psi, \bar{\psi}]} ,$$

where $Z$ is fixed by the condition $\langle 1 \rangle = 1$. To further prepare for an evaluation of the path integral on a computer, the quark degrees of freedom are integrated out analytically. The expression for $\langle \Omega \rangle$ then becomes

$$\langle \Omega \rangle = \frac{1}{Z} \int D[U] \Omega_{\text{eff}} \{ \text{det}(D[U]) \}^{N_f} e^{-S_G[U]} .$$

$\Omega_{\text{eff}}$ denotes the representation of $\Omega$ in the effective theory, where only gluon fields remain in the path integral measure. Equation (3) requires some further explanation:

- $D[U]$ denotes the Dirac operator, and for simplicity of presentation we have considered QCD with $N_f$ quarks of equal mass $m$, which enters $D[U]$. For hadron physics at small energies the case $N_f = 3$ is most relevant. The heavier quarks decouple from the dynamics up to small effects;

- The gluon field is represented by link variables $U(x, \mu)$ connecting the sites $x$ and $x + a \hat{\mu}$ (shown as $\square$ in Fig. 2). The measure $D[U]$ is the product of integration measures for each link;

- The action $S_G[U]$ is a sum of local terms over the lattice, coupling only gluon variables within one plaquette ( $\square$ in Fig. 2). By contrast, the effective interaction resulting from the integration over the anticommuting quark fields is of infinite range: although the finite difference Dirac-operator $D[U]$ is local, $\text{det}(D[U])$ couples gluons at arbitrary distances.

- Using the representation eq. (3), our expression for $\langle \Omega \rangle$ has the form of a thermal average in statistical mechanics, and one may use stochastic sampling to evaluate the integral if the space-time volume is made finite by restricting $0 \leq x_\mu < L$.

- As an example, consider $\langle \Phi(x) \Phi^\dagger(y) \rangle$ where $\Phi(x)$ is a suitable combination of quark fields at a point $x$, which has the quantum numbers of some hadron. The expectation value is then proportional to the quantum-mechanical amplitude for the propagation of the hadron from point $y$ to $x$, from which the mass of the hadronic bound state may be obtained.

The lattice formulation sketched above leads to a mathematically well-defined expression for $\langle \Omega \rangle$. In particular, the typical infinities which are encountered in field theoretical expectation values are absent.

The exact treatment of the determinant, eq. (3), still presents a major challenge, even on today’s massively parallel computers. In many applications one has therefore set $N_f = 0$. 

152
This — very drastic — approximation defines the so-called quenched approximation. Physically it means that the quantum fluctuations of quarks are neglected and only those due to the gluons are taken (exactly) into account. Although it turns out that the quenched approximation works well at the 10%-level, the proper treatment for $N_f > 0$ is perhaps the most important issue in current simulations.

Whether or not the quenched approximation is employed, one always has to address the problem of lattice artefacts. Let $R$ denote a dimensionless observable, such as a ratio of hadron masses. Then its expectation value on the lattice differs from the value in the continuum by corrections of order $a^p$:

$$R_{\text{lat}} = R_{\text{cont}} + O(a^p),$$

where the power $p$ depends on the chosen discretization of the QCD action. The correction term for typical values of $a$ can be quite large, and an extrapolation to the continuum limit is then required to obtain the desired result. Obviously, such an extrapolation is much better controlled if $p$ is large.

Let us now return to the problem of determining the energy dependence of quantities such as the running coupling $\alpha_s$. Obviously the numerical simulation is possible only if the number of points of the lattice, $(L/a)^4$, is not too large; typical lattice sizes are $L/a \lesssim 32$. Therefore, besides $a$ also the effect of the finite value of $L$ has to be considered. It is known rather well that $L = 2$ fm is sufficient for most quantities, and that much smaller volumes would lead to unacceptably large corrections to the desired $L \rightarrow \infty$ limit. A physical box size of $L = 2$ fm together with $L/a \lesssim 32$ thus implies

$$a \gtrsim 0.05 \text{ fm}. \quad (5)$$

Clearly the existence of a minimal distance $a$ also means that the energies, $\mu$, that can be expected to be treated correctly have to satisfy

$$\mu \ll a^{-1} \lesssim 4 \text{ GeV}. \quad (6)$$

In other words, lattice QCD is well suited for the computation of low energy properties of hadrons, while high energies appear impossible to reach. The ALPHA Collaboration has developed and applied an approach to circumvent this problem, which is described below.

### 3 Running Coupling and Quark Masses

We are now going to outline the strategy which allows to connect the low- and high-energy regimes of QCD in a controlled manner. Results for the $\Lambda$ parameter and the quark masses will then be presented. From now on we will drop the subscript “s” from the strong coupling constant $\alpha_s$.

#### 3.1 Gedankenexperiment

From the above it is evident that the restriction of numerical lattice QCD to low energies is necessary to avoid finite-size effects (FSE) when working with a manageable number of lattice sites. For the computation of an effective coupling, this problem can be circumvented since one has great freedom in the definition of such a coupling. Even the strength...
of finite-size effects serve as a measure of the interactions of the theory, and thus a suitably chosen FSE may be used to define an effective coupling \( \bar{g}(L) \) with \( \alpha(\mu) = \frac{\bar{g}^2(L)}{4\pi}, \mu = 1/L \). (7)

It depends on ("runs with") the energy scale \( \mu = 1/L \). Such a coupling can be computed in the regime \( \mu \ll \alpha^{-1} \) requiring only a moderate resolution of the space-time world; \( L/a = O(10) \) points per coordinate are sufficient.

In a numerical calculation, the success of this general idea will depend on a few properties of the coupling. It must be computable with good statistical precision in a Monte Carlo (MC) simulation, and with small discretization errors. Furthermore, one would like to know its scale dependence analytically for large energies. This is achieved by determining the perturbative expansion of the so-called \( \beta \)-function,

\[
\beta(\bar{g}) = -L \frac{\partial \bar{g}}{\partial L} \bar{g} \sim \bar{g}^3 \{ b_0 + b_1 \bar{g}^2 + b_2 \bar{g}^4 + \ldots \},
\]

(8)
to the order indicated above, or even higher.

In QCD, a coupling with these properties could indeed be found. Its definition starts from QCD in a box of size \( L^4 \). Periodic boundary conditions are imposed for the fields as functions of the three spatial coordinates, and Dirichlet boundary conditions are set in the time-direction, as shown in Fig. 3. The Dirichlet boundary conditions are homogeneous except for the spatial components of the gluon gauge potentials, \( A_k \). With this choice of boundary conditions the topology is that of a 4-dimensional cylinder.

Classically, these boundary conditions lead to a homogeneous (\( x \) and \( t \) independent) colour-electric field inside the cylinder. The walls at time \( t = 0 \) and \( t = T \) act quite similarly to the plates of an electric condensor. The strength of the QCD interactions is conveniently defined in terms of the colour-electric field at the condensor plates:

\[
\bar{g}^2(L) = \frac{E_{\text{classical}}}{\langle E \rangle}.
\]

(9)

Here \( E \) is a special colour component of the electric field.

For weak coupling, i.e. for small \( L \), the path integral is dominated by field configurations which correspond to small fluctuations about the classical solution. On the other hand, for \( L \approx 1 \) fm they may deviate significantly from the classical solution, and this can be realized in a MC-simulation of the path integral.
3.2 The Running Coupling

The energy dependence of the coupling $\bar{g}^2(L)$ can be determined recursively through a number of compute-effective steps. Figure 4 illustrates the implementation of the recursion

$$\bar{g}^2(2L) = \sigma(\bar{g}^2(L)).$$

The function $\sigma$ describes the change in the coupling when the physical box size is doubled. Since the relevant energy scale for the running of the coupling ($\mu = 1/L$) is separated from the lattice spacing $a$, one may compute $\sigma$ recursively over several orders of magnitude in $\mu$, whilst keeping the number of lattice sites at a manageable level, i.e. $L/a = O(10)$.

The results of one horizontal step in Fig. 4, performed on the APE-computers at NIC/Zeuthen, is displayed in Fig. 5. They are shown as a function of the resolution $a/L$. Because the coupling and the details of the discretization were chosen with great care, the dependence on the resolution is tiny and the lattice numbers can be extrapolated to the continuum $a/L \rightarrow 0$. To arrive at the continuum limit, each of the horizontal steps in the figure is indeed repeated several times with different resolutions!

In this way, the continuum function $\sigma(u)$ was obtained for a range of $u = \bar{g}^2(L)$. Starting from a minimal value of $u$, the recursive application of eq. (10) yields the points in Fig. 6. They may be compared to the integration of the differential equation (8) towards low values of $\mu$, starting at the lowest value of $\alpha$. One then makes the important observation, that eq. (8), truncated at the indicated order, is quantitatively verified in the region of low enough $\alpha$. It may hence also be used to compute the
Figure 6. The energy dependence of the coupling as computed from lattice QCD (points) compared to the perturbative expression eq. (8) with the $b_2$-term (full line) and without.

$\Lambda$-parameter from $(\bar{\gamma} = \bar{\gamma}(L = 1/\mu))$

$$\Lambda = \mu \left( b_0 \bar{\gamma}^2 \right)^{-b_1/2b_0} \exp^{-1/2b_0\bar{\gamma}^2} \times \exp \left\{ - \int_0^{\bar{\gamma}} d\gamma \left[ \frac{1}{\beta'(\gamma)} + \frac{1}{b_0\bar{\gamma}^2} - \frac{b_1}{b_0\bar{\gamma}} \right] \right\} ,$$

with negligible errors due to higher-order terms that are not included.

The attentive reader will have noticed that the computation explained so far, “only” determines the dependence of $\alpha$ on the combination $\mu/\Lambda$, while in the figure we show it as a function of $\mu$ in physical units for $N_f = 0$, i.e. in the quenched approximation. The missing link is to connect the lowest energy $\mu = 1/L_{\text{max}}$ contained in the figure, to a low energy – experimentally accessible – property of a hadron. The connection to the decay rate of the K-meson, which we do not describe here, yields the physical units shown in the figure as well as the final result

$$\Lambda = 238(19) \text{ MeV for } N_f = 0 .$$

Currently, the ALPHA Collaboration is extending these computations to the numerically very demanding case of $N_f = 2$. First promising results, which are still awaiting a final check for the absence of discretization errors are shown on the r.h.s. of Fig. 6. The step to connect $L_{\text{max}}$ to a low-energy observable has yet to be performed, and hence the results are plotted as a function of $\mu/\Lambda$. It is worth pointing out that, in order to ensure efficiency and correctness of these $N_f > 0$ simulations, the development and testing of MC-algorithms is very important.

### 3.3 Quark Masses

The masses of the quarks are of a very different nature from the mass of the electron. The fundamental difference is the property of confinement mentioned in section 1. It means that a quark can not be prepared in isolation to perform an experimental measurement of its mass. Consequently, quark masses have to be understood as parameters of the theory, and their proper definition very much resembles that of an effective coupling. In particular, they have an energy dependence similar to the one of $\alpha$. When defined in a natural way,
the energy dependence is the same for all quark flavours; in other words, ratios of quark masses do not depend on the energy. The overall energy dependence has been determined with the same strategy as the one described for $\alpha$ and in fact with similar precision (see Fig. 2 in $^9$).

The numerical values of the quark masses (at a particular low energy scale) are conveniently extracted by relating them to the masses of quark-antiquark bound states with pseudoscalar quantum numbers: $\pi$, $K$, $D$, $B$-mesons. Results of the ALPHA Collaboration obtained in the $N_f = 0$ approximation are shown in Fig. 7 in two different (common) conventions. The upper row shows the running masses $m$ in the so-called $\overline{\text{MS}}$ scheme at a renormalization energy $\mu = 2 \text{ GeV}$, while the lower one shows the so-called renormalization group invariant quark masses. The latter are related to the running quark masses via

$$M = \lim_{\mu \to \infty} \overline{m}(\mu) \left[ \frac{2b_0}{(g(\mu))^2} \right]^{-d_0/2b_0},$$

where $d_0$ characterizes the asymptotic behavior of $\overline{m}(\mu)$ for large energy. It is computable in perturbation theory and has a value of $d_0 = 8/(4\pi)^2$.

4 Improvement

Equation (4) shows that lattice and continuum quantities differ by discretization errors. In Wilson’s formulation of lattice QCD these discretization effects are of order $a$, which can be quite large. People have therefore tried to find an improved version of the Wilson action, so that the deviation of observables computed on the lattice to their continuum values is of $O(a^2)$, or even higher.

$^b$The ratio of strange to light quark masses actually relies on$^{10}$. In the special case of the $b$-quark, whose mass is large compared to the QCD scale, $\Lambda$, an expansion in terms of $1/m_b$ has been used and only the lowest order term was considered$^{11}$. 

157
In order to illustrate the problem, let us consider the solution of differential equations by means of numerical integration. Given the initial value problem

\[ \frac{dy}{dx} = f(x, y), \quad y(x_0) = y_0, \]  

the well-known Euler method provides a formula for its numerical solution, i.e.

\[ y_{j+1} = y_j + a f(x_j, y_j), \quad j = 0, 1, 2, \ldots \]  

(15)

Here the quantity \( a \) denotes the finite step size used to approximate the derivative \( \frac{dy}{dx} \) in terms of finite differences. In order to compute the solution at point \( x = x_0 + na \) for a fixed value of \( a \) one has to perform \( n \) iteration steps. The central question then is by how much this solution deviates from the exact one. For the simple Euler method one can show that this so-called truncation error is proportional to \( a^2 \). Obviously the accuracy of the solution will be larger for small \( a \), but then the number of iterations, \( n \), may become very large.

It is much more efficient to use an improved solution scheme, such as the Runge-Kutta method. Here the function \( f \) in eq. (15) is replaced by a more complicated expression, which is chosen in such a way that the truncation error is proportional to \( a^5 \).

The analogy of this example with lattice QCD is obvious: the step size is now the lattice spacing, and the lattice action must be chosen such that the leading discretization error is of higher order in \( a \). However, in a quantum field theory like QCD the search for an improved action is not straightforward: although one can quite easily write down an action which formally has \( O(a^2) \) artefacts, the interactions between quarks and gluons in the discretized theory can again introduce \( O(a) \) errors at the quantum level.

To be more explicit, let \( S_W \) denote the Wilson action for lattice QCD. Sheikholeslami and Wohlert\(^{15}\) have shown that it is sufficient to add just one more interaction term, corresponding to an action \( \delta S \), to \( S_W \) in order to cancel its leading discretization error proportional to \( a \). To ensure that quantum effects do not introduce \( O(a) \) effects “via the back door”, the term \( \delta S \) has to be multiplied by a coefficient \( c_{sw} \). This improvement coefficient must be suitably tuned in order to achieve the complete cancellation of \( O(a) \) artefacts at the quantum level for observables such as hadron masses. The \( O(a) \) improved version of Wilson’s action then reads

\[ S_{SW} = S_W + c_{sw} \delta S. \]  

(16)

The ALPHA Collaboration has developed and applied a method to determine \( c_{sw} \) in computer simulations\(^{16–19}\). The method amounts to computing the expectation value of a pure lattice artefact as a function of \( c_{sw} \). The value of \( c_{sw} \) at the point where the artefact vanishes then defines the improved action for the corresponding value of the lattice spacing. Thus, non-perturbative methods are not only used to compute physical observables, but also to improve the reliability of the numerical treatment of the theory as such. The improved action is only slightly more complicated to implement in practical simulations. The increased effort is easily offset by the faster convergence to the continuum limit. As a result, the action with \( c_{sw} \) as determined in\(^{18,19}\) has by now become a standard for precision lattice QCD computations. In particular this improvement was essential in obtaining most of the results quoted in the previous section.

Of course, one also has to verify that observables computed with the improved action indeed approach the continuum limit with a rate proportional to \( a^2 \). Tests of this kind have been performed successfully (see, e.g. ref.\(^{20}\)), and other examples for the effectiveness
of the improvement programme for the evaluation of many hadronic observables can be found in\textsuperscript{5}.

5 Machines and Resources

Although conceptual advances like $O(\alpha)$ improvement are of great importance, progress in lattice QCD is also dependent on the availability of sufficient computer resources.

In the 1980s most simulations were performed on vector supercomputers like the Cray X-MP. Since then the demand for increased performance has led to the development of massively parallel machines, which are now widely used. Lattice QCD is a problem which lends itself easily to parallelization: the total volume can be divided into many sublattices, which are distributed over a grid of processors. The latter can perform the same task on independent data. Furthermore, the communication pattern for lattice QCD is simple, since most algorithms only require nearest-neighbour communications.

The ALPHA Collaboration has mostly used parallel computers from the APE family of machines\textsuperscript{21–24}. The latest generation, APEmille, has been developed jointly by INFN and DESY. The smallest entity in the APEmille processor grid consists of a cube of $2 \times 2 \times 2$ processors. These cubes are then connected to form larger grids. Despite its very conservative clock speed of 66 MHz, each processor achieves a peak performance of 528 MFlops, \textsuperscript{c} thanks to the optimization for complex arithmetic: the operation $a \times b + c$, which requires 8 floating point operations for the three 32-bit complex numbers $a$, $b$ and $c$, is performed in one clock cycle. The total peak speed of the current installation of APEmille machines at DESY-Zeuthen amounts to more than 500 GFlops (single precision).

The programming language for APEmille is TAO, which has a FORTRAN-like syntax, but also includes special features designed to facilitate parallelization and coding, and allows to achieve a high proportion of the peak speed. One such feature is the easy access to any desired number of the 512(!) registers. The typical efficiency of ALPHA-programs on APEmille is about 30\% of the peak speed. A similar figure has been achieved by other lattice QCD collaborations on machines like the Cray T3E, but at the price of having to code the core routines in assembler language\textsuperscript{25}. An interesting figure-of-merit is the price/performance ratio. For APEmille it is $(8-10)$ Euro per sustained MFlops and there are efforts under way to reduce this number much further. In addition it was also demonstrated that PC clusters with Myrinet network can achieve about $3-4$ Euro/MFlops sustained\textsuperscript{26}. These developments let us await the future with optimism.

6 Future

Moving towards the realistic case of $N_f = 3$ should soon be possible. Once this has been achieved, the most precise determinations of the fundamental parameters of QCD may come from the low energy hadron spectrum combined with lattice QCD to evaluate the theory. The methods developed in this project are expected to play an important rôle in this program.

Furthermore these methods and related ones\textsuperscript{27–29} will improve the reliability of the determination of properly normalized weak decay (and mixing) amplitudes of hadrons\textsuperscript{30}.

\textsuperscript{c}The unit 1 MFlops denotes one million floating point operations per second.
Their knowledge is again of vital importance in tests of the Standard model and the search for an even more fundamental theory.

The results presented in this article were obtained on the APE100 and APEmille installations at Zeuthen. The total CPU time was about $10^7$ processor hours on APE100 and $0.2 \cdot 10^7$ processor hours on APEmille, respectively.

References

Hadronic Structure from Lattice QCD

G. Schierholz

– QCDSF Collaboration –

John von Neumann - Institut für Computing NIC
Deutsches Elektronen-Synchrotron DESY, 15735 Zeuthen, Germany
Deutsches Elektronen-Synchrotron DESY, 22603 Hamburg, Germany

In this contribution I highlight recent lattice calculations of the nucleon structure functions and form factors with two flavors of dynamical quarks done by the QCDSF Collaboration.

Wisdom from Frank Wilczek

“... the Higgs particle (or the doublet) is certainly not – despite much loose talk to the contrary – the Origin of Mass. (Still less is it the God Particle, whatever that means.) Most of the mass of ordinary matter is concentrated in protons and neutrons. It arises from an entirely different, and I think more profound and beautiful, source. Numerical simulation of QCD shows that if we built protons and neutrons in an imaginary world with no Higgs mechanism – purely out of quarks and gluons with zero mass – their masses would not be very different from what they actually are. Their mass mostly arises from pure energy, associated with the dynamics of confinement in QCD, according to relation $m = E/c^2$. This profound account of the origin of mass is a crown jewel in our Theory of Matter.”

hep-ph/0101187

1 Introduction

The lattice formulation of QCD is at present the only known way of obtaining low energy properties of the theory in a direct way, i.e. without any model assumptions. Quantities within the grasp of lattice QCD involving light quarks include the hadron mass spectrum, quark masses, the $\Lambda$ parameter, the chiral condensate, the nucleon sigma term, meson decay constants, the axial and tensor charge of the nucleon, form factors and moments of the polarized and unpolarized structure functions of the nucleon, pion and rho.

Our group, the QCDSF Collaboration\(^a\), has been actively involved for the last few years in determining these quantities, all characterized by their non-perturbative nature. Rather

\(^a\)S. Capitani, M. Göckeler, R. Horsley, W. Kürzinger, D. Pleiter, P. Rakow, G. Schierholz, A. Schiller, and collaborators.
than giving an exhaustive progress report of recent developments in the field I shall, due to lack of space, focus on two topics: nucleon form factors (including the axial charge) and moments of structure functions, and how the lattice method can lead to their determination. The structure functions bear the answer to the question posed by Wilczek, namely how quarks and gluons provide the binding (mass) and spin of the nucleon.

The lattice approach involves first euclideanizing the QCD action and then discretizing space-time (with lattice spacing \( a \)). The path integral then becomes a very high dimensional partition function, which is amenable to Monte Carlo methods of statistical physics. This allows correlation functions, which can be related to QCD matrix elements, to be determined.

Progress in the field is slow. First our ‘box’ must be large enough to fit our correlation functions into. Then, often a chiral extrapolation must be made from a quark mass region around the strange quark mass to the light up and down quarks. Furthermore, the continuum (i.e. \( a \rightarrow 0 \)) limit must be taken. In addition to the above problems, to be able to compare with phenomenological or experimental results, matrix elements must be renormalized. In the statistical mechanics picture, we are approaching a second order phase transition, with all its attendant problems.

In the past, simply to save computer time, the fermion determinant in the action was discarded. This ‘quenched’ or ‘valence’ quark approximation is an uncontrolled approximation. Recently, however, simulations with two flavors of mass-degenerate sea quarks have begun appearing, allowing a first look at ‘real’ QCD and the effects of unquenching.

2 Generalities

Lepton–nucleon elastic scattering, \( lN \rightarrow lN \), in which a photon is exchanged between the lepton (usually an electron) and the nucleon (usually a proton), has been studied for many years. Indeed, there has been a resurgence of interest in these processes as part of the Jefferson Laboratory physics program. The scattering matrix element can be decomposed into a known electromagnetic piece and an unknown QCD matrix element:

\[
\langle \vec{p}', \vec{s}' | J_\mu(q) | \vec{p}, \vec{s} \rangle = \tau(p', s') \left[ \gamma_\mu F_1(Q^2) + i\sigma_{\mu\nu} \frac{q^\nu}{2m_N} F_2(Q^2) \right] u(p, s),
\]

where \( q = p' - p \) is the momentum transfer and \( Q^2 = -q^2 > 0 \). The values at \( Q^2 = 0 \) are \( F_1^p(0) = 1, F_2^p(0) = \mu_p - 1 \) for the proton and \( F_1^n(0) = 0, F_2^n(0) = \mu_n \) for the neutron, where \( \mu \) is the anomalous magnetic moment. Experimentally, it is more convenient to define the Sachs form factors

\[
G_e(Q^2) = F_1(Q^2) - \frac{Q^2}{(2m_N)^2} F_2(Q^2),
\]

\[
G_m(Q^2) = F_1(Q^2) + F_2(Q^2).
\]

Similarly, neutrino–nucleon scattering, for example \( \nu_e n \rightarrow \mu^- p \) mediated by a \( W^+ \) exchange, leads to an unknown axial current hadronic matrix element between neutron and proton states, which, with the use of current algebra and isospin invariance, may be re-written

\[
\langle \vec{p}', s' | A_\mu^{u-d}(q) | \vec{p}, s \rangle = \tau(p', s') \left[ \gamma_\mu \gamma_5 g_A(Q^2) + i\gamma_5 \frac{q_\mu}{2m_N} h_A(Q^2) \right] u(p, s),
\]
where \( A_{\mu}^{u-d} = \pi \gamma_{\mu} \gamma_5 u - \bar{d} \gamma_{\mu} \gamma_5 d \). From the \( \beta \) decay we know \( g_A \equiv g_A(0) = 1.267(4) \).

At higher momentum transfer the nucleon is broken up by the photon (or \( W^\pm \)) probe, and we enter the regime of deep-inelastic scattering (DIS) experiments, \( eN \to eX \) (or \( \nu_\mu n \to \mu^- X \)). The operator product expansion (OPE) leads to relations between moments of the structure functions and nucleon matrix elements of certain operators. For example

\[
\int_0^1 dx x^{n-2} F_2(x, Q^2) = \frac{1}{3} c_n (\mu^2/Q^2) v_n(\mu) + O\left(\frac{1}{Q^2}\right). \tag{4}
\]

Here \( x \) is the Bjorken variable, \( c_n \) are the Wilson coefficients and \( v_n \propto \langle N|\mathcal{O}_n|N\rangle \), where \( \mathcal{O}_n = (i/2)^{n-1} \bar{q} \gamma_\mu D_\mu \ldots D_{\mu_n} q \) are operators bilinear in the quarks, each containing \( n-1 \) covariant derivatives. In parton model language

\[
v_n = \langle x^{n-1} \rangle. \tag{5}\]

All matrix elements can be determined non-perturbatively using lattice QCD, while the Wilson coefficients, which contain the short-distance physics, can be computed in (continuum) perturbation theory. From the moments one can then reconstruct the structure functions.

### 3 Lattice Technicalities

In the last few years lattice QCD has improved in several respects, making it a quantitative tool of analysis.

Cut-off effects can be reduced (from \( O(a) \)) to \( O(a^2) \) by adding irrelevant operators to the Wilson fermion action\(^2\), \( S_F \), and to the operators\(^3\) whose matrix elements one wants to compute:

\[
S_F \to S_F - \frac{a}{4} c_{SW} g \sum_x \bar{\psi}(x) \sigma_{\mu\nu} F_{\mu\nu}(x) \psi(x),
\]

\[
\mathcal{O} \to (1 + c_0 a m) \mathcal{O} + a \sum_{i \geq 1} c_i \mathcal{O}_i, \tag{6}
\]

with \( c_{SW}, c_0, c_1, \ldots \) to be determined with non-perturbative precision\(^4\). This greatly facilitates the extrapolation of the results to the continuum limit.

All calculations which I am going to present have been done with improved fermions (6) and for two flavors of dynamical quarks. The gauge field configurations have been generated in collaboration with UKQCD. Details of our present data sample can be found elsewhere\(^5\).

The lattice operators (and matrix elements) are in general divergent and need to be renormalized, like the Wilson coefficients:

\[
\mathcal{O}(\mu) = Z(\mu, a) \mathcal{O}(a). \tag{7}
\]

There are several possibilities. The axial and vector renormalization constants may be determined by demanding that their (continuum) Ward identities are obeyed, which produces non-perturbative renormalization constants. Perturbation theory can be applied, but due to technical problems only one loop results are known. Even with this restriction, perturbation theory can be improved leading to ‘tadpole improved’ (TI) perturbation theory. However,
it still suffers from systematic errors. The QCDSF Collaboration has developed techniques allowing the renormalization constants to be determined non-perturbatively\(^5\). In Fig. 1 I compare the non-perturbative and TI perturbative renormalization constant for the vector current, which enters in the calculation of the form factors. We see that at our present couplings (\(\beta = 6/g^2 = 5.2 - 5.29\)) TI perturbation theory would lead to a systematic error of \(O(10\%)\).

At present, for the axial\(^7\) and vector currents (see above) most renormalization constants and improvement coefficients are known non-perturbatively. For \(v_n\) we rely on TI perturbation theory\(^3\).

4 Selected Results

We started only a year ago to look at the structure of the nucleon in the presence of dynamical quarks. Below I present some selected results.

The Axial Charge

The axial charge \(g_A\) of the nucleon is known very precisely experimentally. Hence it is a benchmark calculation of lattice QCD.

In Fig. 2 I show our results extrapolated to the continuum (\(a = 0\)) and chiral (\(m_\pi = 0\)) limits, using the fit formula

\[
g_A = A + B(m_\pi r_0)^2 + C(a/r_0)^2,
\]

where \(r_0 = 0.5\text{ fm} = 1/(395\text{ MeV})\) is the force parameter\(^8\), and a correction term is introduced to account for possible \(O(a^2)\) scaling violations, which are not removed by
Figure 2. The chiral and continuum extrapolations of $g_A$ for two flavors of dynamical quarks as a function of $m_\pi$ and $a$, compared to the experimental value (•) and the heavy-quark limit (→).

our improvement program. The agreement with the experimental value is quite good. But it should be noted that we are still far away from the continuum and chiral limits. For comparison, in our latest quenched calculation the smallest lattice spacing was $(a/r_0)^2 \approx 0.01$. To go to similarly small lattice spacings in the dynamical case will require computers with a sustained speed of $O(10)$ Teraflop/s.

Form Factors

Nucleon form factors have been extensively studied, both experimentally and theoretically, for many years. They describe the overall distribution of electric, magnetic and axial charge
Figure 3. The electric and magnetic form factors of the proton, $G_e$ and $G_m$ (solid symbols), as a function of $Q^2$ together with the experimental numbers (open symbols).

Figure 4. The axial form factor $g_A(Q^2)$ (solid symbols) as a function of $Q^2$ together with the experimental numbers (open symbols).

in the nucleon. An interesting aspect of lattice simulations is that one can change the value of the quark (pion) mass and study how the shape of the nucleon changes as one approaches the physical quark mass.

In Fig. 3 I show the electric and magnetic form factors of the proton for one particular coupling ($\beta = 5.25$) and a rather heavy pion mass, $m_\pi \approx 750$ MeV. And in Fig. 4 I show the axial form factor. I compare the results with the experimental numbers. We see that the form factors of the lattice nucleon are shallower than their experimental counterparts, which indicates that the lattice nucleon is somewhat smaller than the physical one. Indeed, a fit of the radius gives $r_{\text{rms}} \approx 0.70$ fm, which is to be compared with the phenomenolog-
Figure 5. The chiral and continuum extrapolations of the fraction of the proton’s momentum, $⟨x⟩$, carried by the up quarks minus that of the down quarks, for two flavors of dynamical quarks as a function of $m_π$ and $a$, compared to the experimental value (*) and the heavy-quark limit (→).

The value $r_{\text{rms}} = 0.83$ fm. This comes at no surprise. Quarks are tightly bound to each other, and they can leave the core of the nucleon practically only if they are bound to an anti-quark in form of a (light) pion. It happens then that pions are constantly emitted and reabsorbed, forming a pion cloud. And the smaller the pion mass is, the more extended is the pion cloud, and the larger is the charge radius of the nucleon.

At the moment our smallest pion mass is limited to $\gtrsim 500$ MeV. The reason is that the computational cost grows like $\sim 1/m_π^2$. The extrapolation of observables, which require an accurate view of the pion cloud, to the physical pion mass is difficult. It requires input from a low-energy effective theory of the nucleon, including pions and at least the $\Delta(1232)$...
resonance. Furthermore, the pion mass must not be too large. For small $Q^2$ first analytic results on the pion mass dependence of the nucleon form factors are available. We are currently attempting a chiral extrapolation incorporating these results.

### Moments of the Nucleon Structure Function

A wealth of information on the structure of the nucleon is contained in the nucleon structure functions, unpolarized and polarized, and the quark and gluon distribution functions derived from them. These distribution functions give the probability of finding a quark or gluon with a certain momentum or spin in the nucleon. Lattice calculations of the lowest moment, $\langle x \rangle$, of the unpolarized quark distribution functions tells us what fraction of the nucleon’s mass is carried by the quarks. Unbound quarks would only account for a few percent of the mass of the nucleon. The rest of the mass is due to the binding of the quarks and the gluons.

In Fig. 5 I show our results for $\langle x \rangle$. To compare the lattice results with the experimental number, one must extrapolate the data from the lowest calculated quark mass to the physical value, as I have discussed before. A simple, linear extrapolation overestimates the experimental number by $\approx 30\%$, suggesting that important physics is being omitted. Recently, it has been shown that the nucleon’s pion cloud gives rise to non-analytic terms in the quark mass, which may result in a large deviation from linearity as the quark mass tends to zero. A fit to the lattice data of the form

$$\langle x \rangle = A \left( 1 - \frac{3g^2 + 1}{(4\pi f^2)} m^2 \ln \left( \frac{m^2}{m^2 + \Lambda^2} \right) \right) + B (m_r)^2 + C (a/r_0)^2,$$

which preserves the correct (chiral) behavior and fits the experimental value, is also shown in Fig. 5.

These results have significant implications. It appears that calculations of the nucleon structure functions require the pion cloud to be adequately represented on the lattice. Even though one need not calculate at the physical quark (pion) mass, the pion must be light enough that the parameters of the chiral expansion are well determined by the lattice calculations. This is only the case for very light pions less than about 300 MeV, as can be inferred from Fig. 5. Similar results are found for the higher moments.

### 5 Conclusions

To obtain quantitative results from lattice simulations, beyond hadron and quark masses, decay constants and perhaps $\Lambda^3$, is quite hard. But progress is steady. About five years ago, when we started to work on structure functions, our calculations were done with standard Wilson fermions on $16^3 \times 32$ lattices in the quenched approximation and using perturbative renormalization constants. Today we are doing simulations in full QCD with two flavors of dynamical quarks on $24^3 \times 48$ lattices, using a non-perturbatively improved fermion action, improved operators and non-perturbative renormalization constants.

The improvement program has paid off: discretization errors are found to be relatively small. But a continuum extrapolation is still indispensable, which requires to repeat the calculations at smaller lattice spacings.
To safely extrapolate to the chiral limit, we need to do simulations at smaller quark masses with \( m_\pi \lesssim 300 \text{ MeV} \), such as to include an accurate view of the pion cloud.

All this will require computers with a sustained speed of \( O(10) \) Teraflop/s. But it should bring reliable calculations of hadronic structure within the capability of the next generation of computers such as the APE\textit{next} machine.

Acknowledgements

The analysis of the dynamical gauge field configurations has been performed on the Cray T3E at NIC-Jülich, where we have used \( \approx 100.000 \) CPU-hours (through September 2001), while the gauge field configurations themselves have been generated on the Hitachi SR-8000 at Leibniz Rechenzentrum in München and on the Cray T3E at the University of Edinburgh. We acknowledge fruitful discussions and collaboration with the UKQCD Collaboration.

References

QCD Thermodynamics: The Numerical Study of Strongly Interacting Matter under Extreme Conditions

F. Karsch\textsuperscript{1}, E. Laermann\textsuperscript{1}, P. Petreczky\textsuperscript{1}, S. Stickan\textsuperscript{1}, and I. Wetzorke\textsuperscript{2}

\textsuperscript{1} Fakultät für Physik
Universität Bielefeld, 33615 Bielefeld, Germany

\textsuperscript{2} NIC/DESY Zeuthen,
Platanenallee 6, 15738 Zeuthen, Germany

The basic building blocks of the matter we get in touch with in our daily life are atoms; its central part, the nucleus, is made up of protons and neutrons. Stable nuclei contain up to a few hundred of them. This is the form of nuclear matter we have to our disposal in a laboratory. Nuclear matter in bulk exists in the universe at much higher densities for instance in the form of neutron stars. We also believe that extremely hot and dense forms of strongly interacting matter existed in the early universe shortly after the Big Bang. Its properties must have been qualitatively different, protons and neutrons as well as all the other hadrons known to us must have dissolved into their constituents - quarks and gluons. It commonly is expected that this change to the quark-gluon plasma state of strongly interacting matter happened through a phase transition. The theoretical framework to describe this phase transition, the dense nuclear matter and its thermodynamic properties is known to us; it is the theory of strong interactions – Quantum Chromodynamics. The complexity of this theory, however, prohibits a direct analytic analysis of most of the interesting phenomena related to complex thermodynamic processes, in particular the critical behaviour related to phase transitions. To analyze the thermodynamics of strongly interacting matter we need large scale numerical calculations. We will discuss here some of the results obtained from such calculations.

1 Introduction

The conjecture that the structure of strongly interacting hadronic matter will undergo a qualitative change at high temperature and/or densities has been around for a long time. Two basic properties of hadrons were essential for the development of these ideas. In high energy experiments it had been observed that strongly interacting particles produce a large number of new resonance particles. This mechanism, in fact, is so effective that the number of resonances in a given energy or mass interval rises exponentially (resonance production \(\Rightarrow\) Hagedorn’s bootstrap model\textsuperscript{1}). As the average energy of hadrons increases with temperatures copious production of new particles will take place in a hot hadron gas; a dense equilibrated system results from this as a mixture of different particle species distributed according to the exponentially rising mass spectrum. Moreover hadrons are known to be extended particles with a typical size of about \(1\text{ fm} \simeq 10^{-13}\) cm. At high temperature extended hadrons thus would start to “overlap” and loose their identity as independent particles (see Fig. 1). This has been formulated in terms of percolation models\textsuperscript{2} and led to the expectation that some form of new physics has to occur under the extreme conditions that are realized at high temperature and/or densities.

With the formulation of Quantumchromodynamics (QCD) as a theoretical framework for the strong interaction force among elementary particles it became clear that this “new physics” indeed meant a phase transition to a new phase of strongly interacting matter –
the quark-gluon plasma (QGP)\(^3\). QCD is an asymptotically free theory; the interaction between the elementary constituents, quarks and gluons, is known to vanish logarithmically with increasing temperature. It thus is expected that at least at very high temperatures the dominant degrees of freedom are the constituents of QCD – quarks and gluons – which behave almost like free particles. Bulk thermodynamics of the high temperature phase thus asymptotically should look like that of an ideal (free) gas of quarks and gluons. An interesting question then is to understand how our phenomenological picture of the transition which has been derived from bootstrap and percolation models is realized within the framework of QCD. The basis for answering such questions is given by the lattice discretized version of QCD – lattice QCD. This opens the possibility to study complex non-perturbative processes like phase transitions in dense matter with the help of numerical simulation techniques based on more or less standard Monte Carlo or molecular dynamics algorithms. Results from such calculations will be discussed in the next section.

A further important aspect that makes the study of QCD thermodynamics so attractive is that quantitative numerical studies can also be confronted with results obtained in relativistic heavy ion collision experiments. A great challenge, of course, is to directly verify that basic hadron properties indeed change qualitatively in a thermal environment as suggested by our discussion given above. In-medium modifications of hadron properties can be studied in numerical calculations on the lattice either through the analysis of static screening mechanisms that lead to modifications of the potential (binding) energy between quarks or directly through the study of thermal modifications of hadron masses and the width of resonances. These effects lead to experimentally observable modifications of particle production cross sections for heavy quark bound states and light mesons, respectively. We will discuss lattice calculations of these quantities in Section 3.
2 QCD Thermodynamics

Numerical studies of the thermodynamics of strongly interacting matter are performed on 4-dimensional space-time lattices. These calculations have shown that a transition to the quark gluon plasma does take place, although this is a true phase transition only in limiting cases, \textit{i.e.} in the limit where the masses of the quarks either vanish or are infinite. In nature the masses of two quark species (up and down quarks) are small but finite, whereas a third species (strange quark) has a mass which is of the order of the transition temperature itself. The latter thus is expected to play a marginal role in thermodynamic processes which take place in the vicinity of the transition temperature but should become as important as the lighter up and down quarks at higher temperatures. At low temperature only hadronic bound states can exist. The lightest particles which will dominate the thermodynamics are pions with a mass of 140 MeV. The pion mass vanishes in the limit of vanishing quark masses (chiral limit) while all other states remain massive with masses larger than 500 MeV. Their contribution to the thermodynamics is suppressed by the corresponding Boltzmann weight factor $O(\exp[-m_H/T])$. Approximating the pions by massless particles the expectation thus is that bulk thermodynamic observables like the pressure or energy density will change from values corresponding to those of a free pion gas to that of a free quark-gluon gas. As the number of degrees of freedom is much larger in the latter case this should lead to a drastic change in the normalized energy density $\epsilon/T^4$,

\begin{equation}
\frac{\epsilon}{T^4} = \begin{cases} 
(n_f^2 - 1) \frac{\pi^2}{60} & , \ T \to 0 \\
(16 + \frac{21}{2} n_f) \frac{\pi^2}{60} & , \ T \to \infty
\end{cases} \quad ,
\end{equation}

where $n_f$ denotes the number of light quark species (flavours). In fact, the transition between these limiting forms occurs in a rather narrow temperature interval as can be seen in Fig. 2. The numerical calculations performed with different number of light quark species indeed show that the energy density of the QGP rapidly approaches that of an (almost) ideal quark-gluon gas. This strongly suggests that our basic picture about the relevant degrees of freedom in this phase is correct and that hadrons should have dissolved during the transition to the QGP. Moreover, one finds that the transition temperature shifts to smaller values as the number of light degrees of freedom increases. This is particularly drastic, when one compares transition temperatures in the cases $n_f = 0$ and $n_f > 0$. In a purely gluonic world ($n_f = 0$) the transition temperature is found to be about 270 MeV\footnote{100 MeV $\equiv 1.1605 \times 10^{12}$ Kelvin} while it drops to 175 MeV for $n_f = 2$ and 155 MeV for $n_f = 3$.

Despite the significant change in the transition temperature it, however, turns out that the critical energy density is not at all that different. It changes a lot when expressed in units of $T_c^4$; however, this is compensated by a corresponding shift of $T_c$. For the case of two and three light quark flavours shown in Fig. 2 one finds,

\begin{equation}
\frac{\epsilon_c}{T_c^4} = (6 \pm 2) \quad .
\end{equation}

This amounts to an energy density $\epsilon_c \simeq (0.3 - 1.3)\text{GeV/fm}^3$ which is at most three times larger than the energy density in an ordinary nucleon. An exciting prediction of these calculations thus is that the transition takes place at an energy density which can be produced in relativistic heavy ion experiments which currently are performed at the SPS at
Figure 2. The energy density for QCD with two and three light quark species. The figure is based on data obtained in connection with the analysis of the flavour dependence of the pressure. The crosses give an estimate for the temperature dependence of the energy density for QCD with a realistic quark mass spectrum with nearly massless up and down quarks and a strange quark mass \( m_s \approx T_c \). It shows that the thermodynamics is dominantly that of 2 light quark flavours in the vicinity of \( T_c \) and approaches that of 3-flavour QCD at higher temperatures.

CERN (Geneva) and the RHIC at Brookhaven National Laboratory (New York). At RHIC as well as at the future Large Hadron Collider (LHC), which will be build at CERN, it is expected that energy densities corresponding to several times the transition temperature will be produced so that long lived plasma states will be created.

When decreasing the values of light quark masses numerical calculations become increasingly difficult. This generally prohibits to perform numerical calculations directly at the physical values of the light, nearly massless up and down quarks; numerical results are generally extrapolated to the physical values realized in nature. This difficulty, however, can also be turned into a virtue. In a numerical calculation we can analyze the dependence of observables, e.g. the transition temperature, on the quark mass. In particular, we can analyze how the transition temperature depends on the values of the lightest hadron mass, the pseudo-scalar pions \( m_{PS} \). While certain model calculations suggest a strong dependence of \( T_c \) on \( m_{PS} \) the general arguments on resonance production and the related resonance gas models would suggest that such a dependence is only minor and the crucial mechanism for a phase transition in fact arises from the exponential rise of heavy resonance states. A systematic analysis of the quark mass dependence of \( T_c \) is shown in Fig. 3. The line shown in this figure is a representative fit to the 3-flavour data, which gave

\[
T_c(m_{PS}) = T_c(0) + 0.04(1) m_{PS} \quad . \tag{3}
\]

It thus seems that the transition temperature does not react strongly to changes of the lightest hadron masses. Their contribution to the overall energy density, however, rapidly decreases with increasing mass. The light hadrons thus will not be able to induce any form of critical behaviour. The weak dependence of \( T_c \) on \( m_{PS} \) favours the interpretation that contributions of heavy resonance masses are equally important for the occurrence of the transition. In fact, this also can explain why the transition still sets in at quite low temper-
Figure 3. The transition temperature in 2 (filled squares) and 3 (circles) flavour QCD versus $m_{PS}$ using an improved staggered fermion action (p4-action). Also shown are results for 2-flavour QCD obtained with the standard staggered fermion action (open squares). The difference between the two data sets for 2-flavour QCD gives a feeling for the systematic errors still present in current estimates of the transition temperature. The dashed band indicates the uncertainty on $T_c$ in the pure gauge theory ($n_f=0$) which corresponds to the limit $m_{PS} \to \infty$. The straight line is the fit given in Eq. 3.

pletely even when all hadron masses, including the pseudo-scalars (pions), attain masses of the order of 1 GeV or more. Also the observation that the critical energy density is only weakly dependent on the quark mass suggests that properties of the light hadron sector in general are not responsible for the transition to the plasma phase. For the quark masses currently used in lattice calculations a resonance gas model combined with a percolation criterion thus provides an appropriate description of the thermodynamics close to $T_c$. From a theoretical point of view it would, of course, still be interesting to analyze whether the role of the light meson sector becomes more dominant when calculations closer to the massless limit are performed, i.e. in a regime where the mass of the lightest hadronic state would be significantly smaller than the relevant temperatures.

3 Thermal Properties of Hadrons from Lattice QCD

Basic properties of the experimentally known zero temperature hadron spectrum are controlled by symmetries of QCD or, more precisely, by the fact that some of them are spontaneously broken in the vacuum. The spontaneous breaking of symmetries leads to non-vanishing condensates and lifts the degeneracy of mass eigenstates which differ only by a quantum number related to this condensate. The condensates, however, will be temperature dependent and eventually they will disappear in the high temperature phase. As the splitting of mass eigenstates is related to the strength of the condensates, it is natural to expect that hadron properties will change with temperature.

Lattice calculations of hadronic screening lengths and hadronic susceptibilities, which are based on the analysis of hadronic correlation functions did indeed provide evidence for significant changes of hadron properties in a thermal medium. These observables, however, provide only indirect information on thermal modifications of hadron masses
or their widths. If we want to get direct access to thermal changes of these basic hadron properties we have to understand the propagation of hadrons in a thermal medium and their interaction with this medium. In fact, all the information we are interested in is contained in the energy and momentum dependence of the retarded hadron propagator $\tilde{G}_R^H(\omega, \vec{p})$ in momentum space. Its imaginary part defines the spectral function, which in particular provides the information on hadron masses and their width,

$$\sigma_H(\omega, \vec{p}, T) = \frac{1}{\pi} \text{Im} \tilde{G}_R^H(\omega, \vec{p}, T) = \frac{\text{Im}\Sigma(T, \vec{p})}{\left(\omega^2 - \vec{p}^2 - \text{Re}\Sigma(T, \vec{p})\right)^2 + \left(\text{Im}\Sigma(T, \vec{p})\right)^2}. \quad (4)$$

Here $\Sigma(T, \vec{p})$ is the self-energy of the hadron. Its imaginary part will receive thermal in-medium contributions from scattering processes, which in general will lead to a broadening of the width of hadronic states (collision broadening).

In a lattice calculation we can analyze the Euclidean time correlation functions $G_H(\tau, \vec{p})$ which through a dispersion relation also depend on the spectral functions $\sigma_H$,

$$G_H(\tau, \vec{p}) = T \sum_n e^{-i\omega_n \tau} \tilde{G}_R^H(i\omega_n, \vec{p}) = \int_0^\infty d\omega \sigma_H(\omega, \vec{p}, T) \frac{\cosh(\omega(\tau - 1/2T))}{\sinh(\omega/2T)} \quad . \quad (5)$$

It is this correlation function which is accessible to numerical lattice calculations. The main problem thus in general is to reconstruct from the correlation functions the spectral functions from which we can get information on hadron masses and their width. However, even before doing so we can learn a lot from the correlation functions themselves. Comparing correlation functions in different quantum number channels, we directly see the drastic changes that occur when going from the low to the high temperature phase of QCD. Moreover, we see that correlation functions which are clearly different at low temperature become degenerate in the high temperature phase. This is, for instance, the case for correlation functions in the pseudo-scalar and scalar quantum number channels shown in Fig. 4. Here the splitting of the degeneracy at low temperature is due to the spontaneous breaking of chiral symmetries which get restored in the plasma phase.

Similar results as those obtained for the scalar and pseudo-scalar correlation functions can be obtained for vector and pseudo-vector correlation functions. The vector correlation function is of particular interest as it carries the quantum numbers of a photon, which couples to hadrons as well as leptons. The latter are accessible in heavy ion experiments. Once they have been formed they can leave the interaction region without any further strong interaction and thus carry important information about the structure of matter formed in an early stage in these collisions. For this reason, the analysis of dilepton pair production is expected to give important information on thermal effects in heavy ion collisions. The cross section for dilepton (e.g. electron-positron pair) production is directly related to the vector spectral function. For pairs with vanishing total momentum this is given by,

$$\frac{dW}{d\omega d^3p} = \frac{5\alpha^2}{27\pi^2} \frac{1}{\omega^2(\omega/T - 1)} \sigma_V(\omega, T) \quad , \quad (6)$$

\footnote{The scalar correlation function shown in Fig. 4 is the so-called connected part of the correlation function.}
with \( \sigma_V(\omega, T) \equiv \sigma_V(\omega, \vec{p} = 0, T) \). The steps needed to get from the numerical calculation of the vector correlation functions to the experimentally accessible cross section is illustrated in Fig. 5. The crucial part is the reconstruction of the spectral function from the correlation function. The latter is known only at a discrete set of Euclidean times, \( \tau T = k/N_\tau, \ k = 0, 1, ..., N_\tau - 1 \). The inversion of the integral equation Eq. 5 thus is an ill-posed problem. However, using statistical tools like the maximum entropy method (MEM)\(^{11}\) it is, nonetheless, possible to determine the most likely form of the spectral function which is consistent with any given prior knowledge on its structure. In the case of QCD this in general means that we have information about the behaviour of \( \sigma_H(\omega, T) \) for large \( \omega \), i.e. we can control the short distance structure of the correlation functions in perturbation theory. This defines an initial guess for the spectral function; the default model \( m(\omega, T) \). The spectral function is then determined by maximizing the function \( Q \equiv \gamma S - L \) with respect to \( \sigma \) and the additional parameter \( \gamma > 0 \). Here \( S \) is the Shannon-Jaynes entropy

\[
S(\sigma) = \int_0^\infty d\omega \left[ \sigma(\omega, T) - m(\omega, T) - \sigma(\omega, T) \ln(\sigma(\omega, T)/m(\omega, T)) \right]
\]

and \( L \) is the \( \chi^2 \) constructed from the data sample for the correlation function at a discrete set of Euclidean times, \( \{D_i(k) \equiv G_H(k/N_\tau)|k = 0, 1, ..., N_\tau - 1; i = 1, ..., \#\text{conf.}\} \), and the fitting function \( G_H^{\text{fit}}(k/N_\tau) \) constructed from Eq. 5 with the current trial version for \( \sigma(\omega, T) \),

\[
L(\sigma) = \frac{1}{2} \sum_{k,l} (G_H^{\text{fit}}(k/N_\tau) - D(k))C^{-1}_{kl}(G_H^{\text{fit}}(l/N_\tau) - D(l))
\]

with \( D(k) = \frac{1}{\#\text{conf.}} \sum_i D_i(k) \) and \( C_{kl} \) denoting the covariance matrix of the data sample. First tests of this approach at zero temperature\(^{12}\) and in the infinite temperature limit\(^8\) indicated that a MEM analysis of hadron correlation functions is feasible already on lattices with moderate temporal extent.
Lattice Calculation

temporal correlation function of hadronic vector current:

\[ G_V(\tau T) = \int_0^\infty d\omega \sigma_V(\omega, T) \frac{\text{ch}(\omega(\tau - 1/2T))}{\text{sh}(\omega/2T)} \]

Statistical Analysis

reconstructed spectral function based on Maximum Entropy Method

\[ \max_{\gamma, \sigma_V} [\exp(\gamma S - L)] \]

(see Eqs. 7 and 8)

Experimental Observable

predicted thermal contribution to the dilepton cross section:

\[ \frac{dW}{d\omega d^3p} = \frac{5\alpha^2}{27\pi^2 \omega^2(\omega/T - 1)^2} \sigma_V(\omega, T) \]

Figure 5. The vector correlation function \( G_V(\tau T) \) and the corresponding free correlation function for massless quark anti-quark pairs, \( G_{V,\text{free}}(\tau T) \) versus Euclidean time in units of the temperature (a), reconstructed vector spectral functions \( \sigma_V \) in units of \( \omega^2 \) at zero momentum (b) and the resulting zero momentum differential dilepton rate (c) at \( T/T_c = 1.5 \) and 3. Also shown with solid lines is the free spectral function in (b) and the resulting Born rate for thermal dilepton production in (c).
The vector spectral function reconstructed from lattice calculations on large lattices of size $64^3 \times 16$ at two values of the temperature is shown in Fig. 5b and the resulting dilepton rate is given in Fig. 5c. The most exciting aspect of this result is that the rate stays close to the Born rate calculated in leading order perturbation theory and that the rate is cut-off at low energies. In particular, the low energy behaviour is difficult to control in perturbation theory which suffers from infrared singularities. The lattice results, on the other hand, suggest that the low energy part of the spectral function is cut-off, which can be understood in terms of a threshold generated by non-vanishing thermal quark masses. This also is supported by studies of the quark propagators in fixed gauges and its spectral analysis.

4 Conclusions

Numerical studies of lattice regularized QCD allow a detailed quantitative analysis of the thermodynamics of strongly interacting matter. The basic techniques and tools used to analyze the QCD equation of state and the transition to the quark gluon plasma have been developed during the last 20 years. However, it is only now that computing facilities with Teraflops computing power are within reach and will allow us to reach the accuracy needed to make definite predictions that can be confronted with experimental data. The critical temperature and critical energy density of the QCD phase transition clearly will be the most fundamental observables which we are asked to provide. At present the critical temperature still is known to us only with a statistical and systematic error of about 10%. This alone leads to a 45% error in the prediction of the critical energy density. There is, however, no doubt that we will soon be able to improve on this.

Lattice calculations also did provide ample evidence for the modification of hadron properties in a thermal medium. However, it is only now that a direct investigation of thermal masses and their widths, which are of direct experimental interest, is within reach. To some extent this also is due to the improved computational resources and the fact that simulations on large thermal lattices at small values of the lattice cut-off now become possible. Equally important, however, it was to realize that statistical tools, which previously have been successfully applied in statistical physics, could also be of use in our studies of correlation functions in quantum field theory. We only have started to explore the possibilities the maximum entropy method offers for the analysis of static correlation functions. In the context of finite temperature field theory it seems to give access to a whole range of open questions that can be addressed in the future.

Acknowledgments

The work has been supported by the TMR network ERBFMRX-CT-970122 and by the DFG under grant FOR 339/1-2. Numerical studies of the QCD phase diagram and the equation of state have been performed on the APEmille at Bielefeld University. Calculations of thermal correlation functions and the hadronic spectral functions have been performed on the Cray T3E at the NIC-Jülich, typically using a partition of 128 processors. This ongoing project gets supported with 13.475 KE per month, i.e. about 3850 processor-hours per month.
References

4. for a recent review see, F. Karsch, Lattice QCD at high temperature and density, Lectures given at 40th Internationale Universitätswochen für Theoretische Physik: Dense Matter (IUKT 40), Schladming, Styria, Austria, 3-10 Mar 2001, hep-lat/0106019.
Numerical Simulation of High Quark Densities in QCD with two Colours

I. Montvay\textsuperscript{1}, S. Hands\textsuperscript{2}, L. Scorzato\textsuperscript{2}, and J. I. Skullerud\textsuperscript{1}

\textsuperscript{1} Deutsches Elektronen-Synchrotron DESY
Notkestrasse 85, 22603 Hamburg, Germany

\textsuperscript{2} Department of Physics, University of Wales Swansea
Singleton Park, Swansea SA2 8PP, U.K.

The DESY-Swansea Collaboration performed numerical simulations investigating SU(2) lattice gauge theory at non-zero chemical potential with one staggered quark flavour in the adjoint representation. This lattice model has similar features to QCD itself and its study gives interesting insights into some open problems of high density quark matter. In particular the role of the “sign problem” can be clarified in connection with diquark condensation and the phase diagram.

1 Introduction

The fundamental theory of strongly interacting matter is Quantum Chromo-Dynamics (QCD), which is a relativistic quantum field theory with quarks and gluons as quanta of the elementary constituent fields. Physical properties of strongly interacting matter, manifested in high energy heavy ion collisions, in the early Universe or in neutron stars, can, in principle, all be deduced from this theory. An important calculational method in QCD is numerical Monte Carlo simulation on space-time lattices.

In recent years, significant progress has been made in understanding the phase diagram of QCD at high quark- (or baryon-) number density. On the basis of model calculations\textsuperscript{1,2}, it is now believed that the ground state of QCD at high density and low temperatures is characterised by a diquark condensate which spontaneously breaks gauge and/or baryon number symmetries (for recent reviews see\textsuperscript{3}). This leads to phenomena such as colour superconductivity and/or colour-flavour locking which have substantial influence on the physics of high density hadronic (quark- ) matter and have a major impact, for instance, on the properties of heavy neutron star cores. However, although the results appear to be qualitatively independent of the specific model and approximation employed, little can be said quantitatively due to the lack of a first-principles, non-perturbative method that can access the relevant regions of the phase diagram.

Non-perturbative Monte Carlo simulations in lattice QCD, which are first-principles methods, fail when a chemical potential term $\mu n_B$, where $n_B$ is the quark number density, is added to the Lagrangian. The reason is that the Euclidean-space fermion determinant becomes complex. As a consequence, standard simulation algorithms cannot be applied. However, it is possible to study QCD-like theories where the fermion determinant remains real even at non-zero $\mu$. These theories can be used as testbeds to examine the validity of the models used to study real QCD, as well as directly to improve our understanding of phenomena such as diquark condensation and phase transitions in dense matter. Examples of such theories with real quark determinant are two-colour (SU(2) colour) QCD and QCD with adjoint quarks.
The DESY-Swansea Collaboration studied two-colour QCD with staggered quarks in the adjoint representation\textsuperscript{4–6}. The use of a particular lattice fermion formulation (with staggered fermions) is important because for real or pseudoreal representations of the gauge group the pattern of symmetry breaking is expected to be different in the continuum and at non-zero lattice spacing\textsuperscript{7}. In this QCD-like model it is possible to investigate several interesting questions. Its basic features are:

- For an odd number of staggered flavours \( N \), the fermion determinant may be negative. This means that this model has a *sign problem*, which may make simulations at large \( \mu \) difficult. At the same time a sign is simpler than the continuum of phases in QCD. It may thus be feasible to make progress using standard means, or at least expose physical distinctions between the positive and negative sectors.

- For zero quark mass (\( m \)) and vanishing chemical potential \( m = \mu = 0 \) the \( U(N) \times U(N) \) flavour symmetry is enhanced to a \( U(2N) \) symmetry which relates quarks to antiquarks. This symmetry is broken by the chiral condensate, leaving \( N(2N - 1) \) massless Goldstone modes, which become degenerate pseudo-Goldstone states for \( m \neq 0 \).

- When \( N \geq 2 \) these pseudo-Goldstone states include gauge invariant scalar diquarks which, though degenerate with the pion at \( \mu = 0 \), carry baryon number. These models with \( N \geq 2 \) can be studied for \( \mu 
eq 0 \) by chiral perturbation theory (\( \chi PT \))\textsuperscript{8}. The main result is that for \( \mu > m_\pi/2 \), where \( m_\pi \) is the pion mass, the chiral condensate rotates into a diquark condensate (the two being related by a \( U(2N) \) rotation) while the baryon density increases from zero.

- The model with \( N = 1 \) is not expected to contain any diquark pseudo-Goldstones and is not accessible to \( \chi PT \). One expects an onset transition as some \( \mu_0 \approx m_b/n_q > m_\pi/2 \), where \( m_b \) is the mass of the lightest baryon and \( n_q \) its baryon charge.

- For \( N = 1 \) the gauge invariant scalar diquark operator is forbidden by the Pauli Exclusion Principle; there is, however, a possibility of a gauge non-singlet, and hence colour superconducting, diquark condensate at large chemical potential.

The numerical simulation of lattice gauge theories is performed on a hypercubic lattice in four dimensional Euclidean space. The fourth coordinate, besides the three space coordinates, is imaginary time. In the path integral formulation the Euclidean action is needed which gives the weight of lattice field configurations. For the gauge field part the standard Wilson action, weighted by the inverse gauge coupling \( \beta \), is taken\textsuperscript{9}. The fermionic part of the lattice action, with quark mass \( m \) and chemical potential of quark charge \( \mu \), can be defined as follows:

\[
S = \sum_{x,y} \bar{\chi}^p(x) D_{x,y}[U,\mu] \chi^p(y) + m \bar{\chi}^p(x) \delta_{x,y} \chi^p(y) \\
\equiv \sum_{x,y} \bar{\chi}^p(x) M_{x,y}[U,\mu] \chi^p(y),
\]

(1)
where the index $p$ runs over $N$ flavours of staggered quark, and $D$ is given by

$$D_{x,y} = \frac{1}{2} \sum_{\nu \neq 0} \eta_{\nu}(x) \left( U_{\nu}(x) \delta_{x,y-\nu} - U_{\nu}^\dagger(y) \delta_{x,y+\nu} \right)$$

$$+ \frac{1}{2} \eta_{0}(x) \left( e^{\mu} U_{0}(x) \delta_{x,y-0} - e^{-\mu} U_{0}^\dagger(y) \delta_{x,y+0} \right). \tag{2}$$

The $\chi, \bar{\chi}$ are single spin component Grassmann objects, and the phases $\eta_{\mu}(x)$ are defined to be $(-1)^{x_{0}+\cdots+x_{\mu}-1}$. The link matrices in the adjoint representation $U_{\mu}$ are real $3 \times 3$ orthogonal matrices. The Grassmann variables in the path integral can be integrated out, resulting in the determinant of the fermion matrix $M$ appearing in (1). This fermion determinant takes into account the effects of virtual fermion-antifermion pairs on the gauge field.

The main difficulty in numerical simulations of gauge theories with fermions is to include the fermion determinant. We have used two different simulation algorithms, the hybrid Monte Carlo (HMC) algorithm\textsuperscript{10} which is the standard choice for QCD, and a Two-Step Multi-Bosonic (TSMB) algorithm\textsuperscript{11} which has been developed recently in connection with a study of supersymmetric Yang-Mills theory.

![Figure 1. Typical eigenvalue spectrum of the fermion matrix on a $4^3 \cdot 8$ lattice at ($\beta = 2.0$, $m = 0.1$, $\mu = 0.35$).](image-url)
2 The Sign of the Determinant

At zero chemical potential $\mu = 0$ the complex eigenvalues of the fermion matrix $M$ in (1) lie on a line parallel to the imaginary axis with real part equal to the quark mass $\text{Re}\lambda = m$. Since the eigenvalues come in complex conjugate pairs, the determinant of the fermion matrix at $\mu = 0, m > 0$ is positive. After introducing a non-zero chemical potential the eigenvalues spread out into an elliptical region of the complex plane (see, for instance, figure 1).

At large enough chemical potential the region occupied by the eigenvalues of the fermion matrix reaches the origin of the complex plane (this is actually the situation in figure 1), which has important consequences for the behaviour of the system. From the technical point of view this is manifested by the appearance of very small eigenvalues of the squared fermion matrix (see figure 2), which makes the numerical simulation difficult. The very small eigenvalues occur because there are sign changes of the fermion determinant. It turns out that this can only be seen in the TSMB simulation because the HMC algorithm is based on small steps in configuration space and cannot in practice change the determinant sign. This is an important observation showing that under these circumstances the standard HMC algorithm is not applicable because it is not ergodic. As we shall see below, the HMC simulation is very similar to the restriction of the TSMB simulation to the positive determinant sector. The omission of the gauge configurations with negative determinants has important consequences for the physical results.

In simulations with the TSMB algorithm the qualitative change for increasing chemical potential can also be clearly seen in the distribution of the reweighting factors (see figure 3). These are required because the polynomial approximations applied in this algorithm are not exact near zero eigenvalues. The reweighting factors, which also include the sign

Figure 2. The lowest eigenvalues of the hermitean fermion matrix, for $\mu = 0.3$ (left) and $\mu = 0.4$ (right). The lines indicate the lower bound $\epsilon$ of the polynomial approximation employed.
of the fermion determinant, are correcting for this. As it can be seen in figure 3, the reweighting factors are strongly peaked near 1 for a $\mu$ value where the spectrum does not yet touch zero ($\mu = 0.3$). At $\mu = 0.37$, where the spectrum just reaches the origin, some configurations with negative reweighting factor (i.e. negative determinant) start to appear. Finally, at $\mu = 0.4$ the frequency of both signs is already almost equal and the sign problem is serious.

3 Simulation Results

Having a suitable simulation algorithm (TSMB) which samples both positive and negative determinant sectors properly one can compare results with and without taking into account the determinant sign. Most of the results of the DESY-Swansea Collaboration have been obtained at $\beta = 2.0$, $m = 0.1$ on a $4^3 \times 8$ lattice. It turned out that the results in the positive sector of the TSMB simulations are close to the HMC simulation results and both are qualitatively similar to the results of chiral perturbation theory for theories with diquark Goldstone modes (for instance, $N = 2$ flavours). This is not surprising because taking the absolute value of the fermion determinant in the path integral measure is equivalent to consider $\sqrt{\det(M^\dagger M)}$ instead of $\det(M)$. An example of the good agreement between the results of HMC simulations and chiral perturbation theory predictions is shown by figure 4 in case of the chiral condensate $\langle \bar{\psi}\psi \rangle$.

The inclusion of the fermion determinant sign, in the single staggered flavour ($N = 1$) model we are interested in, changes the physical results qualitatively. The main effect is that the phase transition at chemical potential $\mu \simeq \frac{1}{2} m_N$, which is signalled for instance by the sharp drop of the chiral condensate in figure 4, disappears due to a cancellation between the contributions of positive and negative sectors. As an example one can consider the behaviour of the quark density shown in figure 5. The increase beyond $\mu = 0.35$ disappears once the signs are taken into account properly. Unfortunately, the severe sign problem does not allow us to go much beyond $\mu = 0.4$ where the onset phase transition of the $N = 1$ model is expected.

A natural explanation of the agreement between the HMC and positive determinant results and $\chi$PT predictions is that the positive determinant sector mimics a theory with two flavours: one quark and one ‘conjugate quark’\textsuperscript{13}. We may cast further light on this by
studying the behaviour of the diquark condensate of the \( N = 2 \) diquark operator

\[
qq_3 = \frac{i}{2} \left[ \chi^{\mu \nu}(x)\varepsilon_{\mu\nu} \chi^q(x) + \bar{\chi}^p(x)\varepsilon_{pq} \bar{\chi}^{q \nu}(x) \right],
\]

(3)

where \( p, q = 1, 2 \) are explicit flavour indices and \( \varepsilon \) the antisymmetric tensor. A non-zero expectation value \( \langle qq_3 \rangle \neq 0 \) is expected in the two-flavour theory, violating the symmetry of the Lagrangian under global U(1) baryon number rotations. Physically this implies a superfluid ground state and (via Goldstone’s theorem) exactly massless scalar excitations. This accounts for the proliferation of small eigenvalues observed in section 2. The superfluid condensate may be determined by introducing a diquark source term in the action, which now describes two flavours,

\[
S[j] = S + j \sum_x qq_3(x).
\]

(4)

One has to extract the condensate \( \langle qq_3(j) \rangle = V^{-1} \partial_1 \ln Z[j] / \partial j \), and extrapolate the results to \( j = 0^{14} \). The HMC simulations do support \( \langle qq_3 \rangle \neq 0 \) for \( \mu \geq m_\pi \), as expected. In the TSMB simulations taking into account the determinant sign the diquark condensate \( \langle qq_3 \rangle \) is suppressed.

Figure 4. Rescaled chiral condensate vs. chemical potential obtained in HMC simulations together with the \( \chi PT \) predictions.
Figure 5. The fermion density from TSMB and HMC simulations. In case of TSM simulations the result in the positive determinant sector is shown separately (open symbols).

4 Conclusions

We have studied two-colour QCD with one flavour of staggered quark in the adjoint representation. We have employed two different simulation algorithms, and have gained insight into the optimal tuning of the TSMB algorithm in the high density regime. This is the preferred algorithm not only because it is capable of maintaining ergodicity via its ability to change the determinant sign, but also because it more effectively samples small eigenmodes, which are important in the presence of a physical Goldstone excitation. We find that the positive determinant sector behaves like a two-flavour model, and exhibits good agreement with chiral perturbation theory predictions for such a model. At higher chemical potentials there are some indications of a breakdown of \( \chi \)PT, and possible signs of a further phase transition. However, data from larger volumes and smaller bare quark masses would be needed to make these observations definitive.

Above the onset transition in the positive determinant sector, we have successfully obtained a signal for a non-zero two-flavour diquark condensate \( \langle \bar{q}q_3 \rangle \), indicating a superfluid ground state for \( \mu > m_\pi/2 \). The chiral condensate rotates into this diquark condensate, in good quantitative agreement with the behaviour predicted by \( \chi \)PT. This feature also enabled us to achieve reasonable control over the necessary \( j \to 0 \) extrapolation\(^6\).

When the negative determinant configurations are included in the measurement, the onset transition and diquark condensation disappear. This is what we would expect for the one-flavour model and is consistent with the Pauli Exclusion Principle. There is a
strong evidence for this scenario, providing a conclusive demonstration, should any still be
needed, that the determinant sign plays a decisive role in determining the ground state of
systems with $\mu \neq 0$. Unfortunately, the severity of the sign problem means we have not
been able to locate the real onset transition for this model.

The numerical calculations presented in this report have been performed on the Cray
T3E of the NIC, Jülich since 1999. The total CPU time was about 180,000 hours.

References

1. R. Rapp, T. Schäfer, E.V. Shuryak, M. Velkovsky, *Diquark Bose condensates in high
density matter and instantons*, Phys. Rev. Lett. **81**, 53 (1998); *High density QCD and

2. M. Alford, K. Rajagopal, F. Wilczek, *QCD at finite baryon density: nucleon droplets
and color superconductivity*, Phys. Lett. **B422**, 247 (1998); *Color flavor locking and

3. K. Rajagopal, F. Wilczek, *The condensed matter physics of QCD*, in *At the Frontier
of Particle Physics: Handbook of QCD*, edited by M. Shifman, p. 2061, Singapore,


6. S. Hands, I. Montvay, L. Scorzato, J. Skullerud, *Diquark condensation in dense ad-

7. S.J. Hands, J.B. Kogut, M.-P. Lombardo and S.E. Morrison, *Symmetries and spectrum
of SU(2) lattice gauge theory at finite chemical potential*, Nucl. Phys. **B558**, 327
(1999).

8. J.B. Kogut, M.A. Stephanov, D. Toublan, J.J.M. Verbaarschot, A. Zhitnitsky, *QCD-

1994.


Phys. Commun. **109**, 144 (1998); in *Numerical challenges in lattice quantum chro-

13. M. Stephanov, *Random matrix model of QCD at finite density and the nature of the

14. S. Morrison and S. Hands, *Two colours QCD at nonzero chemical potential*, in *Strong
From Enemies to Friends: Chiral Symmetry on the Lattice

Pilar Hernández, Karl Jansen, and Laurent Lellouch

1 CERN, Theory Division
1211 Geneva 23, Switzerland
E-mail: Pilar.Hernandez@cern.ch

2 NIC/DESY Zeuthen
Platanenallee 6, 15738 Zeuthen, Germany
E-mail: Karl.Jansen@desy.de

3 Centre de Physique Théorique, Case 907, CNRS Luminy
13288 Marseille Cedex 9, France
E-mail: lellouch@cpt.univ-mrs.fr

The physics of strong interactions is invariant under the exchange of left-handed and right-handed quarks, at least in the massless limit. This invariance is reflected in the chiral symmetry of quantum chromodynamics. Surprisingly, it has become clear only recently how to implement this important symmetry in lattice formulations of quantum field theories. We will discuss realizations of exact lattice chiral symmetry and give an example of the computation of a physical observable in quantum chromodynamics where chiral symmetry is important. This calculation is performed by relying on finite size scaling methods as predicted by chiral perturbation theory.

1 Introduction

Nature as well as physicists like symmetries. The somewhat amusing reason for this is that symmetries can be broken. A very important concept is the spontaneous breakdown of a symmetry: here some symmetry is broken down to a situation with less symmetry when tuning a parameter (e.g. the temperature or the coupling strength) of the theory to a critical value. In the process a non-vanishing vacuum expectation value is developed that breaks the symmetry of the interaction. The process is accompanied by the appearance of so-called Goldstone particles (spinwaves) that are massless. This phenomenon comes under the name of the Goldstone theorem. An example is the spontaneous magnetization: a metal at high temperature is in a symmetric state – the elementary magnets or spins can point in any direction such that the net magnetization vanishes. Decreasing the temperature below some critical value, a spontaneous magnetization occurs, the spins point into a preferred direction and the metal becomes magnetic.

Within the standard model of elementary particle interactions we know two places where such a spontaneous symmetry breaking is supposed to have occurred: in the electroweak sector spontaneous symmetry breaking manifests itself in the Higgs phenomenon with the development of a Higgs field expectation value \( \langle \Phi \rangle \), giving mass to the elementary particles, and the appearance of Goldstone particles leading to the W- and Z-bosons.

In our theory of strong interactions, quantum chromodynamics (QCD), it is a chiral symmetry that is assumed to be spontaneously broken. This symmetry allows for an interchange of left-handed and right-handed quarks while leaving physics invariant – at least
when these quarks are massless. In this case a scalar quark-antiquark \( q\bar{q} \) condensate is developed and the Goldstone particles are identified with the light pions that are observed in nature.

As stated above the occurrence of spontaneous symmetry breaking is an assumption. The phenomenon is inherently non-perturbative and cannot be addressed with approximative methods like perturbation theory. However, even with numerical simulations it is difficult to test, whether a certain model exhibits spontaneous symmetry breaking (SSB). The reason for this becomes clear when the way to detect SSB is considered. Let us choose a system that has a finite physical volume \( V \) as would be required for numerical simulations. Further, we couple the system to an external magnetic field. Spontaneous symmetry breaking is tested in a double limit, where first the volume of the system is sent to infinity and then the external magnetic field is sent to zero. If a non-vanishing magnetization remains, spontaneous symmetry breaking is identified. Obviously, such a procedure is unfeasible within the approach of numerical simulations.

The way out is the use of chiral perturbation theory\(^1\). In this approach chiral symmetry breaking is taken as an assumption with the consequences of the appearance of non-vanishing field expectation values and Goldstone particles. A special situation arises when the size of the box becomes comparable to or even smaller than the Compton-wavelength of the Goldstone particle. Then the corresponding field can be considered as being uniform and it is possible to set up a systematic expansion that starts in the lowest order with an effective lagrangian of this constant mode and then taking systematically higher order fluctuations into account\(^2\).

\section{Example of \( \Phi^4 \)-Theory}

Let us give an example of the Ginsburg-Landau or \( \Phi^4 \) theory with \( O(N) \)-symmetry in four dimensions. The action of this theory is defined by

\[ S = \int d^4x \left( \frac{1}{2} \left( \partial_\mu \Phi(x) \right)^2 + \frac{1}{2} m_0^2 \Phi(x)^2 + \lambda_0 \Phi(x)^4 + j^\alpha \Phi^\alpha(x) \right) \]  

with \( \Phi \) a \( N \)-component vector, \( m_0 \) the bare mass, \( \lambda_0 \) the bare quartic coupling and \( j^\alpha \) a constant external source in direction \( \alpha \) of the group \( O(N) \).

Expectation values of observables are computed through the partition function or path integral in finite volume \( V = L^4 \),

\[ \langle \Phi \rangle_{j,V} = \int \mathcal{D}\Phi e^{-S} \]  

Symmetry breaking is detected via a spontaneous magnetization \( \langle \Phi \rangle \neq 0 \) in the double limit

\[ \langle \Phi \rangle = \lim_{j \to 0} \lim_{V \to \infty} \langle \Phi \rangle_{j,V} \].
Chiral perturbation theory can provide now a prediction for the behaviour of $\langle \Phi \rangle_{j,V}$, i.e. the expectation value in finite volume and at non vanishing external field. In a situation when the external field $j$ and $1/V$ are very small, $\langle \Phi \rangle_{j,V}$ is given in terms of the expectation value of $\langle \Phi \rangle$ at $j = 0$ and $V = \infty$:

$$\langle \Phi \rangle_{j,V} = f(u)$$

where $u$ is a scaling variable

$$u = \langle \Phi \rangle j V$$

that contains the infinite volume and $j = 0$ spontaneous magnetization $\langle \Phi \rangle$.

Let us give for completeness the function $f(u)$:

$$f(u) = \frac{u^2 \eta(u)}{j V} \quad \text{with} \quad \eta(u) = \frac{1}{u} \frac{I_2(u)}{I_1(u)}$$

with $I_1$, $I_2$ modified Bessel functions. The important point here is that these functions only depend on a single scaling variable $u$ and hence on the quantity of interest, the magnetization $\langle \Phi \rangle$.

In an already quite old work\(^3\) the prediction of chiral perturbation theory, eq. (4), was confronted with numerical data. In this test the data obtained through numerical simulations were described very well by the theoretical formulae of chiral perturbation theory and the method of chiral perturbation theory turned out to be very fruitful, at least in this example of a relatively simple model. The lesson that is to be learnt from this example is that finite size effects can actually be used to determine properties of an infinitely large system. In this sense, the finite volume system can be regarded as a probe of the target theory in infinite volume. This gives an entirely new perspective on finite size effects: instead of being afraid of them, they can be used to determine important physical information.

### 3 The Example of Quantum Chromodynamics

Chiral perturbation theory as well as the lattice method was developed for quantum chromodynamics in order to deepen our understanding of the strong interactions. Still, until relatively recently, both approaches could not really come together, at least in the region of very small quark masses where chiral symmetry starts to get restored. The reason was that the lattice seemed to be lacking the concept of chiral symmetry and for many years the infamous Nielsen-Ninomiya theorem\(^4\) was telling us that it would even be impossible to implement chiral symmetry in a consistent way on a lattice.

The situation only changed a few years ago, when an old work by Ginsparg and Wilson\(^5\) was rediscovered\(^6\). The Ginsparg-Wilson paper contained actually a clue for answering the problem of chiral fermions on the lattice. The interaction of the fermions is described by some particular operator, the Dirac operator, the details of which should not be discussed here. Now, in the continuum theory this Dirac operator anticommutes with a certain 4-dimensional matrix $\gamma_5 = \text{diag}(1, 1, -1, -1)$. On a lattice with non-vanishing lattice spacing $a$, such an anti-commutation property cannot be demanded. If the anticommutation property is insisted on, the fermion spectrum of the lattice theory does not correspond to the one of the target continuum theory.
The suggestion of Ginsparg and Wilson was to replace the anti-commutation condition by a relation (now known as the Ginsparg-Wilson (GW) relation) for a lattice Dirac operator $D$:

$$\gamma_5 D + D\gamma_5 = aD\gamma_5 D.$$ \hfill (7)

Clearly, in the limit that the lattice spacing vanishes the usual anti-commutation relation of the continuum theory is recovered.

The fact that renders the relation eq. (7) conceptually extremely fruitful is that it implies an exact chiral symmetry on the lattice even if the value of the lattice spacing does not vanish. The notion of a chiral symmetry on the lattice is a conceptual breakthrough and renders the lattice theory in many respects to behave like its continuum counterpart with far reaching consequences.

However, as nice as the theoretical progress that followed the rediscovery of the Ginsparg-Wilson relation was as much of a challenge are realizations of operators $D$ that satisfy the Ginsparg-Wilson relation (see for reviews). Let us give a particular example for such a solution as found by H. Neuberger from the overlap formalism, based on the pioneering work of D. Kaplan. To this end, we first consider the standard Wilson Dirac operator on the lattice:

$$D_w = \frac{1}{2} \left\{ \gamma_\mu \left( \nabla_\mu^* + \nabla_\mu \right) - a\nabla_\mu^* \nabla_\mu \right\}$$ \hfill (8)

with $\nabla_\mu$, $\nabla_\mu^*$ the lattice forward, backward derivatives, i.e. nearest neighbour differences, acting on a field $\Phi(x)$

$$\nabla_\mu \Phi(x) = \Phi(x + \mu) - \Phi(x)$$

$$\nabla_\mu^* \Phi(x) = \Phi(x) - \Phi(x + \mu).$$

We then define

$$A = 1 + s - D_w$$ \hfill (9)

with $0 < s < 1$ a tunable parameter. Then Neuberger’s operator $D_N$ with mass $m$ is given by

$$D_N = \left\{ 1 - \frac{m}{2(1+s)}D_N^{(0)} + m \right\}$$ \hfill (10)

where

$$D_N^{(0)} = (1+s) \left[ 1 - A(A^\dagger A)^{-1/2} \right].$$ \hfill (11)

What is important here in the definition of Neuberger’s operator is the appearance of the square root of the operator $A^\dagger A$. This means that $D_N$ connects all points of the lattice with each other. Note, however, that despite this the operator is a local operator in the field theoretical sense. In practice, the operator $A^\dagger A$ is represented by a matrix that is, unfortunately, very large. Having a physical volume of size $L^4$, the number of sites is $N^4 = (L/a)^4$. As the internal number of degrees of freedom per lattice point is 12, we end up with $A$ being a $(12N^4) \otimes (12N^4)$ complex matrix with $N$ typically in the range $10 < N < 30$ for present days simulations. Hence we have to construct the square root of a
very large matrix. What is worse, to compute relevant physical observables, we don’t need the operator $D_N$ itself but its inverse. Such an inverse is constructed by iterative methods like the conjugate gradient algorithm and its relatives\cite{14}.

The square root can be constructed by a polynomial expansion, normally based on a Chebyshev approximation, or by rational approximations. Both methods give comparable performances in practice. The convergence of the approximation to the square root is determined by the condition number of the (positive definite) matrix $A^\dagger A$. When the matrix $A^\dagger A$ is normalized such that the largest eigenvalue is one, the condition number is given by the inverse of the lowest eigenvalue. We show in figure 1 the low-lying eigenvalues of $A^\dagger A$ for different values of $\beta$ where the parameter $\beta$ is inversely proportional to the coupling strength of the theory.

It can clearly be observed that very small eigenvalues can occur resulting in large condition numbers. In such situations the convergence of the approximations chosen can be rather slow and special tricks have to be implemented to accelerate the convergence. The most fruitful improvement is to treat a part of the low-lying end of the spectrum exactly by projecting this part out of the matrix $A^\dagger A$\cite{16,17}. Further improvements can be implemented, examples of which are discussed in ref.\cite{17}.
Despite all these technical improvements it is found that a typical value for the degree of a polynomial is $O(100)$ and a typical value for the number of iterations to compute the inverse of $D_N^{-1}$ the Chebyshev polynomial has to be evaluated, this means that for a value of a physical observable on a single configuration ten thousand applications of a huge matrix on a vector has to be performed. To compute the expectation value of some observable, this observables has to be averaged over many gluonic configurations. Clearly, this results in a very demanding computational effort and gives rise to a numerical challenge well suited for NIC.

The only thing that helps a lot in this problem is that the matrix $A$ is sparse. Only the diagonal and a few sub-diagonals are actually filled, a fact that finally makes the problem manageable – although, still a very large amount of computer resources are needed to tackle it.

4 The Scalar Condensate

The existence of an exact lattice chiral symmetry allows the use of finite size effects to test for SSB in QCD as in the case of the $\Phi^4$-theory. Using a chiral invariant formulation of lattice QCD, it is possible to reach the region of very small quark masses where it is to be expected that chiral symmetry starts to get restored.

The “magnetization” in the case of quantum chromodynamics is the condensate of a quark-antiquark state $\Sigma = \langle \bar{\psi} \psi \rangle$. The role of the external magnetic field is played by the quark mass $m_q$. We have developed a fully parallelized code with all technical improvements implemented. This allowed us to compute the scalar condensate as a function of the quark mass at several volumes. A (standard) caveat here is the fact that all computations are done in the so-called quenched approximation where all internal quark loops are neglected. In QCD there is a peculiarity: the field configurations can have topological properties, characterized by the so-called topological charge which can be measured – unambiguously – through the number of zero modes of the operator $D_N$. In fact, the formulae from quenched chiral perturbation theory are parametrized by the topological charge and it is hence very important to be able to identify the topological charge of the gauge field configurations. Without the special properties of lattice Dirac operators that satisfy the Ginsparg-Wilson relation such an identification would be very difficult.

Let us give the complete theoretical formula from quenched chiral perturbation theory in lowest order:

$$\Sigma_{\nu=\pm 1}(m_q,V) = \Sigma \ z \ [I_{\nu}(z)K_{\nu}(z) + I_{\nu+1}(z)K_{\nu-1}(z)] + C \cdot m_q/a^2.$$  \hspace{1cm} (12)

The only important thing to notice here is that this relatively involved combination of Bessel functions do, as in the case of the $\Phi^4$-theory, only depend on one scaling variable

$$z = \Sigma m_qV$$  \hspace{1cm} (13)

that contains the quantity of interest, namely the infinite volume, chiral limit scalar condensate $\Sigma$. The additional term $C \cdot m_q/a^2$ is a power divergence that comes from the renormalization properties of the theory. We will not discuss this field theoretical aspect here but just notice that this term has to be included in the fit.
In figure 2, we show the result of our numerical computation of the scalar condensate\textsuperscript{18} in a fixed topological charge sector $|\nu| = 1$ as a function of the quark mass at several volumes. The solid line is a fit to the prediction of chiral perturbation theory, eq.(12). We find that the simulation data are described by this prediction very well. This means that we find evidence for the basic assumption on which the theoretical prediction relies: the appearance of spontaneous chiral symmetry breaking in (quenched) QCD.

We want to remark that this work that has been performed at NIC was the first of this kind. The project consumed 1400 CPU hours on a typical distribution of the lattice on 128 nodes. After this work, a number of other groups repeated such an analysis\textsuperscript{21–23} and it was reassuring to observe that very consistent results were found. In a subsequent work\textsuperscript{19,20} we developed also a quite general method for renormalizing the value of the bare scalar condensate as extracted from the finite size scaling analysis performed here.
5 Conclusion

In this contribution we have demonstrated that by a combination of theoretical ideas, improved numerical methods and the use of powerful supercomputer platforms it is possible to test basic properties of field theories. Of particular interest was the question of whether the phenomenon of spontaneous symmetry breaking does occur in certain field theories important in elementary particle physics. The phenomenon of SSB leads to far reaching consequences in theories like QCD or the scalar sector of the electroweak interactions. In the work performed here, we found strong evidence for the appearance of spontaneous chiral symmetry breaking in quenched lattice QCD. This conclusion is the result of the fact that two theoretical concepts, the lattice approach to quantum field theories and chiral symmetry, met finally – and that enemies became friends.

We are very much indebted to Hartmut Wittig for numerous discussions, suggestions and finally participating in the project at the stage when the non-perturbative renormalization of the scalar condensate was computed. L.L. thanks the INT at the University of Washington for its hospitality and the DOE for partial support during the completion of this work. This work was supported in part by the EU TMR program under contract FMRX-CT98-0169 and the EU HP program under contract HPRN-CT-2000-00145.

References

10. P. Hernández hep-lat/0110218.
Numerical Simulation of
Supersymmetric Yang-Mills Theory

I. Montvay\textsuperscript{1}, F. Farchioni\textsuperscript{2}, A. Feo\textsuperscript{2}, T. Galla\textsuperscript{2}, C. Gebert\textsuperscript{1}, R. Kirchner\textsuperscript{1}, G. Münster\textsuperscript{2}, R. Peetz\textsuperscript{2}, and A. Vladikas\textsuperscript{3}

\textsuperscript{1} Deutsches Elektronen-Synchrotron DESY
Notkestr. 85, 22603 Hamburg, Germany

\textsuperscript{2} Institut für Theoretische Physik, Universität Münster
Wilhelm-Klemm-Str. 9, 48149 Münster, Germany

\textsuperscript{3} INFN, Sezione di Roma 2, Università di Roma “Tor Vergata”, 00133 Rome, Italy

The DESY-Münster-Roma Collaboration performed the first large-scale simulation of supersymmetric Yang-Mills (SYM) theory. Theoretical expectations about the behaviour of non-perturbative dynamics near the supersymmetric limit at zero gaugino mass have been confronted with numerical results.

1 Introduction

In recent decades research on the structure of matter has revealed that all known matter is composed of a small number of fundamental constituents, the so-called quarks and leptons. Moreover, the physics of these particles is governed by four types of forces: gravitation, electro-magnetic forces, the weak and the strong interactions. In subatomic physics only the latter three play a role. The particles and their interactions are theoretically summarized in the so-called Standard Model. Almost all known experimental data in the presently available energy range can be described by the Standard Model.

In spite of this, many theorists are convinced that by the increase of the energy of particle accelerators within the next ten years new phenomena will emerge which go “beyond the Standard Model”. In particular, the large number of free parameters in the Standard Model is not entirely satisfactory. There are also experimental facts which point towards some extension of the Standard Model: the recently established non-zero mass of neutrinos and the observed baryon matter excess in the Universe which cannot be explained in the minimal Standard Model. In addition, the present theoretical framework has to be extended, if gravitation is to be described in a manner consistent with quantum theory.

Many of the possible extensions of the Standard Model beyond the presently known energy range are based on supersymmetry. Supersymmetry is a concept beyond the usual notion of symmetry. It connects bosons, particles with integral spin, to fermions, which have half-integral spin\textsuperscript{1}. In case of perfect supersymmetry the bosons and fermions are grouped in supermultiplets, in which the members have equal masses. Since such mass degeneracies are not observed in the present experiments one has to assume that supersymmetry is broken. An important consequence of supersymmetry breaking is the heavy mass of the supersymmetric partners of the presently known elementary particles. This would explain why they are not observed up to now.

Although supersymmetry does not seem to immediately solve the problem of proliferation of free parameters, the only known framework which has the potential of incorporating...
the quantum theory of gravity is based on supersymmetric string theories. It is generally assumed that the scale where supersymmetry becomes manifest is near to the presently explored electroweak scale and that the supersymmetry is spontaneously broken.

The Standard Model of elementary particles is based on gauge field theories. Supersymmetric extensions involve supersymmetric gauge theories. A common method to investigate field theories is perturbation theory, which amounts to an expansion of the quantities of interest in powers of a small parameter. However, many interesting features are of a non-perturbative nature. An attractive possibility for spontaneous supersymmetry breaking is to exploit non-perturbative mechanisms in supersymmetric gauge theories. This is the basis of a strong theoretical interest for investigating supersymmetry non-perturbatively.

The motivation for a study of non-perturbative features of supersymmetric gauge theories partly also comes from the desire to understand relativistic quantum field theories better in general: the supersymmetric points in the parameter space of all quantum field theories are very special since they correspond to situations of a high degree of symmetry. The seminal work of Seiberg and Witten and others showed that there is a possibility to approach non-perturbative questions in four dimensional quantum field theories by starting from exact solutions in some highly symmetric points and treat the symmetry breaking as a small perturbation. Beyond this, the knowledge of non-perturbative dynamics in supersymmetric quantum field theories can also be helpful in understanding the greatest puzzle of the Standard Model, with or without supersymmetric extensions, namely the existence of a large number of seemingly free parameters. As we know from Quantum Chromodynamics, the theory of the strong interactions of quarks, strong interactions in non-abelian gauge theories are capable of reproducing a large number of dynamically generated parameters for quantities characterizing bound states from a small number of input parameters. This is a possible solution also for the parameters of the Standard Model if new strong interactions are active beyond the electroweak symmetry breaking scale.

The simplest supersymmetric gauge theory is the supersymmetric extension of Yang-Mills theory, SYM. It describes the carriers of gauge interactions, the gauge particles, together with their supersymmetric partners, the gauginos. Gauginos are massless Majorana fermions, which are in the adjoint representation of the gauge group. The DESY-Münster-Roma Collaboration has mainly investigated the simplest non-abelian SYM theory with gauge group SU(2)\(^3\). (Some first results have also been obtained with SU(3) gauge group.) A non-trivial problem to be solved was to perform the numerical simulations with sufficiently light (Majorana) fermions. A suitable simulation algorithm has been developed which is applicable for light dynamical gauginos and, more generally, also for a broad class of fermionic quantum field theories including Quantum Chromodynamics (QCD).

Our investigations have revealed the basic non-perturbative features of supersymmetric Yang-Mills theory, which will be described below.

## 2 SYM Theory on the Lattice

The numerical simulation of SYM theory is performed on a hypercubic lattice in four dimensional Euclidean space-time. The fourth coordinate, besides the three space coordinates, is the imaginary time. In the path integral formulation we need the Euclidean action which gives the weight of lattice field configurations in the path integral. For a Majorana
fermion in the adjoint representation the fermionic part of the Wilson lattice action is
\[ S_f \equiv \frac{1}{2} \sum_x \left\{ \bar{\lambda}_x \lambda_x - K \sum_{\mu=1}^4 \left[ \bar{\lambda}_{x+\mu} V_{ab,x,\mu} (1 + \gamma_\mu) \lambda_b^a + \bar{\lambda}_{x} V^T_{ab,x,\mu} (1 - \gamma_\mu) \lambda_b^{a+\mu} \right] \right\}. \] (1)

Here \( K \) is the hopping parameter which determines the gaugino mass, \( \gamma_\mu \) denotes a Dirac matrix and \( V_{x,\mu} \) is the gauge field variable in the adjoint representation of the gauge group. The gaugino field \( \lambda \) satisfies the Majorana condition
\[ \bar{\lambda}_x = \lambda_x C, \] (2)
with the charge conjugation Dirac matrix \( C \).

The full lattice action is the sum of the pure gauge part and fermionic part:
\[ S = S_g + S_f. \] (3)

The standard Wilson action for the gauge field \( S_g \) is a sum over the plaquettes
\[ S_g = \beta \sum_{pl} \left( 1 - \frac{1}{N_c} \text{Re} \text{Tr} U_{pl} \right), \] (4)
with the gauge coupling given by \( \beta \equiv 2 N_c / g^2 \).

The path integral over the fermion field \( \lambda \) can be carried out and the result is a Pfaffian:
\[ \int [d\lambda] e^{-\frac{1}{2} \bar{\lambda} Q \lambda} = \int [d\lambda] e^{-\frac{1}{2} \bar{\lambda} M \lambda} = \text{Pf}(M), \] (5)
where \( M \) is the antisymmetric matrix defined as
\[ M \equiv C Q = - M^T. \] (6)

The square of the Pfaffian \( \text{Pf}(M) \) is equal to the determinant of the fermion matrix \( Q \):
\[ \det(Q) = \det(M) = [\text{Pf}(M)]^2. \] (7)

In order to perform Monte Carlo simulations of SYM theory one needs a positive measure on the gauge field which allows for importance sampling of the path integral. Therefore the sign of the Pfaffian has to be taken into account separately. Since the absolute value of the Pfaffian is the non-negative square root of the determinant the effective gauge field action is\(^3\):
\[ S_{CV} = \beta \sum_{pl} \left( 1 - \frac{1}{N_c} \text{Re} \text{Tr} U_{pl} \right) - \frac{1}{2} \log \det Q[U]. \] (8)

The factor \( \frac{1}{2} \) in front of \( \log \det Q \) shows that we effectively have a flavour number \( N_f = \frac{1}{2} \) of adjoint fermions. The omitted sign of the Pfaffian can be taken into account by reweighting the expectation values according to
\[ \langle A \rangle = \frac{\langle A \text{ signPf}(M) \rangle_{CV}}{\langle \text{signPf}(M) \rangle_{CV}}, \] (9)
where \( \langle \ldots \rangle_{CV} \) denotes expectation values with respect to the effective gauge action \( S_{CV} \).
The sign of the Pfaffian may lead in principle to a *sign problem* if the contributions with opposite sign in the reweighting formula (9) cancel each other. This cancellation may lead to an intolerable increase of the statistical error. We have shown\textsuperscript{3–5} that the monitoring of the sign of the Pfaffian can be done with reasonable numerical effort and for positive gaugino masses practically no sign problem occurs because the positive contributions dominate.

3 Low Energy Dynamics in SYM Theory

On the basis of its similarity to QCD one can assume that the basic features of SYM dynamics are similar to QCD: confinement of the coloured degrees of freedom and spontaneous chiral symmetry breaking. As in QCD, a central feature of low-energy dynamics is the realization of the global chiral symmetry. There is only a single Majorana adjoint "flavour" therefore the global chiral symmetry of SYM is an abelian symmetry $U(1)_\lambda$. This is, however, only a symmetry of the classical Lagrangian and it is not fully realized in the quantum field theory. Therefore it is an anomalous symmetry. The remnant symmetry in the quantum theory is $\mathbb{Z}_{2N_c}$ for a gauge group $SU(N_c)$.

**phase structure**

$$\langle \lambda \lambda \rangle < 0$$

$$\langle \lambda \lambda \rangle > 0$$

Figure 1. Expected phase structure of Yang-Mills theory with a Majorana fermion in adjoint representation in the $(\beta, K)$-plane. The dashed-dotted line $K = K_{cr}(\beta)$ is a first order phase transition (or cross-over) at zero gaugino mass.

Similarly to QCD the discrete chiral symmetry $\mathbb{Z}_{2N_c}$ is expected to be spontaneously broken to $\mathbb{Z}_2$ by the non-zero *gaugino condensate* $\langle \lambda \lambda \rangle \neq 0$. The consequence of this
spontaneous chiral symmetry breaking pattern is the existence of a first order phase transition at zero gaugino mass $m_{\tilde{g}} = 0$ (see figure 1). For instance, in case of $N_c = 2$ at $m_{\tilde{g}} = 0$ there exist two degenerate ground states with opposite signs of the gaugino condensate. The symmetry breaking is linear in $m_{\tilde{g}}$, therefore the two ground states are exchanged at $m_{\tilde{g}} = 0$ and there is a first order phase transition. In fact, the $Z_{2N_c}$ chiral symmetry is only expected to be exact in the continuum limit $\beta \to \infty$ (at zero lattice spacing). It is possible that for finite $\beta$ there is a cross-over which becomes a genuine first order phase transition only at $\beta = \infty$.

The DESY-Münster Collaboration performed a first lattice investigation of gaugino condensation in SYM theory with the SU(2) gauge group. The distribution of $\langle \lambda \lambda \rangle$ has been studied for fixed gauge coupling $\beta = 2.3$ as a function of the hopping parameter $K$, which determines the bare gaugino mass, on a $6^3 \cdot 12$ lattice.

A first order phase transition (or cross-over) shows up on small to moderately large lattices as a metastability expressed by a two-peak structure in the distribution of the gaugino condensate. By tuning $K$ one should be able to achieve that the two peaks are equal (in height or area). This is a possible definition of the phase transition point in finite volumes. By increasing the volume the tunneling between the two ground states becomes less and less probable and at some point practically impossible.

The observed distributions are shown in figure 2. One can see that the distributions cannot always be described by a single Gaussian which would correspond to a single phase. The volume dependence of this signal has not been studied up to now. It is left for further studies which could distinguish between a phase transition and a cross-over.

Confinement of the fundamental colour charge is shown by the linear increase of the potential between static colour sources. This fundamental static potential has been extracted, as usual, from Wilson loops and shows indeed a linear rise in agreement with expectations.

The consequence of confinement is that the low energy spectrum in SYM theory consists of colourless hadronic bound states as in QCD. In the supersymmetric limit the states have to belong to supermultiplets with degenerate mass. For non-zero gaugino mass the supersymmetry is softly broken and the supermultiplets are split up in mass. In analogy with the effective chiral Lagrangian of QCD it is possible to derive low energy effective Lagrangians also in SYM theory which reflects the symmetries and is based on an assumption about the nature of the lowest mass states. The simplest assumption gives the Veneziano-Yankielowicz effective Lagrangian but generalizations with more involved supermultiplet structure are possible.

The spectrum of SYM theory with SU(2) gauge group has been studied in numerical simulations by the DESY-Münster Collaboration. A summary of the results on the masses of light bound states is shown in figure 3. The results at the smallest gaugino mass are consistent with two light supermultiplets split up in mass but further investigations in a larger lattice volume are necessary for a more definite statement.

4 Realization of Supersymmetry

An important feature of lattice regularization is that some symmetries are broken for non-zero lattice spacing and are expected to be recovered in the continuum limit. The details of the lattice formulation, which also influence the degree of symmetry breaking, are not
relevant in the continuum limit because of the universality of critical points. A basic set of symmetries broken by the lattice and restored in the continuum limit is the (Euclidean)
Lorentz symmetry including rotations and translations. It is clear that on any regular lattice these symmetries are always broken. Internal symmetries as, for instance, global chiral symmetry are sometimes broken and sometimes conserved on the lattice, depending on the actual formulation. From the point of view of symmetry realization supersymmetry is expected to work similarly to the Lorentz symmetry: at finite lattice spacing it is broken but it becomes restored in the continuum limit. This similarity is quite natural since there is an intimate relation of supersymmetry to the Lorentz symmetries of space-time, shown, for instance, by the fact that the anticommutators of the supersymmetry charges give translations.

In the framework of quantum field theory the symmetries can be exploited by the corresponding Ward-Takahashi (WT) identities. In the case of SYM theory this way of realizing supersymmetry was first considered by Curci and Veneziano. At zero gaugino mass both supersymmetry and anomalous chiral symmetry has to be manifest in the corresponding WT identities. Generally speaking one expects the existence of a renormalized supercurrent operator $S_{R\mu}$ satisfying the Ward-Takahashi-type identity

$$\partial_\mu S_{R\mu} = 2m_R \chi_R.$$  \hspace{1cm} (10)

Here $m_R$ is the renormalized gaugino mass and $\chi_R$ is a suitably defined spinorial density. At zero gaugino mass $m_R = 0$ the supercurrent is conserved and the supersymmetry is exact.

The DESY-Münster-Roma Collaboration studied the SUSY WT-identities in numerical simulations. Omitting $O(a)$ terms which vanish in the continuum limit $a \to 0$ the lattice

---

Figure 3. The lightest bound state masses in lattice units as function of the bare gaugino mass parameter $1/K$ for fixed $\beta = 2.3$. The shaded area at $K = 0.1955(5)$ is where zero gaugino mass and supersymmetry are expected.
version of the SUSY WT-identity can be written as\textsuperscript{12}

\[
\langle \nabla_\mu S_\mu(x)O(y) \rangle + \frac{Z_T}{Z_S} \langle \nabla_\mu T_\mu(x)O(y) \rangle = \frac{m_S}{Z_S} \langle \chi(x)O(y) \rangle .
\]

(11)

\(S_\mu(x)\) is a suitably defined supercurrent and \(T_\mu(x)\) is an admixture with the same quantum numbers as \(S_\mu(x)\). Eq. (11) has to be valid for every gauge invariant function \(O(y)\) which is defined in a point \(y\) in such a way that there are no common points with the functions defined at \(x\). Considering (11) with different \(O(y)\) a system of linear equations is obtained for the two unknowns

\[
z_{TS} \equiv \frac{Z_T}{Z_S}, \quad z_{MS} \equiv \frac{m_S}{Z_S}.
\]

(12)

(Note that \(z_{MS}\) is proportional to the renormalized gaugino mass \(m_R \equiv m_S/Z_m\).) Since there are many different possible \(O(y)\)'s, the expectation that (12) has, in the continuum limit, a unique solution pair \(z_{TS}, z_{MS}\) is highly non-trivial. Its numerical investigation can strongly support (or possibly contradict) our expectations about the realization of supersymmetry in SYM theory.

Figure 4. The variable \(z_{MS}\) as a function of the bare gaugino mass. An extrapolation to zero defines the point where supersymmetry is realized in the continuum limit.

The numerical data of the DESY-Münster-Roma Collaboration (see for instance in figure 4) demonstrate the feasibility of implementing lattice SUSY WT identities in order to verify supersymmetry restoration in a non-perturbative framework.
5 Conclusions

In the context of our investigations of supersymmetric Yang-Mills theories on a lattice we have developed an efficient algorithm for simulating dynamical Majorana fermions, corresponding to $N_f = 1/2$ fermion flavours, on massively parallel supercomputers. Basic features of the model have been studied on lattices of different sizes. This includes the phase structure, gaugino condensate, sign of the Pfaffian, confinement potential, spectrum of low-lying bound states, and supersymmetric Ward-Takahashi identities.

Most of the numerical calculations presented in this report have been performed on the Cray T3E of the NIC, Jülich. The total CPU time since 1997 was about $2 \cdot 10^6$ hours.

References

Computer simulations offer immense scope to study the properties of materials. A central goal is the understanding of macroscopic properties on the basis of microscopic material parameters, and the latter are often more accessible to numerical calculation than to experimental measurement. Nevertheless, severe problems confront the computational scientist who is seeking this goal, and the first is provided by the great range of length and time scales involved. For some properties a continuum description of the structure might be appropriate, allowing phenomena to be studied over macroscopic lengths and time scales of the order of seconds or much longer. At the other extreme, we have properties – such as excited states or reactivity – that are determined by the motion of the individual atoms, involving lengths of the order of interatomic separations and time scales of \(10^{-13}\) seconds or shorter. The answers to such questions require a detailed understanding of the geometry and the electronic structure of the material involved.

The development of a numerical method that spans these ranges continuously is out of the question and would not make much sense even if possible. Nevertheless, there is much effort devoted at the present time to building bridges between the methods now established in different length and time regimes. These efforts to develop “multi-scale” methods for simulations are reflected in the frequent use of this word in conference announcements.

An idea of the range of scales involved can be obtained from the contributions to this symposium. Molecular dynamics (MD) simulations are now fifty years old and enjoy a conceptual simplicity. The forces in the system are described by (usually) simple force laws or potential functions, and the motions of the components are determined by solving Newton’s equations, which introduce a natural time scale. They differ in this regard from Monte Carlo methods, which are purely stochastic and have no built-in time scale. In the contribution of Yamakov et al., MD calculations are performed on polycrystalline aluminium with small grain sizes (under 100 nm). The motion of the grain boundaries is crucial for the mechanical properties of these materials and can be followed with this technique during low-temperature deformation. The calculations provide insight into the motion of the atoms during an extremely complicated process.

The interatomic potential used in these simulations is fitted to experimental data including the elastic constants. The question: “Do the results of the simulations depend on the particular parameterization or the choice of experimental data?” is oft-repeated, but the size and complexity of the simulation make simplifications unavoidable. In the contribution of Brinkmann et al., where metallic glasses and glass-forming melts are simulated under external shear deformations or reduced dimensions, the force field for Ni\(_{0.5}\)Zr\(_{0.5}\) and related systems used short-ranged pair potentials based on other calculations. It can be seen that much information can be found from these calculations concerning the effects of changing the temperature, the composition, or the dimensionality of the samples.
It is natural to ask whether molecular dynamics can be used in the context of electronic structure theory. Since the work of Car and Parrinello\(^1\) in 1985 the answer has been “yes”. The “Car-Parrinello method” uses the density functional (DF) approach to calculate the energy and forces in the system, using the latter as the basis of molecular dynamics. This approach avoids the parameterization of the force fields mentioned above, but the calculations are numerically very demanding and are limited by present computational technology to around 200 atoms and time scales of the order of \(10^{-12}\) seconds. These restrict greatly the range of phenomena that can be studied, but the predictive capability of the density functional method is widely accepted and has important advantages.

An oral presentation of an application of this method was given by the present author during this symposium\(^2\) and described the reaction of a lithium phenoxide molecule with a ring oligomer of an important industrial polymer, bisphenol A polycarbonate (a). The complex process of breaking the ring and creating a chain with a Li ion at the end (b) is described in detail by these calculations, and the energy barriers found are consistent with experimental evidence. These calculations have been extended to even more complex reactions where little experimental information is available, and the reliable data bases provided by the calculated energy surfaces can be used to develop or refine force field parameters for simulations of much larger systems. Density functional calculations require an approximation for the exchange-correlation energy, and it would be surprising if the relatively simple functionals often used (the local density approximation and its modifications) could describe all problems of interest. One important problem where the local density approximation gives qualitatively incorrect predictions is the subject of the contribution of Schmid et al. The question of whether an effective single-particle theory can describe the antiferromagnetic transition metal oxides has haunted density functional workers for many years, as the DF predictions of metallic behaviour run counter to the insulating state found experimentally. The approach taken by Schmid et al. is to separate \(E_{xc}\) into exact exchange and a DF description of correlation. This work is still in progress, and extensions must be made before a final assessment can be made.

References

Application of Implicit Density Functionals
to 3d Transition Metal Monoxides

R. N. Schmid, E. Engel, and R. M. Dreizler

Institut für Theoretische Physik, J.W.Goethe-Universität Frankfurt
Robert-Mayer-Str.8-10, 60054 Frankfurt am Main, Germany
E-mail: {rschmid, engel, dreizler}@th.physik.uni-frankfurt.de

Standard density functional calculations on the basis of the local density approximation (LDA) predict the type II antiferromagnetic phases of FeO and CoO to be metallic, in contradiction to experiment. Within the framework of density functional theory this failure must be attributed to the shortcomings of the LDA, in particular to its incorrect treatment of the self-interaction. In this contribution we report results for FeO obtained with an implicit density functional including the exact exchange, which guarantees a complete elimination of the self-interaction.

1 Introduction

It is a long-standing question whether an effective single-particle description of Mott insulators is possible. Although an enormous amount of work has been devoted to this question over the years (see e.g.,), a clear-cut answer is still not available. This is true not only for excitations and phase diagrams, but even for the ground states at $T = 0$. In this contribution we reconsider the ground state problem within density functional theory (DFT), which currently is the most successful effective single-particle approach.

The prototype Mott insulators are the transition metal monoxides MnO, FeO, CoO and, in particular, NiO. In order to illustrate the challenge which these systems pose to a single-particle approach we consider NiO in more detail. Like MnO, FeO and CoO, NiO crystallizes in a rock salt structure with a lattice constant of 7.93 Bohr, as shown in Fig.1 (with a minor distortion in the case of FeO which will be ignored in the following). The basic features of the band structure of NiO, and thus its insulating property, can be understood as follows. The transition metal compounds typically exhibit a strong overlap

![Figure 1. AFII structure of NiO.](image_url)
of the metal 4s, 4p orbitals with their r-expectation values of about 3 Bohr and the oxygen 2s, 2p states. This leads to a strong hybridization of these states, so that the two 4s electrons of the metal are transferred to the oxygen 2p band and the gap between the 2s, 2p and the 4s, 4p bands becomes rather large. In between one finds, more or less well separated from the 2s, 2p and 4s, 4p bands, the 3d bands which, in the case of NiO, are occupied by 8 electrons per atom. The detailed structure of these bands can be explained by sequentially adding the various relevant effects to the d level of an unpolarized atom\(^6\), as illustrated in Fig.2. First of all, the exchange interaction splits the atomic level into a fully occupied, more strongly bound majority (↑) and a partially occupied, less bound minority (↓) spin level. For NiO this splitting is of the order of 1.6 eV. Next one has to take into account the octahedral symmetry of the lattice, i.e. the six doubly charged oxygen ions which surround each cation. This crystal field effect splits the five spatially distinct d orbital of each spin into an energetically lower triplet of \(t_{2g}\) symmetry \((xy, yz\text{ and }zx\text{ character})\) and an energetically higher doublet of \(e_g\) symmetry \((x^2 - y^2\text{ and }3z^2 - r^2\text{ character})\). The crystal field splitting is also of the order of 1 eV, so that no clear separation of up- and down-spin states emerges. At this point the description of NiO within a band theory framework seems to run into a fundamental problem: If one includes the periodicity of the system all subbands acquire a certain width \(W\). Due to the small overlap of the d-orbitals from neighboring atoms the size of \(W\) can be estimated on the basis of the tight binding method: One finds that \(W\) is directly proportional to the nearest neighbor hopping integral and thus, indirectly, to the size of the crystal field splitting. As a result the \(t_{2g}\) and \(e_g\) subbands seem to overlap, so that a metallic state seems to emerge, in contradiction to experiment.

In order to understand the insulating nature of NiO its antiferromagnetic ordering has to be taken into account. All four transition metal monoxides of interest here are type II antiferromagnets (AFII) for which the majority spins of the atoms in (111) planes are
aligned, while the direction of the alignment alternates from plane to plane (see Fig.1), i.e.
one has two magnetic sublattices and thus an effective doubling of the lattice constant. The
impact of the antiferromagnetic structure on the $d$ bands can be extracted from an analysis
of the relevant interaction hopping integrals, which characterize the coupling (and thus
hybridization) of the $d$-states of neighboring Ni atoms. One finds that the strongest hopping
integral is the $(dd\sigma)$-type coupling between second-nearest neighbor cations (along the
cartesian axes) which is mediated by the $p_x$, $p_y$ and $p_z$ orbitals of the anions located in the
middle between the cations. As Fig.1 shows, the $(dd\sigma)$-interaction couples cations states
of opposite polarization in the case of the AFII ordering. As a consequence the $(dd\sigma)$-
interaction always links a majority spin $e_g$ state of one magnetic sublattice with a minority
spin $e_g$ state in the other sublattice (intersublattice coupling). Their hybridization then
leads to an energetic lowering of the occupied majority spin $e_g$ states of a given magnetic
sublattice and a corresponding shift of the unoccupied minority spin $e_g$ states to higher
energies. The $t_{2g}$-states, on the other hand, are not affected as they are fully occupied and
only contribute to $(dd\pi)$- and $(dd\delta)$-coupling. The fact that the $(dd\sigma)$-interaction only
couples cations from different sublattices is also responsible for the rather narrow width of the $e_g$ bands in the AFII structure (as compared with the width of the $t_{2g}$-bands or the
ferromagnetic phase). Thus, combining the exchange and crystal field splittings with the
antiferromagnetic hybridization shifts, the occupied bands are sufficiently well separated
from the unoccupied $e_{g,1}$ band that the banding effect (intrasublattice coupling) cannot
close the gap between these levels.

In the same fashion one can explain the insulating nature of MnO. In this case the
large exchange splitting of roughly 3.8 eV is sufficient to produce a gap between the fully
occupied majority spin $e_g$- and the completely unoccupied minority spin $t_{2g}$-band. None
of the arguments above is in conflict with the single-particle picture, and, in fact, standard
density functional calculations yield gaps for both NiO and MnO.

More difficult to describe with single-particle methods are the remaining two systems
FeO and CoO for which the minority spin $t_{2g}$-band is partially occupied. Focusing on
FeO, a single-particle description requires that for each site in a given sublattice two of the
$t_{2g,1}$-states are empty, thus breaking the octahedral symmetry. For both spins the $t_{2g}$-level
is thus split into a singly occupied state $t_{2g}$ (e.g. the $xy$-orbital) and a doubly degenerate
level $e_g$. The impact of this broken symmetry is most easily seen within the Hartree-Fock
(HF) approximation. In its most simple-minded version one can consider the $3d$ states in a
net crystal field $v_{CF}$ produced by the nuclei and all other electrons. The single-particle
energies $\epsilon_{k,\sigma}$ of the $d$-levels are then given by

$$\epsilon_{k,\sigma} = \int d^3r \phi_{k,\sigma}^\dagger(r) \left\{ -\frac{\nabla^2}{2m} + v_{CF}(r) \right\} \phi_{k,\sigma}(r) + e^2 \sum_{l\tau} \Theta_{l\tau} [(k\sigma, l\tau|k\sigma, l\tau) - (k\sigma, l\tau|l\tau, k\sigma)],$$

where $\phi_{k,\sigma}$ represents the localized $d$-orbital of symmetry $k$ and spin $\sigma$, $\Theta_{k,\sigma}$ denotes its occupation and $(ab|cd)$ is the standard Slater integral,

$$(k_1\sigma_1, k_2\sigma_2|k_3\sigma_3, k_4\sigma_4) = \int d^3r d^3r' \frac{\phi_{k_1,\sigma_1}^\dagger(r) \phi_{k_2,\sigma_2}^\dagger(r') \phi_{k_3,\sigma_3}(r) \phi_{k_4,\sigma_4}(r')}{|r - r'|}.$$  

Starting from the atomic $d$-level, the first line of Eq.(1) leads to the $t_{2g}$ and $e_g$ states in the
crystal field. The differences between the two spin orientations and between occupied and unoccupied levels result from the second line. On a qualitative level the effect of the second line can be extracted within the most simple approximation to the relevant Slater integrals: Let $S$ denote the self-interaction energy of any of the $10d$ states, $U$ be the direct interaction between two spatially distinct states and $X$ be the corresponding exchange integral,

$$S = (k\sigma, k\tau || k\sigma, k\tau) \quad (3)$$
$$U = (k\sigma, l\tau || k\sigma, l\tau) \quad (k \neq l) \quad (4)$$
$$X = (k\sigma, l\sigma || l\sigma, k\sigma) \quad (k \neq l). \quad (5)$$

This crudest possible approach ignores both the differences between the up- and down-orbitals and those between $t_{2g}$ and $e_g$ states. The second line of Eq.(1) then contributes to the eigenenergies of the various states as

$$t_{2g, \downarrow}^d, e_{g, \downarrow} : (S + 5U) - X \quad (6)$$
$$t_{2g, \downarrow}^s : (2S + 4U) - S \quad (7)$$
$$t_{2g, \uparrow}^s : (2S + 4U) - (S + 4X) \quad (8)$$
$$t_{2g, \uparrow}^d, e_{g, \uparrow} : (S + 5U) - (S + 4X). \quad (9)$$

If one now takes into account the fact that $S, U \gg X$ a gap of the order $U - X$ opens between the $t_{2g, \downarrow}$ level and the unoccupied states. Due to the strong localization of the $d$-states, which leads to a rather large $U$, this gap survives the broadening of the levels by the intrasublattice coupling. Thus, assuming that the coupling between the cells leads to a well-defined spatial pattern of occupied $t_{2g, \downarrow}$-states, this consideration shows that a single-particle description of FeO and CoO is not impossible.

An equivalent argument can be given in the case of DFT. In order to set the stage for the subsequent discussion let us first briefly summarize the essentials of DFT at this point.

### 2 Density Functional Theory

In DFT one uses the unique relation between the ground state of the system and its ground state density to recast the many-particle problem in the form of an auxiliary single-particle problem which reproduces the ground state density and energy of the actual many-body system exactly (but not its ground state)\(^9\). All complicated many-body effects are relegated to the so-called exchange-correlation (xc) energy $E_{xc}$ which, in the spin-dependent version of DFT (which we will use throughout this paper), is a universal (i.e. system-independent) functional of the spin-up ($n_\uparrow$) and spin-down ($n_\downarrow$) densities, $E_{xc}[n_\uparrow, n_\downarrow]$. The xc-energy is introduced as the difference between the exact ground state energy $E_{tot}$ of the interacting system and suitable single-particle components of $E_{tot}$,

$$E_{tot} = T_s[n_\uparrow, n_\downarrow] + \int d^3r \, v_n(r)n(r) + \sum_{\alpha<\beta} \frac{Z_\alpha Z_\beta e^2}{|R_\alpha - R_\beta|}$$
$$+ \frac{e^2}{2} \int d^3r \int d^3r' \, \frac{n(r)n(r')}{|r - r'|} + E_{xc}[n_\uparrow, n_\downarrow]. \quad (10)$$
Here \( n = n_\uparrow + n_\downarrow \) is the total density, \( v_n \) represents the potential of the nuclei at the positions \( R_\alpha \),

\[
v_n(r) = -\sum_\alpha Z_\alpha e^2 \frac{1}{|r - R_\alpha|},
\]
and \( T_s[n_\uparrow, n_\downarrow] \) is the kinetic energy of the auxiliary single-particle system. \( T_s \) can be specified explicitly in terms of the single-particle orbitals \( \phi_{k,\sigma} \) of the auxiliary problem which yields the exact spin-densities,

\[
n_\sigma(r) = \sum_k \Theta_{k,\sigma} |\phi_{k,\sigma}(r)|^2
\]

\((13)\)

\[
\Theta_{k,\sigma} = \begin{cases} 
1 & \text{for } \epsilon_{k,\sigma} \leq \epsilon_{F,\sigma} \\
0 & \text{for } \epsilon_{F,\sigma} < \epsilon_{k,\sigma}. 
\end{cases}
\]

\((14)\)

The \( \phi_{k,\sigma} \) are determined by a minimization of \( E_{tot} \), Eq.(10), which leads to the single-particle equations of the auxiliary problem, the Kohn-Sham (KS) equations,

\[
\begin{aligned}
&\left\{ \frac{-\nabla^2}{2m} + v_n(r) + v_H([n]; r) + v_{xc,\sigma}([n_\uparrow, n_\downarrow]; r) \right\} \phi_{k,\sigma}(r) = \epsilon_{k,\sigma} \phi_{k,\sigma}(r),
\end{aligned}
\]

\((15)\)

with the Hartree (H) and exchange-correlation (xc) potentials

\[
v_H(r) = e^2 \int d^3 r' \frac{n(r')}{|r - r'|},
\]

\((16)\)

\[
v_{xc,\sigma}(r) = \frac{\delta E_{xc}[n_\uparrow, n_\downarrow]}{\delta n_\sigma(r)}.
\]

\((17)\)

Selfconsistent solution of Eqs.(13-17) with the exact \( E_{xc}[n_\uparrow, n_\downarrow] \) leads to the exact \( n_\sigma \) and \( E_{tot} \) of the actually interesting system.

In view of the multiplicative nature of the potential in (15) the existence of a gap in FeO requires a strong spatial variation of \( v_{xc,\downarrow} \); \( v_{xc,\downarrow} \) must be much more attractive for the occupied \( t_{2g,\downarrow} \)-state than for the unoccupied \( t_{2g,\uparrow} \)-states. From Eqs.(6)-(9) it is clear that the driving force for the opening of a gap is the self-interaction correction component in the exchange energy: While it is present for the \( t_{2g,\downarrow} \)-state, it is missing for the \( t_{2g,\uparrow} \)-orbital. Thus the local manifestation of the self-interaction correction in \( v_{xc,\downarrow} \) must be responsible for the spatial structure attracting the \( t_{2g,\downarrow} \)-state. It is thus no surprise that the most simple approximation for \( E_{xc}[n_\uparrow, n_\downarrow] \), the local density approximation (LDA), predicts FeO and CoO to be metallic \(^7,^8\) (while the insulating ground states of MnO and NiO are correctly reproduced): The LDA, which treats the actual inhomogeneous system locally as if it was a homogeneous electron gas, is well known for its insufficient self-interaction correction in the atomic and molecular context \(^9\).

Unfortunately, the standard extension of the LDA by inclusion of density gradients in the form of the generalized gradient approximation (GGA) does not really resolve the difficulties with FeO and CoO. Both the band structures and the magnetic moments obtained with the most frequently applied GGAs \(^11\) are very similar to their LDA counterparts \(^7,^8\). On the other hand, a GGA whose form has been optimized to reproduce atomic exchange
potentials as accurately as possible does predict FeO and CoO to be antiferromagnetic insulators. While the size of the gaps found is much too small, this result nevertheless provides a proof of principle that the mechanism described above can work in practice. Moreover, there is one further indication that the inappropriate handling of the self-interaction is responsible for the failure of the LDA: An explicitly self-interaction corrected form of the LDA gives the correct ground states.

This leads to the question whether an exact treatment of exchange is possible within DFT? The exact exchange energy functional $E_x$ is defined via the Fock expression, using the KS-orbitals,

$$E_x = -\frac{e^2}{2} \sum_{k} \sum_{\sigma} \Theta_{k,\sigma} \Theta_{l,\sigma} \int d^3r \int d^3r' \frac{\phi_{k,\sigma}^\dagger (r) \phi_{l,\sigma} (r') \phi_{l,\sigma}^\dagger (r') \phi_{k,\sigma} (r)}{|r - r'|}. \quad (18)$$

Eq.(18) represents an implicit density functional as the $\phi_{k,\sigma}$ are unique functionals of $n_{\sigma}$ (this concept has already been used for $T_s$). Explicit use of (18) automatically guarantees the exact elimination of the self-interaction energy and should thus be ideally suited for the description of Mott insulators. The question which immediately arises in view of the orbital-dependent expression (18) is how to evaluate the corresponding xc-potential?

### 3 Optimized Potential Method

In contrast to the case of explicitly density-dependent functionals, for which the required functional derivative $\delta E_{xc}/\delta n_{\sigma}$ can be taken analytically, a more indirect, numerical procedure must be used for implicit functionals. The crucial equation of this so-called Optimized Potential Method (OPM) is most easily derived by transforming the functional derivative with respect to $n_{\sigma}$ into a derivative with respect to $\phi_{k,\sigma}$, using the chain rule for functional differentiation. One ends up with a linear integral equation for $v_{xc,\sigma}$,

$$\int d^3r' \chi_{s,\sigma}(r, r') v_{xc,\sigma}(r') = \Lambda_{xc,\sigma}(r), \quad (19)$$

whose ingredients are the static response function of the KS auxiliary system,

$$\chi_{s,\sigma}(r, r') = -\sum_k \Theta_{k,\sigma} \phi_{k,\sigma}^\dagger (r) G_{k,\sigma}(r, r') \phi_{k,\sigma} (r') + c.c. \quad (20)$$

$$G_{k,\sigma}(r, r') = \sum_{l \neq k} \frac{\phi_{l,\sigma}^\dagger (r) \phi_{l,\sigma}^\dagger (r')}{\epsilon_{l,\sigma} - \epsilon_{k,\sigma}}, \quad (21)$$

and an inhomogeneity which contains the derivative $\delta E_{xc}/\delta \phi_{k,\sigma}$,

$$\Lambda_{xc,\sigma}(r) = -\sum_k \int d^3r' \phi_{k,\sigma}^\dagger (r) G_{k,\sigma}(r, r') \frac{\delta E_{xc}}{\delta \phi_{k,\sigma}^\dagger (r')} + c.c. \quad (22)$$

Eq.(19) determines $v_{xc,\sigma}$ only up to some additive constant, as

$$\int d^3r' \chi_{s,\sigma}(r, r') = 0. \quad (23)$$

The overall normalization of $v_{xc,\sigma}$ must be fixed by some additional constraint. Eq.(19) has to be solved selfconsistently together with the KS equations, i.e. its solution for fixed $\phi_{k,\sigma}$
replaces the insertion of $a_\sigma$ into conventional xc-functionals. Eqs. (19)-(23) have been formulated for arbitrary implicit xc-functionals, including a possible correlation component $E_c$. However, to date no suitable orbital-dependent approximation for $E_c$ is available, so that the exact $E_c$ is usually combined with a standard density functional for $E_x$.\textsuperscript{18–21}

The evaluation of $\chi_{s,\sigma}$ is computationally rather demanding, as a large number of excited states have to be taken into account in the evaluation of the Greens function (21). Thus an approximate (semi-analytical) solution of Eq. (19) is obviously of interest. Such a solution can be obtained by use of a closure approximation for $G_{k,\sigma}$, which leads to the Krieger-Li-Iafrate approximation (KLI)\textsuperscript{16, 22}

$$\chi^{KL}_k(r) = \frac{1}{2n_{\sigma}(r)} \sum_k \left\{ \frac{\delta^4}{\delta \phi_k^{\dagger}(r)} \delta E_{xc} + c.c. \right\} + \left| \phi_k^{\dagger}(r) \right|^2 \Delta v_k,\sigma \left( \Theta_k,\sigma \right) \right\} \right\} (24)$$

$$\Delta v_{k,\sigma} = \int d^3r \left\{ \Theta_{k,\sigma} |\phi_{k,\sigma}(r)|^2 v^{KL}_{x,\sigma}(r) - \phi_{k,\sigma}^{\dagger}(r) \frac{\delta E_{xc}}{\delta \phi_k^{\dagger}(r)} \right\} + c.c. \right\} (25)$$

Eqs. (24), (25) can either be solved iteratively or resolved via a set of linear equations for the unknown component of the right-hand side, $\int d^3r |\phi_k|^2 v_{xc}$. The accuracy of the KLI approximation is most easily demonstrated for atoms. In Table 1 we list the ground state energies obtained with the exact $E_x$, either on the basis of the full OPM or with the KLI-approximation, as well as with the LDA and a widely used GGA for a number of atoms. In all calculations correlation has been neglected (x-only limit). For comparison the corresponding HF numbers are also given. Table 1 demonstrates that x-only OPM and HF energies are very close to each other, reflecting the limited importance of the nonlocality of the HF exchange potential for the variational freedom (in both cases the same energy expression is minimzed). More important in the present context is the observation that the KLI-approximation is extremely accurate. The LDA and GGA, on the other hand, exhibit the well-known deviations from the exact result.

The accuracy of the KLI-approximation can also be examined on a more microscopic level by analyzing the exchange potential itself. In Fig. 3 we plot the exchange potentials obtained by selfconsistent OPM, KLI, LDA and GGA calculations for neon. For all $r$ the KLI potential is very close to the exact OPM result. Only the very pronounced shell-structure in the OPM potential is smoothed out (note, however, that this feature is
Figure 3. Exchange potential of Atoms: KLI approximation versus full OPM solution. Selfconsistent OPM, KLI, LDA and GGA results for Ne.

completely absent in the LDA and GGA potential). The OPM and KLI potentials are particularly close in the asymptotic regime, in which they both decay like \(-1/r\). This behavior reflects the fact that both potentials are self-interaction free (compare\(^{25}\)). An analogous comparison for metallic aluminum is provided by Fig.4. Again the accuracy of the KLI-approximation is apparent. One can thus reliably replace the solution of Eq.(19) by the KLI form (24).

4 Application of Exact \(E_x\) to FeO

In order to study the implications of the exact exchange for the transition metal oxides we have applied the combination of the exact \(E_x\) with the LDA for correlation to the AFII structure of FeO. In view of the enormous computational demands of an OPM calculation for this system we have resorted to (a) the KLI approximation and (b) a pseudopotential approach (analogous to\(^{19,21}\)). The valence space consisted of the iron 3\(p\), 3\(d\) and 4\(s\) states. Normconserving pseudopotentials of Troullier-Martins type\(^{26}\) have been used (compare\(^{27}\)). A plane-wave basis with an energy cut-off of 250 Ryd was required for an accurate description of the 3\(p\) states. Test calculations with the LDA showed that 3 special \(k\)-points were sufficient for the integration over the Brillouin zone\(^{28}\).

The resulting band structure is shown in Fig.5. As in the case of the LDA, a metallic ground state is found. On the other hand, many details of the band structure are rather different from the LDA result (compare\(^{7,8}\)). In particular, the exact \(E_x\) reproduces the separation of the oxygen 2\(p\) from the iron 3\(d\) bands observed in experiment\(^{29}\). The same effect has been seen in a calculation with the exact \(E_x\) for MnO\(^{20}\).

In view of this somewhat ambiguous result the question arises to what extent the technical limitations of the calculation might affect the band structure. As we have demonstrated the accuracy of the KLI-approximation in Section 3 and explicitly verified by LDA calculations that the basis set size and the Brillouin zone sampling are sufficient, only two aspects appear to be worth a closer examination. On the one hand, there are indications
that the combination of the exact $E_x$ with the LDA for $E_c$ suffers from the overestimation of correlation energies and potentials by the LDA (in a pure LDA calculation this effect fortuitously cancels with the LDA's underestimation of exchange energies and potentials\cite{25}). However, as the discussion of Section 1 identifies the exchange as the component which is responsible for the opening of the gap, it is unlikely that the failure of the present calculation is due to the shortcomings of the LDA for $E_c$.

On the other hand, the exclusion of the 3s electrons from the valence space seems to be critical: It is well known from LDA pseudopotential calculations for iron that the 3s electrons must be included (at least in the form of nonlinear core corrections — compare\cite{30,31}). This fact is illustrated in Fig.6 where the spin-up eigenvalues of diatomic FeO are shown. The all-electron values are compared with three pseudopotential results, based on different valence spaces (the LDA is utilized for this comparison as all-electron OPM results for FeO are not available). It is obvious that the levels obtained with the 3p3d4s-type valence space are much less realistic than those found with the more complete valence space including the 3s electrons (nonlinear core corrections are not used). A detailed study of the ground and lowest excited states of FeO shows that the excitation energies of this molecule depend very sensitively on the size of the valence space\cite{31}. In fact, with the 3p3d4s-type valence space an incorrect ground state is predicted if nonlinear core corrections are not included (as in our OPM calculation for solid FeO). On the other hand, the 3s3p3d4s valence space
Figure 5. Band structure of FeO in AFII structure ($a = 8.145$ Bohr): Plane-wave-pseudopotential calculation with exact $E_x$, using (a) the KLI-approximation, (b) the LDA for $E_x$, (c) $E_{cut} = 250$ Ryd, (d) a valence space including the iron $3p$, $3d$ and $4s$ states, (e) no nonlinear core corrections.

Figure 6. Spin-up eigenvalues of diatomic FeO (for $R = 3.0$ Bohr): Various pseudopotentials (PP) without nonlinear core corrections in comparison with all-electron (AE) result ($\sigma$-levels — solid lines, $\pi$-levels — dashed lines, $\delta$-levels — dotted lines).

yields the correct ground state. A definitive conclusion concerning the importance of the exact $E_x$ for the insulating nature of FeO thus requires the inclusion of the $3s$ states, in spite of the resulting increased computational demands.
Acknowledgments

Computer time for this work has been provided by the John-von-Neumann Institute for Computing at the Research Centre Jülich and the Hochschulrechenzentrum of the Univ. of Frankfurt. Financial support by the Deutsche Forschungsgemeinschaft (grant Dr 113/20-3) is gratefully acknowledged.

References

Dislocation Processes and Deformation Twinning in Nanocrystalline Al

Vesselin Yamakov\textsuperscript{1,2}, Dieter Wolf\textsuperscript{2}, Simon R. Phillpot\textsuperscript{2}, and Herbert Gleiter\textsuperscript{1}

\textsuperscript{1} Forschungszentrum Karlsruhe
76021 Karlsruhe, Germany
E-mail: Erika.Schuetze@int.fzk.de
\textsuperscript{2} Argonne National Laboratory
Argonne, IL 60439, USA
E-mail: \{yamakov, wolf, phillpot\}@anl.gov

Using a recently developed, massively parallel molecular-dynamics (MD) code for the simulation of polycrystal plasticity, we analyze for the case of nanocrystalline Al the complex interplay among various dislocation and grain-boundary processes during low-temperature deformation. A unique aspect of this work, arising from our ability to deform to rather large plastic strains and to consider a rather large grain size, is the observation of deformation under very high grain-boundary and dislocation densities, i.e., in a deformation regime where they compete on an equal footing. We are thus able to identify the intra- and intergranular dislocation and grain-boundary processes responsible for the extensive deformation twinning observed in our simulations. This illustrates the ability of this type of simulations to capture novel atomic-level insights into the underlying deformation mechanisms not presently possible experimentally.

1 Introduction

The mechanical behavior of nanocrystalline materials (i.e., polycrystals with a grain size of less than 100 nm) remains controversial\textsuperscript{1}, the observations ranging from greatly enhanced ductility\textsuperscript{2-4} to dramatically increased strength and hardness\textsuperscript{5-7}. While it is commonly accepted that the intrinsic deformation behavior of these fascinating materials must arise from the dynamical interplay between dislocation and grain-boundary (GB) processes\textsuperscript{8}, little is known to date on the specific deformation mechanisms. Here we describe large-scale molecular-dynamics (MD) simulations of nanocrystalline-Al model microstructures which begin to elucidate this intricate, highly non-linear interplay during room-temperature plastic deformation. We hope to demonstrate that these simulations have now advanced to a level where they provide a powerful new tool not only for exposing the atomic-level mechanisms controlling the complex dislocation-dislocation interaction processes in heavily deformed nanocrystals, but also for elucidating the response of the GB network to internal and external stresses.

The common low-temperature plastic-deformation mechanism in most metals and ceramics involves the continuous nucleation of dislocations from Frank-Read sources and their glide, on well-defined slip systems, through the crystal. In a polycrystalline material the size of these sources cannot exceed the grain size. Since the stress needed for their operation is inversely proportional to the size of the source, this deformation mechanism can operate only down to a grain size of typically about 1\(\mu\)m. For a smaller grain size, mobile dislocations must be nucleated from other sources, such as the GBs or grain junctions.
Figure 1. Snapshot of a $<110>$ textured Al polycrystal with a grain size of 30 nm after 27.6 ps of tensile deformation under a stress of 2.3 GPa applied in the horizontal direction at 300 K. A complete, extended dislocation emitted from the GB on the left is traveling across grain 2. Common-neighbor analysis was used to identify perfect-crystal atoms as being either in a local hcp (red atoms) or fcc environment (atoms not shown). Blue atoms are ‘defected atoms’, i.e., atoms neither in an fcc nor hcp environment.

2 Simulation Approach

To address this question, recent MD simulations of rather simple, $<110>$ textured Al polycrystals with a grain size of up to 70 nm revealed the onset of plastic flow when the applied tensile stress exceeds the threshold stress for the nucleation of dislocations from the GBs or grain junctions. Idealized microstructures consisting of four grains of identical size and a regular-hexagonal shape were considered (see Fig. 1); the grain orientations were chosen such that all 12 GBs in the simulation cell are asymmetric high-angle $<110>$ tilt boundaries. The $<110>$ column axis was chosen such that, following their nucleation, dislocations can glide in each grain on either of two (111) slip systems, unimpeded by the three-dimensionally periodic border conditions imposed on the simulation cell.

A many-body interatomic potential for Al fitted, among other parameters, to the elastic constants is used. The potential has been slightly modified for a higher degree of smoothness at the cut-off radius. While leaving other physical parameters practically unchanged, this modification increases the stacking-fault energy from 104 to 122 $mJ/m^2$, i.e., closer to the experimental values ranging between 120 and 142 $mJ/m^2$. 

226
3 Simulation Results

Figure 1 illustrates the emission of an extended 1/2[011] dislocation into grain 2 from a GB in the vicinity of a triple junction. The complete nucleation process involves the successive emission of both the leading and trailing Shockley partials (seen as the terminating blue atoms) connected by a stacking fault (red atoms). Following their complete nucleation, these extended dislocations travel across the grain on one of the available (111) slip planes, until they annihilate in the opposite GB, converting their Burgers vector into plastic strain.

It is interesting to note that for low plastic strain (i.e., in the absence of dislocation-dislocation interactions), the deformation is essentially reversible: on unloading the sample, the dislocation disappears by gliding back to the GB sites from which it nucleated, to be re-incorporated into the GB structure. (A movie demonstrating this process is posted at www.msd.anl.gov/im/movies/deform.html.) As the grainsize decreases down to about 20 nm, nucleation of complete dislocations is no longer possible and the dislocation-slip mechanism seizes to be operational [9], in favor of a GB-based deformation mechanism [13,14].

Deformation twinning represents another powerful elementary deformation process associated with the nucleation of partial dislocations from the GBs. In agreement with the
well-known mechanism described in the literature, the mechanism observed in our simulations (see Fig. 2) reveals that the process occurs by the successive emission of partial dislocations onto adjacent (111) planes from the same GB. The stacking sequences of the three types of (111) planes, labeled A, B and C, in five regions of the twinned grain are indicated in the bottom of the figure. These regions are: 1. perfect fcc crystal; 2. intrinsic stacking fault; 3. extrinsic stacking fault; 4. two twin boundaries separated by two (111) planes, and 5. further broadened twin.

As the plastic strain increases, the dislocation concentration in the grain interiors gradually increases, eventually giving rise to various types of dislocation-dislocation interaction processes associated with the glide of extended dislocations on different slip systems but in the same grain. Most of these processes are well-known from extensive deformation studies in single crystals and coarse-grained polycrystals. Among the best known is the formation of Lomer-Cottrell locks, seen in each of the four grains in Fig. 3 below. These locks are formed at the intersection between two extended dislocations on different slip planes, terminated by Shockley partials that attract each other.

In addition to the effects associated with the interactions between gliding dislocations, for larger strains our simulations reveal a variety of unanticipated and intriguing effects arising from the interaction of dislocations with GBs and with deformation twins. Their net combined effect after ~ 12% plastic strain is shown in the snapshot in Fig. 3 for a grain.
Figure 4. Successive snapshots of the vicinity of the triple junction connecting grains 2, 3 and 4, demonstrating the mechanism by which the new grain A in Fig. 2 was formed. (a) $\varepsilon = 6.09\%$ plastic strain; (b) 6.19%; (c) 6.76% and (d) 8.71%. Four distinct processes labeled (1) – (4) are revealed; these are described in the text.

The most intriguing process in Fig. 3 involves the nucleation of the new grain, A. The sequence of snapshots in Fig. 4 captures in detail the underlying GB and dislocation processes. The nucleation of the new grain starts by the emission of a complete $1/2[110]$ dislocation from the GB between grains 3 and 4 (process (1) in (a)). The complex core structure labeled (1’) was formed by two such dislocations, emitted however onto different slip planes. As seen in (b), this new, virtually immobile complex core structure

size of 45 nm.

The considerable roughness of the initially flat GBs is particularly noticeable (compare with Fig. 1). Moreover, beyond a plastic strain of about $8\%$, a new grain (labeled A) is seen to nucleate at those GBs that had been particularly active during these emission/absorption events.

The roughness of the GB network arises in part from the continuous nucleation and annihilation of dislocations. In addition, Fig. 3 demonstrates a mechanism for what, at first sight, appears to be GB splitting (see the GB between grains 2 and 3). However, detailed analysis reveals that the ordered deformation substructure in grain 2 arises from the emission of a series of extrinsic stacking faults from the GB, by the nucleation of double-Shockley partials (compare with region (3) in Fig 2). Interestingly, these cores are aligned in a straight line, presumably due to elastic interactions among them, and eventually leading to the formation of a new dislocation boundary.

This nucleation of double-Shockley partials represents another elementary deformation process associated with the nucleation of dislocations from the GBs. It is known, however, that this process requires very large local stresses.

The sequence of snapshots in Fig. 4 captures in detail the underlying GB and dislocation processes. The nucleation of the new grain starts by the emission of a complete $1/2[110]$ dislocation from the GB between grains 3 and 4 (process (1) in (a)). The complex core structure labeled (1’) was formed by two such dislocations, emitted however onto different slip planes. As seen in (b), this new, virtually immobile complex core structure
subsequently begins to continuously emit partials, producing a growing twin lamella by "partial-dislocation breakaway"\textsuperscript{17}, process (2). This lamella grows further by absorbing additional 1/2[110] dislocations emitted from the same GB (process (1) in (b)), leading to its increased size in (c). As also seen from (c), the development of the new grain also involves the emission of another twin lamella (labeled (3)) together with extrinsic stacking faults terminated by double-Shockley partials (labeled (4); see also process 7 in Fig. 1). Finally, the rather complex deformation substructure thus formed subsequently coalesces to form the final grain A in (d) and Fig. 3.

In coarse-grained Al, the propensity for deformation twinning observed in Figs. 1 - 4 would be very surprising, given that Al has a rather high stacking-fault energy. However, it is also known that the relationship between the stacking-fault energy and the occurrence of deformation twinning is rather indirect and, in fact, entirely unexplored in nanocrystalline materials\textsuperscript{18}. Our observations strongly suggest that a reexamination of the basic models for twinning, with particular emphasis on nanocrystalline grain size, may be timely.

4 Conclusions

In large part, our simulations cover entirely new, experimentally as of yet unexplored ground, not only as far as the small grain size is concerned but also in other important aspects. By contrast with typical in-situ transmission-electron-microscopy experiments, the insights gained from Figs. 3 and 4 capture inherently bulk deformation behavior, unencumbered by any mechanical stresses and surface effects that inadvertently affect the observations in thin-film specimens. Also, even for coarse-grained materials it would be extremely difficult to extract from experiments, in the type of dynamical and atomistic detail available from Figs. 3 and 4, the underlying deformation mechanisms.

The above simulations of well-characterized, albeit highly idealized model systems illustrate the exciting new opportunities offered by large-scale MD simulations towards unraveling the complex interplay between the dislocation and GB processes in polycrystalline materials. However, it is equally important to be aware of the fundamental limitations inherent to this approach. Apart from being limited to relatively small systems, by their very nature MD simulations are restricted to very high stresses and strain rates, many orders of magnitude higher than experimental rates. For example, a strain of 1% occurring in 10 ns of simulation time corresponds to a strain rate of 106 s\textsuperscript{-1}.

Concerning the present simulations, in spite of the large tensile stresses (of 2.3 GPa) and strain rates (∼ 10\textsuperscript{7}s\textsuperscript{-1}) under which deformation was observed, typical dislocation-glide velocities in our simulation (of about 500 m/s) are well below the sound velocity (of about 3664 m/s in the [100] direction obtained by simulation, compared to the experimental value of 3050 m/s). More importantly, however, the resolved shear stresses, of the order of 1 GPa on the (111) slip planes of our system, are well below the theoretical shear strength, σ\textsubscript{th}, for Al. Frenkel’s perfect-crystal shear model gives an estimate of σ\textsubscript{th} = G/2π; with G = 32.5 GPa for Al at 0K (to which our potential was fitted\textsuperscript{19}), we obtain a value of σ\textsubscript{th} = 5.18 GPa at 0K, and slightly lower at room temperature. This value is in good agreement with recent nanoindentation experiments on Al thin films which yielded values of σ\textsubscript{th} in the range of 4.2–4.5 GPa\textsuperscript{19,20}, leading the authors to conclude that the Frenkel formula is well satisfied for Al.
In summary, our simulations provide atomic-level insights into dislocation - GB interaction processes in nanocrystalline materials never before seen in either experiments or simulations, and with a degree of mechanistic detail not presently possible by experiments. They thus offer a glimpse of how, in the near future, such simulations may revolutionize our understanding of plastic deformation processes in heavily deformed materials. In particular, it should soon be possible to elucidate the physical mechanisms controlling technologically important processes, such as superplastic forming. Moreover, the ability of such simulations to completely characterize the highly inhomogeneous state of internal stress will undoubtedly spur the development of better materials-physics based deformation models.

Acknowledgments

V.Y., D.W. and S.R.P. are supported by the US Department of Energy, BES-Materials Science under contract W-31-109-Eng-38. The computations were performed with a grant of computer time provided by the Research Centre Jülich. V.Y. is also grateful for support from the DOE/BES-MS Computational Materials Science Network (CMSN). We benefited from useful discussions with A. K. Mukherjee (University of California).

References

Metallic Glasses and Glass Forming Melts under External Forces or Reduced Dimension: Molecular Dynamics Simulations

K. Brinkmann, K.-B. Küchemann, and H. Teichler

Institut f. Materialphysik
Universität Göttingen, 37073 Göttingen, Germany
E-mail: teichler@umpa06.gwdg.de

Results are reported from molecular dynamics computer simulations for metallic glasses and glass-forming melts under external shear deformations or reduced dimensions. The simulations provide rather unique possibilities to study on the atomic level modifications of the amorphous state of matter under external perturbations. Modelling constant-shear-stress deformations makes visible details of the atomistic mechanisms of plastic deformations of the glasses. Constant-strain-rate processes reveal a deformation induced change of the amorphous structure into a state of easy deformability. This latter state here is ascribed to the interior of the shear bands observed experimentally under ambient temperatures. From the simulations, its properties are characterised in some detail. Local changes of the amorphous state are also studied in the second part of the project, which is concerned with thin films of glass-forming melts and glasses. In particular it is described that for simulated Ni$_{0.2}$Zr$_{0.8}$- and Ni$_{0.81}$B$_{0.19}$-films a chemical demixing takes place near the film surface while earlier studies of mid-concentration Ni$_{0.5}$Zr$_{0.5}$-films revealed in a nm-layer near the surface a significant change of the atomic dynamics by two orders of magnitude without change of the chemical composition.

1 Introduction

Metallic glasses exhibit technologically interesting combinations of properties such as good corrosion resistance, high strength and stiffness at ambient temperatures and high strain-rate superplasticity and workability at elevated temperatures. With regard to the presumed innovative technological perspectives, in particular of the so called ‘bulk metallic glasses’ developed in the last ten years, strong efforts are made worldwide to understand in detail the atomic mechanisms governing the behaviour of this class of materials in order to enlighten the fundamentals for design and control of their properties.

Regarding recent advantages in computing technology — parallel computers on the hardware side, robust algorithms on the software side, and significant improvements in reliably modelling realistic systems — computational approaches play an increasing role in this field, as they can provide more complete and detailed data than often experiments can do for such complex and highly non-linear systems. In accordance with this trend, the present project uses large-scale computer simulations to gain information on the atomic level about the response of glassy materials and their melts to either mechanical external forces or to inhomogeneities like surfaces or interfaces.

In case of mechanical forces, the absence of long range order and of translation symmetry for the glassy state makes inapplicable the standard models of plastic response developed for crystalline solids. Important contributions to modelling the plastic behaviour of glasses have been provided in the past, e.g., by Spaepen$^1$, who described the plastic deformation as arising from single atom jumps, by Argon$^2$, where the plastic response is
viewed to be due to reorientation of clusters containing a larger number of atoms, and Tomida and Egami\textsuperscript{3}, promoting the bond-exchange mechanism. But there remain fundamental questions, particularly regarding the relationship of these models to the present successful approaches in describing diffusive matter transport and relaxation dynamics of metallic glasses (for some details see, e.g.,\textsuperscript{4}).

The report communicates results from molecular dynamics (MD) simulations for a three dimensional model amorphous solid concerning its response against shear deformations (for further details, see\textsuperscript{5}). The study is aimed at elucidating to which extend the plastic response is carried by the original, relaxed amorphous state and to which extend state changes take place under deformations towards a mobile configuration of the glass. Local state changes of related type are the topic of the second part of the project, which is concerned with effects of spatial inhomogeneities, like free surfaces, on the local state in highly viscous liquids and glasses.

2 Dynamics of Metallic Glasses under Shear Deformations

2.1 Simulations

Modelling of the behaviour of a glassy system under shear was carried out for a system adapted to Ni\textsubscript{0.5}Zr\textsubscript{0.5}. The latter is well known as a glass forming metallic system from experiment and MD simulations. Details of the applied model can be found in the literature\textsuperscript{6}.

In the present simulations a medium size system was considered containing an ensemble of 5184 atoms subjected to periodic boundary conditions. The MD modelling was carried out with the TABB code\textsuperscript{7}, a simulation program optimised for use on parallel machines with shared memory conditions. It makes use of dynamical domain decomposition with distribution of the atomic objects to the domains such that communications take place between adjacent domains only, where each domain means one node on the hardware side. The distribution of atoms takes care of equal computation load for each node and can be refreshed with preselected frequency.

Two different deformation processes have been considered in the simulations: deformations under constant-shear-stress and deformations under constant-shear-strain rate. Constant-shear-strain simulations are carried out at three deformation rates ($\dot{\epsilon} = 0.013 \, \text{ns}^{-1}$, $0.035 \, \text{ns}^{-1}$, and $0.13 \, \text{ns}^{-1}$) each at two temperatures (1000 K and 810 K). Deformation under constant-shear-stress was modelled for 6 stress values (ranging from 200 to 1200 MPa) at the two mentioned temperatures. In order to study the macroscopic dynamics of the system under realistic loading conditions, the simulations covered time windows of 40 ns, each corresponding to $2 \cdot 10^7$ MD integration steps of the system with its more than 15,000 degrees of freedom.

Additional simulation runs were used for preparing well-relaxed initial configurations of the system, including long-time runs over more than 100 ns, for verification whether stationary dynamics already are reached during the processes, for analysing the behaviour under time-dependent loading conditions, and for studying the relaxation and recovery behaviour after deformations.
2.2 Results of Constant-Shear-Stress Simulations

For demonstrating our observations during constant-shear-stress simulations, Fig. 1a displays the results of loading the system with 1150 MPa at 810 K. Displayed is the shear strain in the system under constant stress as function of time, the creep curve. The individual events of the plastic shear deformation are visible in the creep curve by peaks and jumps. Comparison of this creep curve with the mean square displacements of Zr-atoms (Fig. 1b) reveals a strong correlation between the macroscopic (shear) strain and the displacements of the Zr atoms, which form the rigid matrix in the considered binary amorphous structure.

Apparentely there are two modes of deformation: jump-like transitions denoted in Fig. 1a and 1b by A to C, and a continuous creep-like behaviour between these events. For the jump C, Fig. 1d presents the atoms that carry out significant displacements during this jump. Obviously, the atoms taking place in this jump are heterogeneously distributed in space and form a loose, local cluster, a ‘mobile spot’. For comparison, Fig. 1c presents...
the atoms that carry out significant movements in the time interval between the jumps A and B. Their distribution is rather diluted in space. A closer analysis shows that the underlying atomic dynamics in this latter creep-like regime can be interpreted as a sequence of individual chain processes, during each of which a chain of atoms, embedded into a more rigid matrix of host material, collectively carries out a transition towards a more favourable configuration. There is evidence that the mobile spots mean local avalanches consisting of a number of such chain processes, similar as discussed by Teichler as local relaxation processes of the glassy state during aging.

In addition to these descriptive results, the simulated creep curves can be used to estimate the activation volume for creep under constant stress. According to Spaepen and Argon, under stationary creep there holds the relationship between strain rate \( \dot{\varepsilon} \), shear stress \( \sigma \), and temperature \( T \), with activation volume \( \Omega \), namely

\[
\dot{\varepsilon} \sim \sinh\left(\frac{\sigma \Omega}{k_B T}\right)
\]

Analysing simulated strain rates obtained under varying shear stress and temperature yields an activation volume of \( 7 - 8 \cdot 10^{-29} \text{ m}^3 \), corresponding to approximately 5 atomic volumes, in good agreement with experimental estimates.

### 2.3 Results of Constant-Shear-Strain Rate Simulations

As a typical example, Fig. 2a displays the stress-strain curve obtained from simulations under constant-strain-rate conditions at nominal 1000 K with a deformation rate of 13 % per ns. The stress-strain curve displays a well-developed yield point after initial linear, elastic increase and then turns into a stationary mode. Our results are qualitatively and quantitatively in good agreement with very recent experimental findings at massive glasses under constant-strain-rate deformations.

The observations have to be considered as a clear indication for a strain induced transition of the glass from its initial rigid configuration into a state of easy deformation under constant load. The transition clearly is demonstrated by Fig. 2b. There simulated constant-stress deformation curves are shown of two different pre-deformation states of the system, one from the initial linear-elastic regime and one from the easy-deformation regime beyond the yield point. While the former remains stable under the applied load of 664 MPa,
the latter shows marked creep under the same conditions and reflects the transition of the system into the deformation-induced easy-deformation state. Similar observations have been obtained at lower strain rates as well as at 810 K.

Our computer studies allow to characterise on the atomic level the properties of the easy-deformation state. Compared to the initial relaxed glassy state, the easy-deformation state shows much higher atomic dynamics, visible in the decay time of the structural $\alpha$-decay, an increased inner energy of the system (see Fig. 3a), and an increased Wendt-Abraham parameter. The latter means a compact way to measure changes in the radial distribution (comp. Fig. 3b) of atoms and thus in the structure of the system. The observed increased parameters are characteristic for undercooled melts at higher temperatures. Hence, although the atomic vibrations in the deformed system correspond to the temperature of 1000 K, the mentioned global parameters refer to a structure of the system at higher temperature.

2.4 Discussion of the Shear Deformation Results

For metallic glasses it is known that deformation under loading takes place in two different ways, depending on temperature and strength of load: at high temperatures and low loads homogeneous deformations are found, at lower temperatures and higher loads deformation shows a heterogeneous behaviour. The heterogeneous deformation mode is characterised by the formation of planar shear bands of easily deformable material.

Our simulations are carried out well in the temperature regime for which homogeneity is expected. We observe, nevertheless, two distinct modes of deformation in our model system. This fact deserves some discussion. It seems to be related to the large stress and strain rates accessible to computer simulations. According to our impression, the critical quantity is the ratio of energy increase in the system due to plastic deformation compared to the energy dissipation through the structure. The latter depends on the state of the system and should be larger in high-energy configurations. Steady-state deformation demands that both energy contributions are outbalanced. A further increase of the deformation rate
by additional loading leads to an increase of energy in the system and thus automatically turns the system into a state of reduced stiffness but larger energy dissipation. It may be a question of details whether for the larger load a new point of stationarity is obtained or whether an instability takes place.

According to this picture, the easy-deformation state found in the simulated constant-strain-rate processes corresponds to a high-strain-rate structure, that means to the interior of a shear band rather than to the experimental state of homogeneous deformation. The experimentally observed heterogeneous deformation thus means a local transition into the here described ‘flow state’. As indicated in the preceding section, the computer simulations allow to identify the microscopic properties of this flow-state of glassy matter and to distinguish it from the relaxed state of the glass in a much more detailed way than it can be expected from experiments.

3 Glasses and Glass Forming Melts at Reduced Dimension

3.1 Simulated Ni_{0.5}Zr_{0.5}-Films with Free Surfaces

In a preceding study\textsuperscript{10, 11} we had shown by MD simulations that a change of state takes place close to free surfaces of amorphous and liquid Ni_{0.5}Zr_{0.5}-films. This change is visible in the atomic dynamics, e.g., in the local diffusion coefficients or local ‘hopping rates’ (the latter being measured in terms of the life time of the nearest-neighbour relationship). While in the interior of the film the bulk values of the diffusion coefficients are found, the diffusion increases continuously by more than two orders of magnitude in a surface layer of about 1 nm thickness when approaching the film surface. This behaviour has some resemblance to the phenomenon of premelting known at (110) surfaces of some fcc metals. There is, however, the particular difference that in case of premelting an additional inner interface is found while in the amorphous film — and in simulated films of highly viscous liquids — the change takes place continuously. According to these findings, significant changes of the properties of the glassy state are expected in case of reduced dimension, that means once the extension of the glassy material in one dimension is comparable or less than twice the surface layer thickness. It is the objective of the present studies to investigate which changes of the sketched phenomena are to be expected once the chemical composition of the film or the chemical nature of the constituents are modified.

3.2 Simulated Ni_{0.2}Zr_{0.8} and Ni_{0.81}B_{0.19}-Films with Free Surfaces

MD simulations for Ni_{0.2}Zr_{0.8} and Ni_{0.81}B_{0.19}-films revealed a significantly different behaviour of these materials compared to Ni_{0.5}Zr_{0.5}. Both materials are characterized by the fact that there is a fraction of about 20 % of smaller atoms dispersed in an amorphous matrix of larger atoms. For both systems, simulated annealing of a film of finite thickness revealed similar features: During annealing, in the interior of the film there takes place an enrichment of the smaller atomic species while at the surface a pure layer of the larger atoms is formed. A particular example of this process is shown in Fig. 4. — The corresponding MD simulations on the Cray T3E of the NIC at Jülich have been carried out by use of the TABB-code addressed in section 2.1.

For the Ni-B system, the freshly formed surface layer of the larger Ni-atoms was removed during the simulations a couple of times, yielding a series of systems with slightly
decreasing Ni- and increasing B-content. Under annealing, in each case the freshly formed mixed Ni-B-surface relaxed into a pure Ni-surface with further increase of the B-content in the interior. These observations indicate that — apparently — formation of the larger atoms surface layer is a particular property of the surface structure for the considered glassy states. By formation of this structure, the density of energetically favourable NiB-bonds (resp. ZrNi-bonds in the Ni-Zr system) is increased on account of a loss of entropy of mixing in the surface sheet.

It is to be expected that in the Ni-Zr-system the demixing tendency near the surface decreases with decreasing Zr-concentration of the melt, leading to a homogeneous chemical composition for mid-concentration films. The details of this process are, however, not yet known.

### 3.3 Sandwich Structures of Amorphous NiZr and Crystalline Zr

While the preceding section considers glassy and viscous liquid films with free surfaces, the present section reports some results obtained for viscous liquid Ni-Zr films between crystalline Zr-walls. During annealing of such sandwich structures, by either growth of the crystalline Zr-structure or partial dissolution of the crystal under growth of the melt, the composition of the melt is adjusted to the concentration of the liquidus curve describing the thermodynamic equilibrium between the highly viscous liquid and the Zr-crystal at the considered annealing temperature. — Fig. 5 demonstrates for a Ni$_{0.2}$Zr$_{0.8}$-melt between crystalline Zr-walls the broadening of the crystalline layers under Ni-enrichment of the melt obtained from MD simulations for nominally 1250 K.

Due to the temperature dependence of the liquidus concentration, a temperature change demands a change of the equilibrium composition of the melt and hence crystal growth or dissolution. With regard to this, a series of relaxations of the system under successive cooling was carried out, yielding a stepwise increase of the Ni-content in the melt. On the other hand, a series of relaxations under stepwise increase of the temperature gave a series of stepwise decrease of the Ni-content in the melt. It is to be expected that the equilibrium liquidus concentration just lies between the Ni-concentrations from cooling and from heating and that the equilibrium concentration thus can be estimated from these simulations. The here sketched simulations open the way to study in detail on the atomistic level the structure and further properties of the crystal-melt interface. In the following we shall exploit, however, a different aspect of these simulations, which is related to the question of stability of the melt against concentration fluctuations.
3.4 Discussion of the Simulations of Glass-Forming Melts with Reduced Dimension

In Section 3.3 we described MD simulations for a sandwich structure of liquid Ni\(_{1-x}\)Zr\(_x\) and crystalline Zr. The resulting continuous shift of the equilibrium Ni-concentration with temperature, as sketched in Fig. 6, indicates for Gibbs thermodynamic potential of the melt in the considered concentration regime a convex curvature with composition. According to this finding the NiZr-melt in this regime will be stable against composition fluctuations without any tendency for demixing.

With this knowledge we now shall consider the simulations described in Section 3.2. There we found near the surface, e.g., of the Ni\(_{0,2}\)Zr\(_{0,8}\)-film a local change in composition yielding a pure Zr layer on the film accompanied by an enrichment in the interior. Regarding the just mentioned results of Section 3.3, a volume instability promoting chemical demixing not seems plausible. Thus, there remains the alternative that the driving force
for the observed surface-demixing lies in the surface itself. Accordingly, the concentration profile near the surface should reflect the way the system balances the competition between the preferred surface structure and the volume stability of the melt. — The observations for the Ni-B-melts display a similar behaviour and thus give support to the assumption that these findings are not limited to the NiZr-system but reflect a more general tendency.

Combining the here considered phenomena with the results discussed in Section 3.1 for mid-concentration melts leads to the conclusion that surfaces of the highly viscous melts may give rise to changes in the amorphous structure, yielding a change either in the mobility and atomic dynamics near the surface or in the chemical composition and dynamics. For the low-concentration case, the observed demixing reflects that the composition couples to the surface in this case. In the mid-concentration case, the increased atomic mobility without chemical demixing indicates, however, the existence of an additional hidden parameter in the amorphous structure which governs this mobility and which couples to the surface\textsuperscript{10,11}.

4 Concluding Remarks

The above described MD simulations on highly viscous, glass-forming melts are concerned, in essence, with the question of modulations of the amorphous viscous state. The first part concerns changes of the state under mechanical shear deformations. Regarding this, the provided results demonstrate in case of the creep-curve, Section 2.2, the relaxation dynamics of the equilibrated viscous state under loading. Beyond this, the constant-strain deformations in Section 2.3 describe a change of the amorphous state towards a state of easy deformability. Analysis regarding its energy and Wendt-Abraham parameter revealed some of the properties of this new state, indicating that its structure is similar to that of the equilibrium viscous states at higher temperatures.

The MD simulations\textsuperscript{10,11} had shown that a change of state takes place close to free surfaces of amorphous and liquid Ni\textsubscript{0.5}Zr\textsubscript{0.5}-films. This change was visible in the atomic dynamics, e.g., in the local diffusion coefficients and local ‘hopping rates’. This change of state was described in terms of a hidden parameter of the amorphous structure\textsuperscript{10,11}. The present MD simulations show that the surface modifies the chemical composition in case of Ni-poor Ni-Zr-films (and Ni\textsubscript{0.81}B\textsubscript{0.19} films). They give a clear picture of the chemical demixing near surfaces and makes obvious that the phenomena in the mid-concentration Ni-Zr-film are rather different from the case of chemical demixing. — Regarding the here sketched modifications of the highly viscous liquid — or glassy — state of matter, computational physics, in particular MD simulations, seem to offer rather unique possibilities to study the phenomena in much more detail than accessible to present experiments. The variety of results may encourage, however, the design of appropriate experimental studies.

Acknowledgements

The authors acknowledge support of this study by the NIC by providing the necessary CPU time on the Cray T3E. Financial support is acknowledged from the DFG in the new SPP "Phasenumwandlungen in mehrkomponentigen Schmelzen" and by the Sonderforschungsbereich 345 der DFG "Festkörper weit weg vom Gleichgewicht".

241
References

Condensed Matter physics deals with fundamental properties of solids and liquids. The solution of problems in this field clearly bears high technological potential but it also provides a substantial intellectual challenge. The formation of structures, the appearance of different types of spatial order and disorder, the occurrence of transitions between various phases of condensed matter, the temporal evolution and the characterization of thermal and quantum fluctuations associated with various states of condensed matter are confronting us with many important unanswered questions. It is not so much the individual phenomenon occurring in a specific substance but rather the systematics within classes of materials and across these classes which require explanation.

In principle all properties of interest in condensed matter physics depend simply on the interaction of the electrons which can be regarded as the glue between the ionic cores. However, the exponential growth of the quantum-mechanical problem with the number of particles so far prohibits a numerical solution of an interacting quantum-system of more than some hundred particles even in simplified model-systems. One must, therefore, resort to effective interactions between atoms in order to treat large systems. Characteristic aspects of the effective interactions between the particles forming the solid or liquid are the typical range of interaction, local degrees of freedom, local and global symmetries of order parameters.

Typical phenomena of condensed matter physics are directly associated with the enormous number of interacting particles involved, leading to the complicated correlations between the particles even when the direct interactions are of short range. This is a main reason for the continuous need of virtually unlimited computer-resources for the detailed investigations of model-systems for condensed matter phenomena, and it is a strong driving-force for the development of highly efficient numerical algorithms.

Computer-simulation in condensed matter problems start at the quantum-mechanical level with density-functional models for metals or semiconductors, the so-called models for first-principles calculations. With these models one may now calculate electronic and structural ground-state properties of large classes of materials with surfaces and layer-structures, including compound-materials. A second class of quantum-models starts from an effective Hamiltonian, e.g. of Hubbard-type, which gives the interactions of the many-body system in simplified form. The main objective for these model-studies is the explanation of correlation effects giving rise to collective phenomena like electronic excitations, metal-insulator transitions or superconductivity. A third set of models defines effective interactions between atoms or molecules, serving as basis for classical simulations with molecular dynamics or Monte Carlo methods. The various levels of large-scale computation in condensed matter physics and the enormous progress achieved in recent years are reflected in the following articles.
Starting from a Hubbard model, the Mott-Hubbard metal-insulator transition is studied in the article of Blümer et al. using the quantum-mechanical dynamical mean field approximation (DMFT), with extensions for realistic materials by combining DMFT with the local-density-functional approximation (LDA). This LDA+DMFT Quantum-Monte-Carlo calculation gives remarkably good results for photo-emission and X-ray absorption associated with the strongly correlated electrons in vanadium-oxide.

The mechanism of high-temperature superconductivity (HTSC) is addressed in the contribution by Dorneich et al. This is a numerical study of an effective bosonic model which describes the low-energy physics of a ‘projected’ SO(5)-theory. The theory unifies the three-dimensional order parameter of the antiferromagnetic phase with the two-dimensional order parameter of the superconducting phase. The large-scale numerical simulations use a technique of stochastic series expansions. The resulting phase diagram looks very similar to the one of real HTSC-cuprates.

Electronic correlations coexisting with dynamical electron-phonon interactions is a central mechanism in physical properties of novel materials like organic polymers, charge-ordered nickelates or colossal magnetoresistance manganites. In the article of Fehske et al. exact diagonalization results for the typical Hamiltonians are presented. As a successful interdisciplinary cooperation the implementation of efficient parallel diagonalization methods for very large matrices also resulted from this project.

A microscopic many-body theory describing the optical and electronic properties of semiconductors and heterostructures is described in the contribution of Meier et al.. The importance of correlations beyond the Hartree-Fock-level for polarization-dependent pump-probe and four-wave mixing experiments have been convincingly demonstrated by these massively parallel computer simulations. Optical properties of semiconductors are of high technological significance in information technology. A very efficient method to calculate optical spectra for complicated geometries such as quantum wires, superlattices in electric and magnetic fields or excitations on rough solid surfaces is presented by Glutsch. The calculations have given convincing agreement with experiments.

The structure of solids and magnetism on an atomic scale is discussed by Hergert et al. on the basis of quantum-mechanical density-functional calculations for electronic properties, accompanied by molecular dynamics simulations for metallic nanostructures. It is demonstrated that a large number of materials properties like surface diffusion barriers, magnetic moments and strain and stress effects can now be reliably calculated for real materials like cobalt deposited on copper and other metals like gold and silver.

Janke et al. study classical spin-glasses as typical representatives of disordered materials, where the possible states of the system are separated by high barriers of the free energy. A multi-overlap algorithm gave substantial new insight concerning the excitations over these barriers, statistically rare events. The method proved to be particularly powerful in the study of the tails of the distribution function.

A related investigation by Binder et al. discusses similarities and differences between glasses and spin-glasses using Monte Carlo and molecular dynamics studies on various model systems. A crucial point is the dramatic slowing down of thermal equilibration, as the glass-temperature is approached from above. Detailed comparisons also with experiments on silica-glass are presented.

Liquid crystals provide a wealth in possible phases and dynamical processes due to the non-spherical symmetry of their molecules. Because of their unusual electric, elastic and
rheological properties they are of high technological interest. As an example, the interface between the nematic and the isotropic phase was studied by Akino et al. via molecular dynamics. An unexpectedly complex spectrum of capillary waves was observed, with significant deviations from capillary wave theory.

Systems of liquid crystals with different molecular shapes like chiral or banana-shape were studied by Memmer in order to unveil the relationship between molecular symmetry and macroscopic phases. A rich polymorphism of phases could be characterized in dependence on temperature, pressure, chirality parameter and molecular structure.

The examples presented here span a whole range of condensed matter physics, from the electronic structure of "hard" materials such as semiconductors and metals to the phase diagrams of "soft matter" - systems, such as liquid crystals. This underlines the significance of large-scale simulations as an increasingly flexible tool which can yield important insights into problems that are relevant both for physics and materials science.
Numerical Investigations of the Glass Transition and Glassy State of Structural Glasses and Spin Glasses

Kurt Binder\(^1\), Jürgen Horbach\(^1\), and Walter Kob\(^2\)

\(^1\) Institut für Physik, Johannes Gutenberg–Universität Mainz
Staudinger Weg 7, 55099 Mainz, Germany
E-mail: kurt.binder@uni-mainz.de, horbach@komath.physik.uni-mainz.de

\(^2\) Laboratoire des Verres, Université Montpellier II
Place E. Bataillon, cc069, 34095 Montpellier, France
E-mail: walter.kob@ldv.univ-montp2.fr

Monte Carlo and molecular dynamics simulations of various models of glass–forming systems are described, aiming at a better understanding of the amorphous state of matter, and to clarify how a liquid transforms into a glass. Our studies range from chemically realistic models of silica (SiO\(_2\)) and SiO\(_2\)–Na\(_2\)O glasses over simpler models like binary Lennard–Jones fluids to abstract spin models such as the so–called Potts glass. The motivation to study such models will be briefly addressed, and the reason will be explained why massive computational efforts on large–scale supercomputer facilities are crucial to make progress in this field. Results will be presented on bulk properties as well as on systems confined to small droplets or thin films.

1 Introduction

While within solid state physics the properties of crystalline materials are rather well understood, the amorphous state of matter still poses significant challenges: in a crystal the atoms form a periodic lattice, while in a glassy material the geometric arrangement of the atoms exhibits only order on short length scales, but no (known) long range order. Thus, while single crystals have a regular external shape with facets meeting under particular angles, glasses have no particular shape, and that is why these materials are called “amorphous”. While in a crystal lattice different directions are usually not equivalent, and physical properties hence exhibit more or less anisotropy, glasses are completely isotropic, just like fluids. This isotropy (and also their homogeneity on mesoscopic and macroscopic scales) is, e.g., very desirable for many optical applications, such as lenses or prisms (e.g., obviously it would be difficult to make eyeglasses from crystals because of their birefringence), although in the following we will not discuss these issues further.

If one analyzes the structure of glass on an atomistic scale (by performing scattering experiments of X–rays or neutrons), one obtains a picture very similar to the corresponding liquid. Therefore it is sometimes said that a glass is a “frozen liquid”, where atoms are no longer allowed to move over larger distances. But what mechanism prevents this motion? And already in the liquid phase one finds a dramatic slowing down if the liquid is cooled down towards the temperature where the glass transition occurs. As an example, Fig. 1 shows the diffusion constants of silicon (Si) and oxygen (O) atoms in molten silica\(^1\): one observes a decrease of these diffusion constants \(D\) (and this corresponds simply to an increase of the structural relaxation times, \(\tau \propto D^{-1}\)) over about 15 decades! Over most of this range, the structure of the fluid changes very little (at least as far as properties as
pair correlation functions between atoms, as accessible via scattering experiments, are concerned), and hence this dramatic slowing down is very puzzling. Fig. 1 also illustrates that it is very difficult to address this problem by means of computer simulations: in a liquid, local motions are rather fast (atomic vibration times are of the order of $10^{-13}$ sec), while near the glass transition $\tau$, the time scale for the structural relaxation, is of the order of $10^2$ sec.

If the fluid is cooled below the glass transition temperature $T_g$, it falls out of equilibrium and stays more or less frozen as a solid, i.e. $\tau$ diverges. A molecular dynamics simulation amounts to solve numerically Newton's equations of motion of classical mechanics for all the interacting atoms in the model system. (Typically one uses a simulation box containing a few thousand atoms and applies periodic boundary conditions to make the system quasi-infinite.) Since the time step $\delta t$ of the integration routine must be distinctly smaller than the shortest characteristic time of the system, $10^{-13}$ sec, for the simulation data in Fig. 1 a time step $\delta t = 1.6 \cdot 10^{-15}$ sec had to be used. Thus one has to make many of these steps if one wants to reach mesoscopic time scales. Furthermore also the calculation of the forces with which the atoms interact needs a lot of computer time due to the long range of the Coulomb interactions between the ions (these forces cannot be truncated and therefore necessitate the use of the so-called Ewald summation technique). Therefore the data of Fig. 1, where one needed to extend the simulation to $2 \cdot 10^{-8}$ sec, i.e. $1.2 \cdot 10^7$ time steps, could only be generated on a high performance supercomputer (the CRAY-T3E of the NIC was used with parallelization of the calculation of the forces). Despite this effort it is obvious that there is a gap between the temperature range at which simulations can be done in equilibrium and the temperature scale accessible to experiments, and at the present time it is obvious that there is a gap between the temperature range at which simulations can be done in equilibrium and the temperature scale accessible to experiments, and at the present
Figure 2. Schematic explanation how quenched disorder by random dilution leads, (a) from a ferromagnet to a spin glass or, (b) from a molecular crystal to a quadrupolar glass. The ferromagnet corresponds to Europium Sulfide, showing only one plane of the face centered cubic lattice, where magnetic exchange interactions occur essentially only between nearest ($J_{nn}$) and next nearest ($J_{nnn} \approx -J_{nn}/2$) neighbors. Random dilution with non–magnetic Strontium leads to the spin glass Eu$_x$Sr$_{1-x}$S (if the concentration $x$ of Europium lies in the range $0.13 < x < 0.5$). Arrows indicate the orientation of the magnetic moments in each case. The long range ordered molecular crystal refers to ortho–hydrogen, showing again one lattice plane of the crystal only. Ellipsoids indicate the orientation of the quadrupole moments of the hydrogen molecules. Dilution with para hydrogen, which has no quadrupole moment, yields a quadrupolar glass.

While in SiO$_2$ really small relaxation times (small means here $\tau \approx 10^{-13} - 10^{-10}$ sec) are only reached at extremely high temperatures, that are not accessible to experiment as Fig. 1 shows, in the so–called “fragile glasses” $\tau$ decreases from $\tau \approx 10^2$ sec at $T_g$ to small values very rapidly, in a temperature interval of the order of 100 K, and hence are well accessible to experiment. These fragile glassformers typically are materials formed from organic molecules, including polymers which are the basis of the ubiquitous plastics materials. Due to the complicated chemical structure of all these molecules, chemically realistic models of these systems pose similar difficulties to simulation as in the case of SiO$_2$. But experiment shows that all these fragile glassformers behave qualitatively very similar. There is a remarkable degree of universality, irrespective of the differences in the chemical structure. This fact suggests to simulate a simpler model, a fluid consisting of two types of atoms denoted symbolically as A and B, interacting with short range pair potentials $U_{AA}(r)$, $U_{AB}(r)$, and $U_{BB}(r)$. These potentials are “freely invented” and chosen such that neither crystallization nor phase separation occurs, at least not on any time scales relevant for simulation. We shall present a few results for this model that have clarified important theoretical questions in Sec. 3.

But also in this model it is not possible to equilibrate the system in the vicinity of $T_g$ or at even lower temperatures. In order to test issues pertaining to such low temperatures,
Figure 3. Static structure factor of quartz glass ($\text{SiO}_2$) at room temperature ($T = 300$ K) plotted vs. wavenumber $q$ ($q$ is related to the scattering angle $\theta$ as $q = (4\pi/\lambda)\sin\theta$, $\lambda$ being the wavelength of the radiation used for the scattering experiment). The full curve is the molecular dynamics simulation of Ref. 1, using the measured neutron scattering lengths of Si and O atoms, and the circles are a neutron scattering experiment.

one has to restrict oneself to even more abstract models, that in a strict sense do not really represent amorphous materials, but rather crystalline solids in which other degrees of freedom exhibit glass–like relaxation. Examples for such systems are the so-called “spin glasses” and “orientational glasses”\(^2,3\). Spin glasses are created by random dilution of a ferromagnet with competing interactions (Fig. 2): disorder plus frustration cause the spins to freeze in random directions\(^2\). Similarly, random dilution of molecular crystals creates orientational glasses: electric dipole or quadrupole moments freeze in random orientations (Fig. 2). We shall also present some very interesting simulation results about such models (Sec. 4), before we summarize and conclude (Sec. 5).

2 Simulation of Molten Quartz and Sodium Silicate Glasses

$\text{SiO}_2$ is suitable for a molecular dynamics simulation, since accurate (pairwise) potentials could be extracted from quantum chemistry methods\(^4\) that faithfully simulate the covalent bonding in this material, and reproduce well many known properties of crystalline $\text{SiO}_2$. While we can equilibrate $\text{SiO}_2$ only at rather high temperatures (e.g. $T = 2750$ K), cooling such melts down to $T = 300$ K with the slowest rates acceptable for a simulation (still as high as $10^{12}$ K/sec!) at constant density yields a structure factor $S(q)$ in very good agreement with scattering experiments (Fig. 3). Here $S(q)$ represents the intensity of the radiation observed with a momentum transfer $\hbar q$. In the comparison of Fig. 3 no adjustable parameters whatever enter, and thus this comparison shows that our simulation can reproduce experiment. But more importantly, the simulations can go beyond experiment in various ways: e.g., for the mixture $\text{Na}_2\text{O}–\text{SiO}_2$ we have measured all 6 partial structure factors $S_{\alpha\beta}(q)$ ($\alpha, \beta = \text{Na}, \text{O}, \text{Si}$), which are not accessible experimentally\(^6\), but are needed to develop a detailed picture of the atomistic structure. Another relevant infor-
Figure 4. Snapshot of a SiO$_2$ cluster with 432 ions at two different temperatures. The red and yellow spheres correspond to silicon and oxygen atoms. Covalent bonds between them are drawn as little sticks if their distance does not exceed 2.3 rA. From Roder et al.$^7$.

formation is nature of the local ordering at free surfaces of molten (or frozen) SiO$_2$ against vacuum. We can study this problem by simulating a finite cluster with a free surface (Fig. 4). Note that in the ideal random SiO$_2$ network there are chemical rules concerning the nearest neighbor coordination of each atom: each Si atom wants to be linked by covalent bonds to 4 oxygen atoms, and in turn each oxygen atom wants to have two Si neighbors. Thus, the glass structure can be viewed as a network of distorted tetrahedra (with Si in the center of the tetrahedra, oxygen atoms at their corners), linked together at the tetrahedra corners with almost random angles. In such a structure, there are almost no “dangling bonds”, i.e. oxygen atoms that have a single silicon neighbor. However, while such a structure seems natural in the bulk, it is not clear how this structure will change in the vicinity of a free surface against vacuum (we disregard here the problem encountered in practice that the vacuum is not perfect and water molecules or other molecules may supply ions to saturate the dangling bonds at the surface). The snapshots (Fig. 4) show that for very high temperatures indeed many “dangling bonds” are found, while for the still rather high temperature $T = 2750$ K the dangling bonds are already rather scarce, and a closer analysis shows that the network adjusts by an enhanced density of oxygen atoms in the surface region, to maintain a perfect coordination.

Note that the coordination is not truly perfect in the bulk phase of molten SiO$_2$: there does occur a small fraction of Si atoms with five oxygen atoms as nearest neighbors and a corresponding fraction with three neighbors$^1$. The motion of these coordination defects and associated breaking of covalent bonds is in fact responsible, that a network as shown in Fig. 4 is not just an elastic solid but still a fluid! In fact, the self–diffusion constant of oxygen atoms is inversely proportional to the average “lifetime” of the covalent Si–O bonds$^1$.


The computer simulation of SiO$_2$ and silicate melts is very expensive regarding computer time because of the long range of the Coulomb interaction between the ions (the potential energy decays with distance $r$ as $1/r$). In molecular fluids like ortho-terphenyl or
in polymer melts, intermolecular forces are much shorter ranged, but the chemical structure of the molecules itself presents a complication. Qualitatively, however, one still can capture the essential features of the glass transition even if one disregards this complication and assumes spherical particles interacting with simple Lennard–Jones (LJ) potentials,  

\[ V_{\alpha\beta} = \epsilon_{\alpha\beta} \left[ \left( \frac{\sigma_{\alpha\beta}}{r} \right)^{12} - \left( \frac{\sigma_{\alpha\beta}}{r} \right)^6 \right], \]

where the parameters \( \epsilon_{\alpha\beta}, \sigma_{\alpha\beta} \) describe the strength and the range of this pairwise interaction between particles of type \( \alpha \) and \( \beta \). Here we have anticipated that it is necessary to work with a binary mixture of two kinds of particles A and B (so three types of pairs \( \alpha\beta = AA, AB, \) and \( BB \) occur): a single component Lennard–Jones fluid would be a good model of a rare gas fluid, and rare gases crystallize easily, so they are no glassformers! However, tuning the parameters \( \{ \epsilon_{\alpha\beta}, \sigma_{\alpha\beta} \} \) for the binary mixture suitably, an excellent model for a glass–forming fluid is obtained\(^8\), well suited for a test of various theoretical concepts about the glass transition.

Some of these concepts are sketched in Fig. 5: According to the entropy theory of the glass transition\(^9,10\), glassy freezing is caused by the vanishing of the configurational entropy of the supercooled fluid at the temperature \( T_0 \) \(< T_g \), which should lead to a divergence of the viscosity according to the Vogel–Fulcher law,  

\[ \log \eta(T) \propto (T - T_0)^{-1}. \]

Although experimental data for the entropy\(^11\) and viscosity\(^12\) are often compatible with such an interpretation, it is very speculative because always \( T_0 \) is distinctly lower than \( T_g \) (e.g. \( T_g - T_0 \approx 50 \) K), and the mathematical approximations involved\(^9,10\) are rather crude and questionable.

Figure 5. Schematic plot of the viscosity \( \eta(T) \) of a fluid vs. inverse temperature \( 1/T \). Characteristic temperatures are indicated: melting temperature \( T_m \), critical temperature \( T_c \) of mode coupling theory (MCT), glass transition temperature \( T_g \) — defined empirically via \( \eta(T = T_g) = 10^{13} \) Poise — and the Vogel–Fulcher temperature \( T_0 \), respectively. The lower left part shows the schematic behavior of the Fourier transform of density correlations \( \Phi_q(t) \), the so–called “non–ergodicity parameter”. The upper left part shows the normalized entropy difference \( \Delta S/S_m \) \((S_m \) is the entropy difference at \( T_m \)) between fluid and crystal as a function of temperature: often these data are compatible with a linear extrapolation according to which \( \Delta S \) vanishes at \( T_0 \), giving rise to the so–called “Kauzmann paradox”\(^9\)\(^–\)\(^11\).
A complementary concept is the mode coupling theory (MCT) of the glass transition\textsuperscript{13}, which tries to describe the dynamics of the supercooled fluid on the atomic scale. I.e., it focuses on the time–displaced density–density correlation function $\phi_q(t)$ for wave-vector $q$ (Fig. 5). According to the idealized version of MCT, $\phi_q(t)$ decays quickly to zero at high temperatures, while there exists a critical temperature $T_c$ ($T_c > T_g$) where the behavior changes drastically: a plateau develops in $\phi_q(t)$ for $T \geq T_c$, and the “lifetime” of this plateau diverges as $T \to T_c$. For $T < T_c$, the system gets “stuck”, $\phi_q(t)$ does no longer decay to zero any more, but relaxes only to the “nonergodicity parameter” $f_c(q)$. The physical explanation for this behavior is the “cage effect”: in a dense fluid, the motion of each atom is hindered by the atoms in its neighborhood, and as the temperature gets lower, it is no longer possible for an atom to escape from this “cage”. However, this theory would imply that the structural relaxation time (and hence the viscosity) would diverge according to a power law as $T \to T_c$ which disagrees with experiment. In the more refined “extended MCT” this divergence at $T_c$ is replaced by a smooth crossover from a power law to a simple Arrhenius behavior, and this is perhaps compatible with most experiments.

Now the “time window” corresponding to $10^{-1} \leq \eta(T) \leq 10^{2}$ Poise, where idealized MCT should work, is well accessible to molecular dynamics simulations of simple fluids as the binary LJ mixture introduced above. Thus, a stringent test of this theory could indeed be performed (Fig. 6)\textsuperscript{14}. Since the static structure factor measured in the simulation is the only input needed for the theory, one again notes excellent agreement without adjusting any fit parameters. A prediction of the theory, namely that $f_c(q)$ is independent of the nature of the underlying microscopic dynamics could be tested directly by comparing a stochastic dynamics simulation (where a random force and a damping force is added to the equations of motion) to the standard (Newtonian) molecular dynamics simulation\textsuperscript{14}.
This model is also suitable to address a question that currently finds much attention\(^\text{15}\), namely the question how a confinement of the liquid in cylindrical pores or in slit pores affects the glass transition. Experimentally this problem is very difficult, since one does not know precisely the interactions between the liquid and the walls of the pore and also the density of the fluid in the pore is hard to control. Simulations have a distinct advantage here: we can create a pore\(^\text{16}\) by introducing a virtual (cylindrical) wall in a bulk system, and freezing all particles outside of this wall at their positions at this instant of time. In this way, the interactions between particles in the fluid, and the particle–wall interactions are very well controlled, the geometric structure of the confined fluid remains almost unaffected, and the effect of confinement on the dynamics can be nicely analyzed! This is an example that shows that with simulations one can prepare more ideal conditions for an “experiment” than it would be possible in reality, and thus precise conclusions can be drawn, while in experiments it is rather ambiguous how the different effects can be disentangled\(^\text{15}\). In this way it was possible to demonstrate that near the confining wall the cage effect is enhanced due to the atomic roughness of this wall\(^\text{16}\), i.e. the relaxation slows down. As one approaches the glass transition, the length scale over which the atoms are affected by this confining wall grows\(^\text{17}\).

4 The 10–State Mean–Field Potts Glass: An Abstract Model for the Structural Glass Transition?

For the model of SiO\(_2\) in Sec. 2, the lowest temperature that could be simulated in thermal equilibrium was \(T = 2750\) K, while the experimental glass transition temperature \(T_g\), i.e. the temperature of interest, is around 1450 K. For the Lennard–Jones mixtures one can
come somewhat closer to the glass transition temperature, namely almost $T/T_g \approx 1.3$. But even for this system it is not yet possible to simulate below $T_c$, and on the scale of the viscosity the 10 orders of magnitude from $\eta \approx 10^3$ Poise to $\eta(T = T_g) = 10^{13}$ Poise cannot yet be explored.

In order to learn something about the behavior at and below $T_g$, one needs to resort to even simpler but rather abstract models of the spin glass type. However, the standard spin glasses (Fig. 2) are not good models of the structural glass transition: they show a gradual onset of glass behavior, i.e. for $T < T_g$ the spin glass order parameter $q_{EA} = \langle S_i^2 \rangle_{av}$ increases gradually, according to a power law with a critical exponent $\beta$, $q_{EA} \propto (1 - T/T_g)^{\beta/2}$. From Fig. 5, however, it is clear that the “nonergodicity parameter” — which plays the role of an order parameter of the glass transition — becomes nonzero via a discontinuous jump.

However, such a case can be realized with spin glass models, too, for instance, the 10–state mean–field Potts glass. In this model, each spin $S_i$ can be in one of 10 discrete states, and the energy $J_{ij}$ is only nonzero if a pair of spins $(S_i, S_j)$ is in the same state, while otherwise $J_{ij} \equiv 0$. A mean–field model is defined by introducing such an energy function between any pair of spins in the system, independent of their distance. The model shows a glass transition by choosing a Gaussian distribution of interactions $J_{ij}$ (with negative mean). Interestingly, this model shows two transitions in the thermodynamic limit (number of spins $N \to \infty$): at $T_D$ a dynamical transition occurs from ergodic to non–ergodic behavior, of the type described by MCT. The time–displaced spin–spin autocorrelation function shows the same behavior as $\Phi_q(t)$ in Fig. 5! In addition, there is a static transition at $T_0 < T_D$, where $q_{EA}$ jumps discontinuously from zero to a nonzero value, and the (configurational) entropy has a kink (if one would extrapolate the entropy of the high temperature phase to $T < T_0$, one indeed would also find a “Kauzmann temperature” $T_K$ below which the entropy would become negative, but this has no physical significance, the “entropy catastrophe” at $T_K$ is preempted by the static transition at $T > T_K$).

A Monte Carlo study of this model is of interest, because computer simulations always deal with finite systems, where phase transitions are shifted in temperature and smeared out (i.e., critical singularities and jump discontinuities cannot straightforwardly be observed, rather one always observes a gradual behavior). Fig. 7 shows the relaxation time and Fig. 8 the entropy of this model, both plotted versus inverse temperature. In the case of the entropy, exact results for $N \to \infty$ are available and included for comparison.

Indeed this model for finite $N$ exhibits a gradual crossover from a power law divergence of the relaxation time (for $T > T_D$) to an exponential divergence (of Arrhenius type) for $T < T_D$. The activation energy $E_A$ increases with increasing $N$. (For this mean–field model the barriers between the different valleys in configuration space have a finite height only if $N$ is finite.) Both static and dynamic properties of this model exhibit considerable finite size effects, which are still understood only partially. At this point, we do not yet know whether analogous size effects are to be expected for the more realistic models of the structural glass transition, considered in the previous sections.

5 Concluding Remarks

In this report, we have demonstrated that large scale computer simulations can indeed give a lot of insight into the challenging problem of glasses and the glass transition, even though
not all questions of interest can be answered yet. While chemically realistic models elucidate the understanding of particular experimental data and complement them in certain directions (e.g. SiO$_2$ at $T \geq 2750$ K, where no measurements can be done), there is still a considerable gap between the range of time scales accessible in experiment and in the simulations. While techniques such as “parallel tempering”$^{23}$ can help to win about a decade in time or so, Fig. 1 demonstrates that there is need for more fundamental ideas to bridge this gap in the time-scale.

A major achievement of the simulation of the binary Lennard–Jones mixture is the extensive test provided of the mode coupling theory. The existing experiments are obscured by many complicating details, because additional molecular degrees of freedom provide additional structure in the relaxation and response functions, and are much more difficult to interpret. In addition, the simulations have begun to yield insight on the fascinating interplay of confinement and surface effects on the glass transition.

Although the simulation of spin glass models is an active topic since 25 years, our recent study$^{21}$ reveals that there are still many open problems, and in particular finite size effects are not yet well understood. Since in many areas of computer simulations the understanding of finite size effects has been crucial, clearly more work in this field is still necessary.

Acknowledgments

We thank C. Brangian, T. Gleim, M. Nauroth, A. Roder, P. Scheidler, T. Stühn, K. Vollmayr, and R. Yamamoto for a fruitful collaboration on the problems described here. We are
grateful to the Deutsche Forschungsgemeinschaft (DFG-SFB 262/D1), the Bundesministerium für Forschung und Technologie (BMBF, grants No. 03N8008C and 03N6015) and SCHOTT Glas for financial support. We are particularly grateful to the NIC Jülich (previously: HLRZ) for generous allocations of computer time.

References

18. For SiO2 one finds $\eta(T_c) \approx 10$ Poise and $T_c = 3330$ K and thus also temperatures below $T_c$ to some extent are accessible\(^1\). However, this case is somewhat pathological, since $T_c \approx 1.6 T_m$, very different from the situation sketched in Fig. 5.
19. Here $S_i$ is the value of spin number $i$ and the notation $\langle ... \rangle_{av}$ stands for averaging over the quenched disorder, e.g. the configurations where the nonmagnetic atoms sit in the Eu$_{1-x}$Sr$_x$S crystal (Fig. 2). Such a quenched disorder, present already at high temperatures, cannot be identified for the structural glass transition at all.
Exact Diagonalization Results for Strongly Correlated Electron-Phonon Systems

Holger Fehske\(^1\), Alexander Weiße\(^1\), and Gerhard Wellein\(^2\)

\(^1\) Physikalisches Institut, Universität Bayreuth, 95440 Bayreuth, Germany
E-mail: \{holger.fehske, alexander.weisse\}@uni-bayreuth.de

\(^2\) Regionales Rechenzentrum Erlangen, Universität Erlangen, 91058 Erlangen, Germany
E-mail: gerhard.wellein@rrze.uni-erlangen.de

Combining Lanczos, density matrix, kernel polynomial moment and maximum entropy algorithms we study the ground-state and spectral properties of interacting electron-phonon systems on massively parallel and vector supercomputers. In order to illustrate the efficiency and reliability of our approach, we present numerical results for the single-polaron band dispersion, the Luttinger liquid to charge-density-wave metal insulator transition, and the interrelation of spin, charge, orbital and lattice degrees of freedom in colossal magnetoresistive manganites. The scalability and performance of our CRAY T3E-1200 implementation is briefly discussed.

1 Motivation and Projects

The role of coexisting electronic correlations and electron-lattice interactions is rapidly becoming central to determining many physical phenomena in important classes of novel materials, including the electronically quasi-one-dimensional (1D) organic polymers, halide-bridged metal chain complexes, and quantum spin (Peierls) systems, the quasi-2D high-$T_c$ cuprate superconductors, as well as the 3D charge ordered nickelates and the colossal magnetoresistance manganites. From a theoretical point of view the challenge is to describe the partly exotic properties of these materials in terms of simplified microscopic models, taking into account the complex interplay of charge, spin, orbital and lattice degrees of freedom. Figure 1 gives a schematic survey of the physical problems and materials investigated in this context by the Bayreuth-Erlangen group at the John von Neumann Institute for Computing (NIC) during the last few years.

Adapting the corresponding model Hamiltonians to real physical situations one is frequently faced with the difficulty that the energy scales of electrons, phonons and their interaction are of the same order of magnitude, causing analytic methods to fail in most of these cases. Thus, at present, the most reliable results came from powerful numerical calculations like finite-cluster exact diagonalizations (ED), which are usually performed on supercomputers. As many other problems in theoretical physics, such ED studies are related to eigenvalue problems involving large sparse matrices. Iterative Lanczos or Jacobi-Davidson subspace methods are commonly used to determine the ground state and some excited eigenstates of the underlying Hamiltonian matrices. In general, the computational requirements of these eigenvalue algorithms are determined by a matrix-vector multiplication (MVM). MVM are also the most time and memory consuming step in the determination of spectral functions by kernel polynomial and maximum entropy methods. Thus the efficient use of modern supercomputers strongly depends on a parallel, fast and memory saving implementation of the MVM within the diagonalization routines. This was the main objective of our NIC project from the technical side.
Figure 1. Summary of the NIC project. For more detailed information see the following original papers: (i) polaron/exciton problem\textsuperscript{1}, (ii) spontaneous dimerization transitions\textsuperscript{2}, (iii) polaronic effects in strongly correlated electron systems\textsuperscript{3}, and numerical methods\textsuperscript{4}.

2 Selected Physical Problems

2.1 Polaron Band Structure

The very fundamental problem of a single conduction electron coupled to the phonon degrees of freedom is still not completely understood. The challenge is to describe the crossover from a weakly (phonon) dressed charge carrier to a strongly mass-enhanced, less mobile polaronic quasiparticle with increasing electron-phonon (EP) coupling strength. This process is termed “self-trapping” since the distortion that tends to bind the carrier in a deformable lattice is induced by the particle itself, i.e., the trapping potential depends on the state of the carrier. Nonetheless, even if the quasiparticle is confined to a single lattice site (small polaron), this type of trapping does not imply localization of the electron. The tunneling between different lattice sites is still relevant and self-trapped states remain itinerant even though the coherent band motion is substantially reduced in the strong EP coupling limit.
Unfortunately, as yet none of the various analytical treatments, based, e.g., on perturbation expansions, are suitable for the investigation of the physically most interesting transition region, where the highly non-linear self-trapping process of the charge carrier takes place. That is because precisely in this situation the characteristic electronic and phononic energy scales are not well separated and non-adiabatic effects become increasingly important, implying a breakdown of the standard Migdal approximation.

We addressed the problem of polaron band formation in terms of the Holstein model by performing systematic Lanczos studies. The aim is to discuss the evolution of the quasiparticle band structure in dependence on the phonon frequency and EP interaction strength. The Holstein Hamiltonian reads:

\[
\mathcal{H} = -t \sum_{\langle ij \rangle} (c_i^\dagger c_j + c_j^\dagger c_i) - g \omega_0 \sum_i (b_i^\dagger + b_i)(c_i^\dagger c_i - \frac{1}{2}) + \omega_0 \sum_i (b_i^\dagger b_i + \frac{1}{2}),
\]  

where \( c_i^\dagger \) and \( b_i^\dagger \) are the annihilation [creation] operators of (spinless) fermions and phonons at Wannier site \( i \), respectively. In (1), the following idealizations of real electron-phonon systems are made: (i) the electron transfer \( t \) is restricted to nearest-neighbour

![Figure 2](image_url)

Figure 2. Band dispersion \( E(K) \) of a single electron described by the Holstein model on 1D rings with \( N \) sites, where \( g = 2.73 \) and \( \omega_0 = 0.4 \). The insets show the low-energy part of the one-particle spectral function \( A_K(E) \) for selected momenta \( K \). The magenta line corresponds to the dispersion of a free particle with a renormalized bandwidth. The orange line gives a least-squares fit to an effective band dispersion \( \bar{E}_K = \sum_{l=0}^{3} a_l \cos lK \).
pairs $\langle ij \rangle$; (ii) the charge carrier is locally coupled to a dispersionsless optical phonon mode ($g$ denotes the EP coupling constant and $\omega_0$ is the bare phonon frequency); (iii) the phonons are treated within a harmonic approximation.

In order to determine the polaron band structure, at first one has to calculate the wavevector resolved spectral density function $A_{\vec{K}}(E) = \sum_n |\langle \Psi^{(1)}_{n,\vec{K}} | c^\dagger_{\vec{K}} | 0 \rangle|^2 \delta(E - E^{(1)}_n)$. Then the so-called “coherent” band dispersion, $E_{\vec{K}}$, can be extracted from the first peak of $A_{\vec{K}}(E)$ having finite spectral weight ($|\Psi^{(1)}_{n,\vec{K}}\rangle$ is $n$-th excited state in a certain $\vec{K}$-sector).

Figure 2 displays a typical polaron band dispersion obtained for the crossover regime. $E_{\vec{K}}$ exhibits several striking features: (i) Although we observe all signs of the famous polaronic band collapse, at intermediate-to-strong EP couplings the coherent bandwidth is by about a factor of 3 times larger than predicted by the standard Lang-Firsov formula $\Delta E_{LF} = 4t \exp\left[-g^2\right]$. (ii) The effective polaronic band dispersion deviates substantially from a simple tight-binding cosine band due to further than nearest-neighbour ranged hopping processes induced by the residual polaron-phonon interaction. As a consequence, the polaronic mass enhancement is substantially weakened. (iii) The flattening of the band dispersion at large momenta, which is normally observed in the weak-coupling case, persists to surprisingly large interaction strengths, even if the renormalized band width is by one order of magnitude smaller than the bare phonon frequency.

### 2.2 Luttinger-Liquid vs. Charge-Density-Wave Behaviour

Low dimensional electronic materials are known to be very susceptible to structural distortions driven by the EP interaction. Probably the most famous one is the Peierls instability of 1D metals: As the temperature is lowered the system creates a periodic variation in charge density, called a “charge-density-wave” (CDW), by shifting the electrons and ions from their symmetric positions. For the half-filled band case the dimerization of the lattice opens a gap at the Fermi surface; as a result the metal is transformed into an insulator.

Most theoretical treatments of the Peierls instability describe the lattice degrees of freedom classically. In a wide range of quasi-1D metals, however, the lattice zero-point motion is comparable to the Peierls lattice distortion, which makes the rigid lattice approximation questionable. By any means lattice dynamical (quantum phonon) effects should be included in the theoretical analysis of transport and optical phenomena.

Let us consider the 1D Holstein model of spinless fermions, Eq. (1), at half-filling ($N_e = N/2$). This model is of physical relevance in the strong interaction limit of the Hubbard model ($U \to \infty$). Based on a density-matrix algorithm for the phonon Hilbert space reduction, we have diagonalized the Holstein model on chains of even length with up to 12 sites using periodic boundary conditions. The resulting phase diagram is shown in Fig. 3. For small $g$ quantum phonon fluctuations destroy the dimerized ground state and the system becomes a metal, more precisely a Luttinger liquid LL, with parameters that vary with the coupling. For large $g$ the system has an energy gap and develops true long-range CDW order in the thermodynamic limit. The phase boundary obtained with our optimized phonon diagonalization methods was confirmed by the density matrix renormalization group (DMRG) results by Bursill et al. (Phys. Rev. Lett. 83, 408 (1999)). In the adiabatic limit $\omega_0 = 0$, the critical coupling converges to zero. In the strong-coupling anti-adiabatic regime, the half-filled Holstein model can be transformed to the exactly soluble
XXZ model, which exhibits a Kosterlitz-Thouless phase transition (long-dashed curve in Fig. 3).

The LL and CDW phases can be characterized in more detail. According to Haldane’s Luttinger liquid conjecture, 1D gapless systems of interacting fermions should belong to the same universality class as the Tomonaga-Luttinger model. As stated above, the Holstein system is gapless for small enough coupling $g$. The left insets of Fig. 3 give the LL parameters as a function of $g$ in the region, where our Lanczos data obey the LL finite-size scaling relations, $E_0(N) = \varepsilon_\infty - \frac{\pi u_\rho}{6N^2}$ and $E_{\pm 1}(N) - E_0(N) = \frac{\pi u_\rho}{2K \rho N}$. Here $\varepsilon_\infty$ ($E_0(N)$) denotes the ground-state energy per site of a infinite (finite) system (of $N$ sites), $K \rho$ is the effective coupling (stiffness) constant, and $u_\rho$ is the velocity of the charge excitations.

A very interesting result is that for small frequencies the effective fermion-fermion interaction is attractive, while it is repulsive for large frequencies, where the system forms a polaronic metal. Obviously there is a transition line in between, where the model de-

Figure 3. Ground-state phase diagram of the 1D Holstein model of spinless fermions at half-filling ($N_e = N/2$), showing the boundary between the Luttinger liquid (LL) and charge-density-wave (CDW) states obtained by ED and DMRG approaches. Left insets show the LL parameters $u_\rho$ and $K \rho$ as a function of the electron-phonon coupling $g$ in the metallic regime; right insets display for a six-site chain the regular part of the optical conductivity $\sigma^{\text{reg}}(\omega)$ (dotted lines) and the integrated spectral weight $S^{\text{reg}}(\omega) = \int_0^\infty d\omega' \sigma^{\text{reg}}(\omega')$ (solid lines) in the CDW region.
scribes “free” particles in lowest order. In the CDW state valuable information about the low-energy excitations can be obtained from the optical conductivity, which can be evaluated by means of very efficient and numerically stable Chebyshev recursion and maximum entropy algorithms. The optical absorption spectrum in the strong EP coupling regime is quite different from that in the LL phase. It can be interpreted in terms of strong EP correlations and corroborates the CDW picture. In the adiabatic region (upper right inset), the broad optical absorption band is produced by a single-particle excitation accompanied by multi-phonon absorptions. The most striking feature is the large spectral weight contained in the incoherent part of optical conductivity. Employing the f-sum rule for the optical conductivity and taking into account the behaviour of the kinetic energy ($\propto u^2 \rho$) as function of $g$, we found that in the metallic LL and insulating CDW phases nearly all the spectral weight is contained in the coherent (Drude) and incoherent (regular) part of $\text{Re } \sigma(\omega)$, respectively.

2.3 Lattice Dynamical Effects in CMR Manganites

The transition from a metallic ferromagnetic low-temperature phase to an insulating paramagnetic high-temperature phase observed in some hole-doped manganese oxides (e.g. in $\text{La}_{1-x}\text{[Sr, Ca]}_x\text{MnO}_3$) is associated with an unusual dramatic change in their elec-

![Figure 4](image_url)

Figure 4. **Upper panel:** (a) Crystal-field and Jahn-Teller splitting of the five-fold degenerate atomic Mn 3d levels (half-filled $t_{2g}$ triplets form local spins $S = 3/2$ interacting ferromagnetically with electrons in single occupied $e_g$ levels); (b) double-exchange model (the hopping amplitude reaches is maximum if the core spins are aligned and vanishes in the case of antiparallel orientation); (c) Jahn-Teller modes. **Lower panel:** Schematic phase diagram of $\text{La}_{1-x}\text{Ca}_x\text{MnO}_3$ reproduced from E. Dagotto et al. [Phys. Rep. 344 (2001)] (FM: Ferromagnetic Metal; FI: Ferromagnetic Insulator; AF: Antiferromagnet; CAF: Canted AF, and CO: Charge/Orbital Ordering (left figure)). Kinetic energy, selected spin and orbital correlation functions for the CMR model on a 4-site cluster, where $t_{(t_{2g})} = 1/3$, $J_H = 1.75$, $U = 15$, and $\omega_0 = 0.0175$ in units of $t_{(t_{2g})} = 0.4$ eV (right figures).
tronic and magnetic properties, including a spectacu-
larly large negative magnetoresistive response to an applied magnetic field, which might have important technological applications.

Early studies on lanthanum manganites concentrated on the link between magnetic correlations and transport, and attributed the low-$T$ metallic behavior to Zener’s double-exchange mechanism, which maximizes the hopping of a strongly Hund’s rule coupled Mn $e_g$-electron in a polarized background of the core spins (Mn $t_{2g}$-electrons) (see Fig. 4). Recently it has been argued that physics beyond double-exchange is important not only to explain the very complex phase diagram of the manganites but also the colossal magnetoresistance (CMR) transition itself. More specifically, orbital and lattice effects seem to be crucial in explaining the CMR phenomenon. There are two types of lattice distortions which are important in manganites. First the partially filled $e_g$ states of the Mn$^{3+}$ ion are Jahn-Teller active, i.e., the system can gain energy from a quadrupolar symmetric elongation of the oxygen octahedra which lifts the $e_g$ degeneracy. A second possible deformation is an isotropic shrinking of a MnO$_6$ octahedron. This “breathing”-type distortion couples to changes in the $e_g$ charge density, i.e., is always associated with the presence of an Mn$^{4+}$ ion. In the heavily doped material, both, breathing-mode collapsed (Mn$^{4+}$) and Jahn-Teller distorted (Mn$^{3+}$) sites are created simultaneously when the holes are localized in passing the CMR metal insulator transition.

To investigate the interaction of charge, spin, orbital and lattice degrees of freedom in CMR manganites without uncontrolled approximations, we performed exact diagonalizations based on the following very general Hamiltonian

$$H = H_{\text{double-exchange}} + H_{\text{electron-phonon}} + H_{\text{2nd order spin-orbital}} + H_{\text{phonon}}$$

$$= \sum_{i,\delta,\alpha,\beta} (a_{i,\delta}^\dagger a_{i+\delta,\delta,\alpha,\beta} + a_{i,\delta}^\dagger a_{i+\delta,\delta,\alpha,\beta}^\dagger) + \sum_{i,\delta,\alpha,\beta} g_0 \left( n_{i,\alpha} - n_{i,\beta} \right) (n_{i,\alpha} + n_{i,\beta}) (c_{i,\alpha}^\dagger c_{i,\alpha} + c_{i,\beta}^\dagger c_{i,\beta})$$

$$+ \sum_{i,\delta,\xi,\zeta} \left( J_{\xi\zeta}^S S_i S_{i+\delta} + \Delta_{\xi\zeta}^c P_i^c P_{i+\delta}^c \right) + \sum_{i,\theta,\epsilon} \left( b_{i,\theta}^\dagger b_{i,\theta} + b_{i,\epsilon}^\dagger b_{i,\epsilon} \right),$$

containing Schwinger bosons $a_{i,\mu,\nu}^{\dagger}$, i.e., $S_i = a_{i,\mu,\sigma}^\dagger a_{i,\mu,\nu}$ ($\mu, \nu \in \{\uparrow, \downarrow\}$), fermionic holes $c_{i,\alpha}^{\dagger}$, phonons $b_{i,\alpha}^{\dagger}$ ($\alpha \in \{\theta, \epsilon\}$), and orbital projectors $P_i^c$.

Exemplarily, in Fig. 4 we show the dependence of nearest neighbour spin and orbital correlations and of the kinetic energy on the strength of the EP coupling $g$. For doping level $x = 0.25$, which corresponds to the metallic regime, we found pronounced ferromagnetic correlations of both actual spins ($S$) and $z$-components of the pseudo-spins $\tau^z$ (representing orbitals, $\tau^z = n_{i,\theta} - n_{i,\epsilon}$). Only at very large EP couplings, a “trapping” of the $e_g$-electrons takes place (cf. the behaviour of $E_{\text{kin}}$ in Fig. 4), accompanied by the development of strong antiferromagnetic spin correlations. In contrast, at $x = 0.5$, a transition from a “ferromagnetic state” to an “antiferromagnetic state” with weak ferromagnetic spin correlations in $x$–direction is observed for a relatively modest EP interaction ($g_c \simeq 1$). Note that in this “phase” spin and pseudospin ordering is anticorrelated, and the absolute value of the kinetic energy shows a rather smooth decrease with increasing $g$. Of course, to be more predictive, ED of larger systems would be highly desirable.
3 Numerical Concepts and Methods

In general, ED studies of microscopic electron-phonon models involve very large sparse matrices, even for small clusters. Thus both continuous access to the most powerful supercomputers and steady improvements of algorithms and implementations are the technical basics of our project. In the past years the algorithms listed below have been successfully implemented on numerous architectures including CRAY T3E, NEC SX-4/5, IBM SP, Fujitsu VPP700 and Hitachi SR8000 supercomputers.

In order to discuss a wide variety of physical quantities and ensure an efficient use of supercomputers at the same time we use a wide variety of algorithms and techniques:

- **Lanczos diagonalization:**
  The standard Lanczos algorithm is employed to compute non-degenerate ground states. In that case the Lanczos algorithm provides rapid convergence and low memory consumption which allows us to determine ground states of matrices up to a dimension of \( D_{\text{tot}} \approx 4 \times 10^9 \) on the CRAY T3E at NIC.

- **Jacobi-Davidson algorithm with pre-conditioning:**
  In a joint project with A. Basermann (NEC) a Jacobi-Davidson algorithm was implemented to compute eigenstates in the low-energy spectrum with high resolution (including degeneracies): In typical production jobs a maximum of 50 eigenvalues with matrix dimension of \( D_{\text{tot}} \approx 10^7 \) can be computed at the same time.

- **Kernel-polynomial expansion and maximum-entropy method**
  A high quality approximation of spectral properties (involving all eigenstates) of large sparse matrices can be obtained using a Kernel-polynomial expansion in combination with a maximum-entropy method. This technique is applied to matrices up to a dimension of \( D_{\text{tot}} \approx 4 \times 10^9 \) and provides results which can be compared qualitatively with experimental data.

- **Density-matrix algorithm:**
  In selected parameter regions the matrix dimension can be substantially reduced by a transformation from bare phonon states to optimal phonon states. This transformation can be determined by a density-matrix algorithm in combination with a sequence of several hundred Lanczos diagonalization steps involving the matrix with reduced dimension.

The memory requirement and compute time of all algorithms is determined by a MVM step involving the large sparse matrix. Of course symmetries and conservation laws are exploited to reduce the matrix dimension. Nonetheless the matrix dimension and the quality of our calculations is mainly limited by the available memory. Therefore the large amount of memory per node on CRAY T3E-1200 at NIC has been a valuable benefit for our project. With regard to the coding of the MVM step two supplementary strategies have been chosen. The *in core* implementation stores the non-zero matrix-elements only, using storage techniques like the *compressed row storage* format or the *jagged diagonals storage* format. Together with a parallel MVM implementation based on non-blocking MPI (message passing interface) communication calls, the *in core* version provides high scalability and numerical performance at the cost of additional memory consumption. On the contrary, the
out core implementation does not even store the non-zero matrix-elements but recomputes the entries in each MVM step. In combination with algorithms (1) and (3) this technique allows us to reach very high matrix dimensions at the cost of additional computational effort. To optimize the scalability of the out core version a lot of work has been done to preserve the natural parallelism (direct product of electron and phonon Hilbert spaces) in the parallel program even for symmetrized basis states. Furthermore we have developed both a portable MVM based on non-blocking MPI communication calls and a CRAY specific version using the shmem library.

The single processor performance of the algorithms used is clearly limited by the available memory-bandwidth, because the sparse MVM step does not allow major cache-reuse. However, the total performance of the parallel implementation strongly depends on a well balanced ratio between local and remote memory bandwidth.

4 Performance Analysis on Supercomputers

The CRAY T3E series is widely regarded as a paradigm for scalability. Thus the efficient use of these systems is closely connected with the ability of algorithms and programs to scale to hundreds of processors. We have chosen an exemplary benchmark for our in core MVM implementation to deliver insight in the average performance characteristics of our computations. As a measure of scalability we use the parallel efficiency

\[ \epsilon_1(N_{\text{proc}}) = \frac{T(1)}{(N_{\text{proc}} \times T(N_{\text{proc}}))} \]

\[ T(N_{\text{proc}}) \] is the time per MVM step on \( N_{\text{proc}} \) processors at fixed problem size (Fig. 5).

For small to intermediate processor numbers (up to 96) a parallel efficiency of more than 75% can be sustained. Although there is a drop in the efficiency at 128 processors, our MVM implementation provides a scalable algorithm because it still exceeds an efficiency
of 70% even for 256 processors and there is only a minor performance dependence on the problem size. Most notably, the CRAY T3E provides outstanding scalability even when compared with next generation supercomputers as can be seen from Fig. 5 (b).

5 Concluding Remarks

The work presented in this report is an example for a successful interdisciplinary cooperation, spanning the fields of physics, algorithmic design and programming techniques. From the numerical point of view the overall goal was to develop valuable tools to perform fast and reliable many-body calculations on modern supercomputers. The implementation of the various optimized program packages on the T3E supercomputer at the NIC provided new and exciting insights to the interaction of electronic and phononic degrees of freedom in the currently most intensive studied novel materials: the quasi-1D metals and charge-density-wave systems, high-$T_c$ cuprates, polaronic nickelates and CMR manganites.

Acknowledgements

We are particularly indebted to the NIC Jülich, LRZ München and HLR Stuttgart for the generous granting of their parallel computer facilities. We acknowledge useful discussions with N. Attig, A. Basermann, E. Jeckelmann, B. Steffen, and R. Vogelsang.

References

   G. Wellein and H. Fehske, Polaron band formation in the Holstein model, Phys. Rev. B 55, 4513 (1997);
   G. Wellein and H. Fehske, On the self-trapping problem of electrons or excitons in one dimension, Phys. Rev. B 58, 6208 (1998);
2. A. Weiße and H. Fehske, Peierls instability and optical response in the one-dimensional half-filled Holstein model of spinless fermions, Phys. Rev. B, 58, 13526 (1998);
   B. Büchner, H. Fehske, A. P. Kampf, and G. Wellein, Lattice dimerization for the spin-Peierls compound CuGeO$_3$, Physica B 259-261, 956 (1999);
   A. Weiße, G. Wellein, and H. Fehske, Quantum lattice fluctuations in a frustrated Heisenberg spin-Peierls chain, Phys. Rev. B 60, 6566 (1999);
   H. Fehske, M. Holicki, and A. Weiße, Lattice dynamical effects on the Peierls transition in one-dimensional metals and spin chains, Advances in Solid State Physics 40, 235 (2000);


Zener Breakdown in Superlattices

Stephan Glutsch

Institut für Festkörpertheorie und Theoretische Optik
Friedrich-Schiller-Universität Jena
Max-Wien-Platz 1, 07743 Jena, Germany
E-mail: stephan@ifto.physik.uni-jena.de

In this paper we provide numerical and experimental evidence for the Zener breakdown in the optical spectrum of superlattices. It is demonstrated that the assumption of Wannier-Stark ladders and the Kane approximation are not justified in the regime of the Zener breakdown. The numerical calculation of the absorption spectrum, including Coulomb interaction, is performed by solving an initial-value problem, where the Hamiltonian in real space is discretized by finite differences. The method is highly efficient and perfectly suitable for vector machines. Storage and computing time scale like $O(N)$, where $N$ is the number of grid points.

1 Introduction

In 1960, Wannier\(^1\) discovered that the energy spectrum of a Bloch electron in an electric field consists of equally spaced eigenvalues, but also argued that the eigenfunctions are not normalizable and should therefore not lead to stationary states. On the other hand, the most common approximations, the discrete model,\(^2\) the tight-binding model,\(^3\) and the one-band approximation,\(^4\) lead to localized wave functions, and it was a common belief for a long time that the spectrum of a Bloch electron in an electric field would be discrete. Eventually the continuous nature of the spectrum was rigorously established by Avron et al.\(^5\) Nevertheless, from the practical point of view, the notion of “Wannier-Stark resonances” is justified when the lifetime broadening is much smaller than the level spacing.\(^6\)

In 1988, Wannier-Stark quantization was observed experimentally in photoluminescence, photocurrent, and photoluminescence excitation spectroscopy on superlattices\(^7,8\) and Bloch oscillations were measured in time-resolved and spectrally resolved four-wave-mixing experiments.\(^9\)

Traditionally, these experiments have been described in terms of Wannier-Stark ladders. It was shown that the Wannier-Stark ladder picture suffices to explain absorption measurements\(^7,8,10-12\) as well as nonlinear optical experiments.\(^9,13\) For high electric fields, Zener tunneling is expected to take place. Signatures of Zener tunneling have been reported in recent years.\(^14-16\) These papers mainly focus on the coupling of a finite number of minibands, also known as resonant Zener tunneling.

In the original paper, the Zener effect describes an open system with tunneling between one band and the continuum of all other bands under the influence of an electric field $F$.\(^17,18\) The tunneling rate $\gamma$ for interminiband transitions is equal to

$$\gamma = \frac{e|F|a}{2\pi\hbar} \exp \left[ -\frac{m_e a (\Delta E)^2}{4\hbar^2 e|F|} \right],$$

(1)

where $m_e$ is the effective electron mass, $a$ is the superlattice period, and $\Delta E$ is the gap between first and second miniband. This result was derived for nearly free electrons, under...
the assumption that the other gaps are significantly smaller than $\Delta E$. The quantity $\hbar \gamma$ can be considered as the imaginary part of the eigenvalues, which leads to an exponential decrease of the wavefunction and to a broadening in the optical spectrum. This non-resonant Zener tunneling has been observed in the optical absorption of strongly coupled superlattices, and has been explained mainly in term of interaction-free electrons and holes.

A particularly difficult problem is the calculation of the optical absorption in low-dimensional semiconductors. The interband absorption is related to the eigenvalue problem of an electron-hole pair, subjected to Coulomb interaction, geometric confinement, and external fields, which is in general a partial differential equation with up to six variables. The spectrum can be entirely discrete or possess a one- or many-parametric continuum. Only in very few cases with high symmetry, the problem can be solved analytically.

In the general case, the optical absorption has to be calculated numerically. This has been done by solving eigenvalue or boundary-value problems by exact or iterative methods. So far, in the case that the continuous spectrum of the electron-hole Hamiltonian is effectively one-parametric, the reformulation as a scattering problem proved to be most efficient, leading to very accurate results.

However, in the case that the continuum is two- or more-parametric, like in superlattices, conventional methods are not applicable. A highly efficient method, based upon discretization of the operator in real space and solution of an initial-value problem, which was developed by the author, is able to calculate optical spectra of the most complicated geometries, such as quantum wires, superlattices in electric and magnetic fields, or excitons on rough interfaces in many dimensions. Equation-of-motion methods are frequently used in quantum chemistry, solid-state physics, and material science. In many cases, memory and computing time can be further reduced considerably by the introduction of absorbing boundary conditions, as demonstrated by Ahland et al.

The paper is organized as follows. In Sec. 2, we consider the one-particle problem of an electron or hole in a superlattice with applied electric field. A brief introduction into methods of calculating the optical absorption of low-dimensional semiconductors is given in Sec. 3. In Sec. 4, we calculate the absorption of a superlattice for realistic parameters. The numerical results are compared with experimental spectra. Finally, a summary and conclusions are given in Sec. 5.

## 2 One-Particle Problem

We consider a particle with charge $q$ and mass $m$ in a one-dimensional periodic potential $U$ with the period $a$, subjected to an electric field $F$. A periodic potential, which is typically realized in a superlattice, is shown in Fig. 1. The eigenvalues and eigenfunctions obey the following equation:

$$\left[-\frac{\hbar^2}{2m} \frac{d^2}{dz^2} + U(z) - qFz\right] \varphi_\lambda(z) = E_\lambda \varphi_\lambda(z).$$

In case that $U(z) = 0$ and $F \neq 0$, the spectrum is entirely continuous and the eigenfunctions are given by the Airy function. For $U(z) \neq 0$ and $F = 0$, the eigenvalues and eigenfunctions $\varepsilon_\lambda(k)$ and $\psi_{k\lambda}$ obey the Bloch theorem. There is no simple rule how the spectrum looks like in the general case $U \neq 0$ and $F \neq 0$, for example, whether the eigenvalues are discrete or continuous. We only know that, if $E$ is an eigenvalue with
normalizable or non-normalizable eigenfunction \( \varphi(z) \), then \( E - qFam; m \in \mathbb{Z} \) is also an eigenvalue, and the corresponding eigenfunction is \( \varphi(z - ma) \). This was already observed in the original paper by Wannier.\(^1\)

In order to get a feeling about the structure of the solution, we construct a discrete model

\[
\sum_{n'} \left[ A \delta_{nn'} - B (\delta_{n,n-1} + \delta_{n,n+1}) + f n \delta_{nn'} \right] w_{n'} = E w_n. \tag{3}
\]

The solution can be found analytically and for a non-zero field \( f \) it holds that\(^2\)

\[
E_m = A + fm; \quad m \in \mathbb{Z} \quad w_{m;n} = J_m n \left( \frac{2B}{f} \right); \quad \sum_n w_{m;n} w_{m';n} = \delta_{m,m'}. \tag{4}
\]

The spectrum is entirely discrete and the eigenvectors are normalizable. The real constants \( A, B, \) and \( f \) can be identified with the parameters of the tight-binding model with next-neighbor interaction. Then, in principle, each below-barrier state of the single quantum well leads to a Wannier-Stark ladder. For above-barrier states, which are not localized, the procedure is not applicable, which suggests that those states give rise to a continuum for \( F \neq 0 \).

In the Kane approximation, the Hamiltonian is projected onto the individual subbands \( \lambda \). Then the eigenvalues and eigenfunctions (Kane or Houston functions) are:\(^4\)

\[
E_{\lambda m} = \frac{a}{2\pi} \int_{-\pi/a}^{+\pi/a} dk \left[ \varepsilon_\lambda(k) - qFZ_\lambda(k) \right] + meFa
\]

\[
\varphi_{\lambda m}(z) = \left( \frac{a}{2\pi} \right)^{1/2} \int_{-\pi/a}^{+\pi/a} dk \tilde{\varphi}_{\lambda m}(k) \psi_\lambda(z),
\]

Figure 1. Sketch of the periodic potential \( U \).
where

\[ \tilde{\varphi}_{\lambda n}(k) = \exp \left\{ \frac{1}{-iqF} \int_{0}^{k} dk' \left[ E_{\lambda m} - \varepsilon_{\lambda}(k') - eFZ_{\lambda \lambda}(k') \right] \right\} \]

(6)

\[ Z_{\lambda \lambda}(k) = \frac{i}{a} \int_{-a/2}^{+a/2} dz \ u_{k \lambda}^*(z) \frac{\partial u_{k \lambda}(z)}{\partial k} . \]

The normalization of the Bloch functions is assumed to be \( \delta_{\lambda \lambda'} \delta(k - k') \) modulo \( 2\pi/a \). This time, there is no difference between below- and above-barrier states: every miniband \( \lambda \) leads to a Wannier-Stark ladder \( E_{\lambda m} \).

With increasing field, the localization of the Kane functions increases. It can be shown\(^{36}\) that the extension of the Kane function is of the form \( \sqrt{a^2 + b^2/F^2} \). For \( |F| \to \infty \), the Kane function goes over in the Wannier function. This picture seems not to be realistic for large fields or weak confinement: if the modulation of the periodic potential is negligible, compared to \( qFa \), then the eigenfunctions should qualitatively behave like the Airy function and the spectrum should be continuous. It can indeed be shown that the spectrum is entirely continuous for arbitrary \( F \), but the proof is intricate.\(^5\) On the other hand, the spectrum can show resonances with a very small width, which behave like discrete eigenvalues. Practically, we expect expect three regions: the region of small fields, when the Kane approximation is justified, a transition region for medium fields, and a region of large fields, when the Kane approximation breaks down and the spectrum becomes a structureless continuum.

In order to study the accuracy of the Kane approximation, we make a comparison be-
between the Kane approximation (5) and the result of a numerical solution of Eq. (2), based upon finite differences. We use the parameters of the conduction electron in a 111/171A GaAs/Ga0.7Al0.3As superlattice: \( q_e = -e, m_e = 0.067 \, m_0, a = 12.8 \, \text{nm}, b = 1.7 \, \text{nm}, \) \( c = 11.1 \, \text{nm}, \) and \( \hbar_e = 237 \, \text{meV}. \) The result is shown in Fig. 2. The probability density \( |\varphi|^2 \) is plotted versus \( z \) for various fields \( F = 5 \ldots 70 \, \text{kV/cm}. \) For low electric fields, \( F \leq 15 \, \text{kV/cm}, \) the Kane function (blue) and the exact solution (red) are virtually indistinguishable. For high fields, \( F \geq 50 \, \text{kV/cm}, \) the Kane function approaches the Wannier function and becomes independent of the electric field. In contrast, the exact solution shows a strong tunneling through the barrier, leading to delocalization. The tail for \( z \to -\infty \) qualitatively behaves like the solution for \( U = 0. \)

The delocalization of the eigenfunctions has consequences for the optical absorption. For localized eigenfunctions, we expect sharp resonances, while for delocalized eigenfunctions absorption should become entirely continuous. In the one-particle picture, the optical absorption can be characterized by the joint density of states

\[
D(\omega) = \frac{1}{L} \sum_{\lambda\lambda'} \left| \int_{-\infty}^{+\infty} dz \, \varphi_{e\lambda}(z) \varphi_{h\lambda'}(z) \right|^2 \frac{\pi}{\hbar} \delta\left[ \hbar \omega - (E_{e\lambda} + E_{h\lambda'} + E_g) \right].
\]

Here, \( \varphi_{e\lambda} \) and \( \varphi_{h\lambda} \) are the eigenfunctions of electron and hole with energies \( E_{e\lambda} \) and \( E_{h\lambda}. \) The eigenvalue problem for the hole follows from Eq. (2) by replacing \( m_e \) by \( m_h \) and \( -e \) by \( +e. \) The normalization length is considered in the limit \( L \to \infty. \) The hole parameters for the GaAs/(Ga,Al)As superlattice under consideration are \( q_h = +e, m_h = 0.45 \, m_0 \) and \( \hbar_h = 63.2 \, \text{meV}. \) The band gap of GaAs is \( E_g = 1.52 \, \text{eV}. \)

The optical density \( D \) as function of the photon energy \( \hbar \omega \) and the electric field is shown in Fig. 3. For low values of the electric fields, we observe two Wannier-Stark
ladders, associated with the (1,1) and (2,2) transitions. The positions of the resonances scales linear with field and the distance between subsequent lines of the Wannier-Stark ladders is \( e Fa \). With increasing field, the resonances shift to lower energies. Furthermore, we observe strong interactions between different Wannier-Stark ladders. For large fields, the Wannier-Stark ladders are destroyed. For the (2,2) ladder, this is the case for \( F \geq 50 \text{kV/cm} \). The (1,1) ladder disappears at about \( F = 60 \text{kV/cm} \). The remaining resonances cannot be directly attributed to a Wannier-Stark ladder and are strongly broadened. The situation is similar to a bulk semiconductor in an electric field, where the spectrum shows Franz-Keldysh oscillations.

3 Numerical Method

The calculation of absorption spectra in low-dimensional semiconductors is very complicated, especially in cases where the continuum is more than one-parametric. This problem was the subject of numerous papers in the last decade. Here, we give a brief overview over different methods of calculating the optical absorption.

The absorption coefficient \( \alpha(\omega) \) is proportional to the imaginary part of the optical susceptibility \( \chi(\omega) \). Generally, the optical susceptibility of a low-dimensional semiconductor (in dimensionless units) is given by\(^{21-23}\)

\[
\chi(\omega) = \sum_{\Lambda} \frac{|\langle \mu | \Phi_{\Lambda} \rangle|^2}{E_{\Lambda} - (\omega + i\epsilon)}
\]

where \( \hat{H} \) is the electron-hole-pair Hamiltonian, which is a differential operator in up to six variables, \( |\mu\rangle \) is the dipole matrix element, and \( \epsilon = +0 \) is a positive infinitesimal. In reality, \( \epsilon \) is finite, and is identical to the homogeneous line broadening. After discretization of \( \hat{H} \), the numerical effort for the diagonalization is in the order of \( N^3 \), where \( N \) is the dimension of the matrix \( H \). The dimension \( N \), which can be handled, is in the order of a few thousands. Iterative methods, based upon the Lanczos algorithm are able to calculate the absorption spectrum in \( O(N^3) \) operations.\(^{24}\)

The representation (8) is equivalent to the following boundary-value problem:\(^{27}\)

\[
\chi(\omega) = \langle \mu | \left[ \hat{H} - (\omega + i\epsilon) \right]^{-1} | \mu \rangle.
\]

This time, a set of equations has to be solved for each \( \omega \). The numerical effort is \( O(N^3) \), if the matrix \( H \) is fully occupied, but can be significantly larger if \( H \) is tridiagonal or banded. Tridiagonal or banded matrices result from the discretization of one-dimensional problems or multi-dimensional problems, when the domain is very anisotropic, like quantum wells or quantum wires. For multi-parametric continua, the solution as a boundary-value problem is not practicable.

Both the eigenvalue and the boundary-value problem can be reformulated as a scattering problem.\(^{10,27,29}\) This method is highly accurate and efficient, but is limited to problems with one continuum direction.

276
By Fourier transform, Eq. (9) goes over into\textsuperscript{30}

\[
\chi(\omega) = i \int_0^\infty dt e^{i(\omega+it)x} \langle \mu | \Psi(t) \rangle ; \quad i \frac{d}{dt} | \Psi(t) \rangle = \hat{H} | \Psi(t) \rangle ; \quad | \Psi(0) \rangle = | \mu \rangle , \tag{10}
\]

which requires the solution of a Schrödinger initial-value problem. If \( \hat{H} \) is discretized in real space, using finite differences, the matrix \( H \) is (irregularly) sparse and the calculation of the spectrum takes only \( O(N) \) operations, independent of the dimensionality and the domain. Furthermore, for regular meshes, the algorithm is ideally suitable for vector machines. The method allows to calculate the optical response for the most difficult geometries.

A significant reduction of the storage and the computational effort can be achieved by means of absorbing boundary conditions. Ahland et al. used a complex coordinate transform, leading to decrease of the wavepacket in the absorption layer.\textsuperscript{34} An alternative formulation, used by the author, is the introduction of a complex effective mass inside the absorption layer.\textsuperscript{37}

### 4 Optical Absorption

The Hamiltonian of the electron-hole pair in a superlattice, in the presence of an electric field in growth direction is

\[
\hat{H} = -\frac{\hbar^2}{2m_\rho} \Delta_\rho - \frac{\hbar^2}{2m_e} \frac{\partial^2}{\partial z_e^2} - \frac{\hbar^2}{2m_{hz}} \frac{\partial^2}{\partial z_h^2} + U_e(z_e) + U_h(z_h)
\]

\[
+ F(z_e - z_h) - \frac{1}{\sqrt{\rho^2 + (z_e - z_h)^2}}.
\tag{11}
\]

This time, we take into account the anisotropy of the hole mass. The parallel and perpendicular hole masses are \( m_{hz} = 0.377 m_0 \) and \( m_{hp} = 0.491 m_0 \). The parallel reduced mass is \( m_\rho = m_e m_{hp} / (m_e + m_{hp}) \). The dipole moment is \( \mu(\rho) \propto \delta(\rho) \delta(z_e - z_h) / (2\pi \rho) \) and the volume element, which enters the definition of the scalar product (10), is \( 2\pi \rho d\rho \, dz_e \, dz_h \).

For the numerical solution, we introduced center and relative coordinates \( Z = z_e \in [0, a] \) and \( z = z_e - z_h \in (-\infty, +\infty) \). This choice ensures that the mesh points for \( z_e \) and \( z_h \) are identical. The Bloch theorem can be applied for the variable \( Z \) and only those eigenfunctions contribute to the optical absorption, which are periodic in \( Z \). The eigenvalue problem was rewritten as an initial-value problem (10) and absorbing boundary conditions were used for the radial direction. The total number of grid points was in the order of 10 million.

We calculated the optical absorption for a 76/39\,\text{GaAs/Ga}_{0.92}\text{Al}_{0.08}\text{As} superlattice. Due to the small Al content of 8\%, the barriers are very shallow, \( h_e = 63.2 \text{meV} \) and \( h_h = 36.8 \text{meV} \), which leads to an increased tunneling probability. We compare the numerically exact solution with the approximate result, when the two-particle Hamiltonian is projected onto the electron and hole Kane function.

The result is shown in Fig. 4a. The numerical exact result (red curve) is compared with the Kane approximation. For zero electric field, we observe an exciton peak and

277
a flat continuum. For small fields, $F \leq 10$ kV/cm, we see one Wannier-Stark ladder, associated with the (1,1) transition. The lines are inhomogeneously broadened due to Fano interference. There is no qualitative difference between the exact and the approximate solution, although the influence of non-resonant subbands, neglected in the approximate solution, leads to an overall increase of the absorption. Starting at $F \geq 20$ kV/cm, exact and approximate solution behave differently. For the approximate solution, the absorption peak narrows and the oscillator strength increases with field. In contrast, in the exact solution, we observe a line broadening and a shift of the maximum to lower energies, which is a result of Zener tunneling. For very large fields, $F \geq 50$ kV/cm, the absorption becomes a flat continuum.

Figure 4 b shows experimental results, obtained by Rosam and co-workers. The experimental spectrum is very similar to the result of the full numerical solution. This is true for the Wannier-Stark ladder, the Fano interference, the line shift, and the broadening for large electric fields. On the other hand, the theoretical spectrum, based upon the Kane approximation, is not able to explain the experimental results for large electric fields.

Finally, we compare the measured linewidth with the prediction of the Zener model (1). For this reason, we compare the linewidth (HWHM) with the function

$$h\gamma = \frac{e}{2\pi} AF \exp\left(-\frac{B}{F}\right) + C$$

where $A$, $B$, and $C$ are fitting parameters. The result is shown in Fig. 5. The function (12) is shown by a solid line, while the blue circles represent the experimental values. For the fitting parameters, we obtain $A = 6.09$ nm, $B = 21.8$ kV/cm, and $C = 0.632$ meV. The values of $A$ and $B$ are in the order of values expected from Eq. (1). The constant $C$ is equal to the intrinsic broadening at $F = 0$, which is not considered in the Zener model.
Figure 5. Linewidth vs. electric field. Blue circles: experimental values. Solid line: fit to Eq. (1) with $A = 6.09$ nm, $B = 21.8$ kV/cm, and $C = 0.632$ meV

The experimental linewidth shows strong oscillations, which result from anticrossing with other Wannier-Stark ladders (cf. Fig. 1).

5 Summary and Conclusions

In this paper we have demonstrated that Zener breakdown is realistic in semiconductor superlattices. The numerical calculation of the spectrum was performed by means of a highly efficient real-space-real-time method. The experimental results are in good agreement with the predictions of the theory. Meanwhile, the Zener breakdown has also been detected in a superlattice with perpendicular electric and magnetic fields. In this geometry, Fano interference is suppressed by the Landau quantization and a transition is observed between entirely discrete and continuum states.\textsuperscript{38}

The calculations were performed on a Cray J 90 or T 90. Depending on the problem, the computing time was between 3 days and 3 weeks.

Acknowledgments

The author is indebted to K. Leo and B. Rosam for contributing the experimental results. Computing time from the John von Neumann Institute for Computing, Research Center Jülich, is gratefully acknowledged.

References

27. R. Zimmermann, phys. stat. sol. (b) 135, 681 (1986).
A Bosonic Model for High-Temperature Superconductivity and Antiferromagnetism: Numerical Simulation Studies

A. Dorneich 1, E. Arrigoni 1, W. Hanke 1, M. Troyer 2, and Shoucheng Zhang 3

1 Institut für Theoretische Physik, Universität Würzburg, Am Hubland, 97074 Würzburg, Germany
E-mail: \{ardornei, arrigoni, hanke\}@physik.uni-wuerzburg.de

2 Institut für Theoretische Physik, ETH Zürich, 8093 Zürich, Switzerland

3 Department of Physics, Stanford University, Stanford, 94305 California

As their most prominent universal feature, high-temperature superconductors always display antiferromagnetism and $d$–wave superconductivity in close proximity, in their phase diagram. A unifying theory has been proposed, according to which these two at first sight radically different phases are “two faces of one and the same coin”. They are unified by a common symmetry principle, the SO(5) symmetry. Recently, it was proposed that this theory had to be supplemented with a so-called “Gutzwiller projection” in order to resolve several inconsistencies with experimental results on high-$T_c$ superconductors. Here, we present a numerical study of an effective bosonic model which describes the low-energy physics of this “projected” SO(5) theory. Our numerical results, obtained by the Quantum Monte Carlo technique of Stochastic Series Expansion, show that this model provides a realistic description of the global phase diagram of the high-$T_c$ superconductors and accounts for many of their physical properties. Moreover, we address the question of asymptotic restoring of the SO(5) symmetry at the critical point.

1 Introduction: The SO(5) Theory

The high-temperature superconductors 1, 2 (HTSC), discovered at the end of the eighties, are characterized by a number of fascinating properties. The most appealing one, from which their name derives, is the fact that they can conduct electrical current without resistance up to relatively high temperatures. In these materials, commonly called cuprates, the dominant charge-carrier dynamics takes place in the two-dimensional (2D) CuO$_2$-planes 3. Each CuO$_2$ unit cell contains an effective magnetic moment of spin $\frac{1}{2}$, essentially due to the Cu ion. At higher temperatures, neighboring Cu-spins form so-called singlets, i.e. pairs of electrons with antiparallel spin (Fig. 1). The energy gain due to the singlet formation, the magnetic exchange $J$, is relatively large $\sim 120\text{meV} \sim 1400\text{K}$. On the other hand, the temperatures ($T$) for the transition into both low-temperature phases, the antiferromagnetic (AF) and the superconducting (SC) phases, $T_N$ and $T_c$ (see Fig. 1), are both significantly lower and of similar magnitude ($\sim 250\text{K}$ for $T_N$ and $\sim 100\text{K}$ for $T_c$). Already this order of magnitude suggests that the mechanism of superconductivity does not directly result from the singlet formation, but is instead related to the mechanism, which results in the antiferromagnet in the insulating situation. In addition, the AF and the SC phases are in close proximity in the phase diagram of Fig. 1. Therefore, it seems tempting to look for a common origin for these low-temperature phases, despite the fact that on first glance, they appear dramatically different: On the one hand, the insulator and, on the other hand, the ideal conductor, i.e. the superconductor.
Figure 1. Generic temperature ($T$) vs. doping phase diagram of the high-temperatures superconductors. The AF (red) and the SC (blue) phases can be seen as a condensation of magnons (red) or hole pairs (blue) on top of the same “spin-liquid” state (see text).

Let us consider, at first, the insulator: At high temperatures $T^* \sim 1000$ K, the singlet pairs are completely disordered. This state is termed, therefore, a spin liquid (Fig. 1). How does one arrive from this disordered state at high temperatures, to an ordered AF state at low temperatures? The idea is that the ordered AF state can be considered as a kind of Bose-Einstein condensation of magnon excitations\(^4\). The SC state, on the other hand, corresponds to a Bose-Einstein type of condensation of hole pairs (Cooper pairs). Magnon excitation and hole-pairs excitation are “two faces of one and the same coin” in similarity to other unifying concepts such as the isospin theory of proton and neutron in nuclear physics\(^5\). The condensation energy of these two excitations yields then the corresponding temperature scales $T_N$ and $T_c$.

These considerations led S.C. Zhang to formulate the so-called SO(5) theory of HTSC\(^6\). This theory unifies the three-dimensional order parameter of the AF phase with the two-dimensional SC order parameter. The interpretation of the AF and SC phases of the HTSC within SO(5) theory is as follows: in a completely SO(5)-symmetric system the superspin vector can rotate within a five-dimensional sphere, and we would expect mixed states of coexisting AF and SC long-range order. In reality, however, the chemical potential (which controls the hole doping) induces an anisotropy between AF and $d$-wave SC and explicitly breaks SO(5) symmetry. The chemical potential in SO(5) theory plays the same role as the magnetic field in angular-momentum with a mechanism similar to the Zeeman effect.

In summary, the basic idea of SO(5) theory is that antiferromagnetism and superconductivity are “two faces of the same coin” – just as the electric and magnetic field in the

---

\(^4\)Magnons are electron pairs with parallel spins (Fig. 1)
theory of relativity or the proton and the neutron in Heisenberg’s isospin concept\textsuperscript{5}.

Unfortunately, the presence of a Mott insulating behavior at half-filling (i.e., zero doping) in the cuprates severely challenges the validity of the SO(5) theory\textsuperscript{7–10}. SO(5) symmetry requires collective charge pair excitations to have the same (vanishing) mass as collective spin-wave excitations. The real cuprates, on the contrary, are Mott insulators at half-filling and possess a large energy gap $U$ of several eV due to electron-electron interaction. The only way to overcome these problems is to project out states with doubly-occupied sites. These states are separated from the ones without double occupancies by an energy scale of more than 10 eV, which is by orders of magnitude higher than the low-energy scales $T_N$ and $T_c$. Therefore, states with double occupancies should not even be important as intermediate states for scattering processes. The resulting models are called ‘projected SO(5)’ or ‘pSO(5)’ models.

2 “Projected” SO(5) Bosonic Hamiltonian

In Ref.\textsuperscript{11} a low-energy effective bosonic model was constructed in which the Gutzwiller constraint of no-double-occupancy was implemented exactly. This is done by projecting out the mode creating particle pair excitations and by retaining only the massless magnon and hole-pair modes. In Ref.\textsuperscript{4} it has been shown that the low-energy SO(5) excitations on the rung of a ladder can be cast into a picture of 5 hard-core bond bosons: three magnon states ($t_{\alpha=2,3,4}$) and particle and hole-pair state ($t_h^\dagger$ and $t_h^\dagger$, respectively). As an effective coarse-grained model, this description can be extended to a two-dimensional system, whereby the excitations are now defined on a $2 \times 2$ plaquette\textsuperscript{11}. The Gutzwiller projection is implemented by restricting the Hilbert space to states with $t_p(x) = 0$. The projected SO(5) Hamiltonian takes the form\textsuperscript{11}

$$H = \Delta_s \sum_{x,\alpha=2,3,4} t_{\alpha}^\dagger(x) t_\alpha(x) + (\Delta_c - 2\mu) \sum_x t_h^\dagger(x) t_h(x)
- J_s \sum_{<xx'>,\alpha=2,3,4} n_\alpha(x)n_\alpha(x') - J_c \sum_{<xx'>} (t_h^\dagger(x) t_h(x')+h.c.),$$

where $n_\alpha = (t_\alpha + t_\alpha^\dagger)/\sqrt{2}$ are the three components of the AF order parameter. $\Delta_s$ and $\Delta_c \sim U$ are the energies to create a magnon and a hole-pair excitation, respectively, at vanishing chemical potential $\mu = 0$. As one can see, the excitation energy for hole pairs can be compensated by $\mu$ in order to have equal energies for spin and hole-pair excitations. Due to this partial compensation, the mean-field ground state of this model\textsuperscript{11} recovers SO(5) invariance at $J_c = 2J_s$ and $\Delta_s = \Delta_c$. However, this invariance is not exact, and a symmetry-breaking effect can already be seen at the Gaussian level\textsuperscript{11}. Nevertheless, this is the simplest bosonic model containing two generic ingredients which are relevant for the high-$T_c$ materials, namely, the Mott gap and the vicinity and, possibly, common origin of the antiferromagnetic and of the superconducting phases.

3 Numerical Analysis

In this Section, we study the physics of the bosonic pSO(5) model introduced in Eq. 1. As already pointed out, numerical simulations are currently the only methods to study
this strongly-correlated system in an appropriate manner, i.e. including all particle-particle interaction effects correctly. Our numerical results for two-dimensional (2D) lattice geometries were obtained by means of Stochastic Series Expansion (SSE, see Ref.\textsuperscript{12}). We were able to simulate systems of up to $40 \times 40$ lattice sites and to evaluate arbitrary dynamical response functions. Our results show that this model gives a realistic description of the global phase diagram of the high-$T_c$ superconductors and accounts for many of their physical properties. Moreover, we address the question of asymptotic restoring of the SO(5) symmetry at certain critical points of the phase diagram. We carry out our analysis with numerical simulations on an isotropic 2D square lattice, a good model for the physics of the cuprates’ CuO$_2$ planes.

3.1 Doping Dependence of the Chemical Potential and Phase Separation

We choose $J_s = J_c/2$, corresponding to the SO(5)-symmetric point in the mean-field approach\textsuperscript{11} and define $J := J_s$ as our unit of energy. We, take $\Delta_s = J$, and shift the chemical potential so that $\Delta_c = \Delta_s$. The mean-field calculation of Ref.\textsuperscript{11} predicts a phase transition from the AF to the SC phase accompanied by a jump in the hole-pair and magnon densities at a chemical potential $\mu_c = 0$. In mean field, the system at $\mu = \mu_c$ is in a coherently mixed AF+SC phase and there is no phase separation. However, Gaussian fluctuations change the picture and predict a first-order transition. Here, we want to study this region in more detail with an appropriate strong-coupling method. In fact, we expect the picture to change appreciably, since no long-range order is allowed in two dimensions. A jump in the density as a function of $\mu$ (or, equivalently, a constant $\mu$ as a function of density) has also been seen in $La_{2-x}Sr_xCuO_4$.\textsuperscript{13} Fig.2 displays a comparison of our numerical calculation with the experimental data of Ref.\textsuperscript{13}. As one can see, our data reproduce the experimental results within error accuracy.

The nature of the phase transition at $\mu = \mu_c$ can be determined by studying histograms of the hole-pair distribution for fixed $\mu = \mu_c$. While in an homogeneous phase the density is peaked about its mean value, at $\mu = \mu_c$, we obtain two peaks indicating a first-order transition with a phase separation between (almost) hole-free regions and regions with high hole-pair density. Interestingly, a phase separation into hole-rich and almost hole-
free phases can also be observed in real cuprates, for example in the HTSC compound La$_2$CuO$_{4+y}$\textsuperscript{14}. For doping concentrations \(y = 0.01\) to \(0.055\) the material displays a mixture of two phases with different oxygen concentrations: there are alternating AF regions with very low doping concentration \((y = 0.01)\) and hole-rich ‘stage 6’ regions \((y = 0.055)\) which become superconducting at low temperatures (Fig. 3b). This is in remarkable accordance with the pSO(5) results in Fig. 3a. Due to the fact that each additional oxygen attracts and immobilizes 2 electrons in the CuO$_2$ planes, thereby introducing 2 holes, the accordance between theory and experiment is perfect even on a quantitative level: the doping densities of \(y = 0.01\) and \(y = 0.055\) of the two phases exactly correspond to the values \(\delta = 2y = 0.02\) and \(\delta = 0.11\) obtained from the pSO(5) model.

### 3.2 Quasi Long-Range Order and ‘Superconducting’ Phase

We have anticipated the existence of a SC phase for \(\mu > \mu_c\). In fact, in two dimensions at \(T > 0\) a true long-range order is prohibited by the Mermin-Wagner theorem. However,
we can still have a so-called Kosterlitz-Thouless\textsuperscript{15} (KT) phase of finite superfluid density $\rho_s$ (see below) at finite temperature, which is identified by a power-law decay of the SC correlation function $C_h(r)$. The transition separates long-range power-law ($C_h(r) \propto r^{-\alpha}$) from rapid exponential decay ($C_h(r) \propto e^{-\lambda r}$). A reliable and accurate distinction between these two decay behaviors requires a finite-size scaling with large system sizes, as well as an efficient Quantum-Monte-Carlo (QMC) estimator for the Green functions appearing in the correlation function. With its non-local update scheme and with our new estimators for arbitrary Green functions, SSE provides both. Fig. 4 demonstrates how precisely a

![Graph showing decay behavior of the SC correlation function](image)

Figure 4. Decay behavior of the SC correlation function $C_h(r)$ for $T/J = 0.5$ and for the chemical potentials $\mu = -0.1$ (red squares) and $\mu = 0.3$ (blue triangles). The $C_h(r)$ data for $\mu = 0.3$ can be fitted by a straight line in a log-log plot of $\log(C_h)$ versus $\log(r)$ (left), indicating a power-law decay, while an exponential decay can be deduced from the linear fit in a semi-logarithmic plot for the $\mu = -0.1$ data.

correlation length can be determined with SSE. For the point $(T = 0.5J, \mu = -0.1)$, which is located in close neighborhood to the phase transition line in fig. 5 the semi-logarithmic plot almost perfectly fits a straight line, indicating an exponential decay. For the neighbored points $(T = 0.4J, \mu = -0.1)$ or $(T = 0.5J, \mu = 0.1)$ we would find the straight fit in a log-log plot instead of the semi-logarithmic plot, which indicates that these points are located in a different phase with power-law decay behavior of $C_h(r)$ (see fig. 5).

In the numerical simulations presented here, the largest system size used for finite-size scaling was $32 \times 32$, in some calculations only $24 \times 24$ lattice sites. Earlier QMC studies of simpler Hamiltonians than (1), e.g. the analysis of phase transitions in the quantum XY model by Ding\textsuperscript{16}, needed much larger lattice sites to determine phase transition lines and coherence lengths with high precision. However, the QMC methods employed in those works suffer from systematic Trotter discretization errors and from rapidly growing autocorrelation times on large system sizes, which considerably blows up the statistical errors. In Ding’s work\textsuperscript{16}, for example, typical autocorrelation times on the largest lattice, the $128 \times 128$ lattice, were 5000 update–measurement sweeps. Together with the total sweep number of twice $6 \cdot 10^5$, this means that only about 200 statistically uncorrelated data points were recorded, so that the relative statistical error of the QMC results was
about $1/\sqrt{200} \approx 7\%$. In SSE, on the contrary, there are no systematic errors, and the loop update mechanism produces autocorrelation times of the order of 1 even on large lattices. Therefore, the recorded finite-size data typically have relative errors of not more than $10^{−3}$. Obviously (see Fig. 4), these high-precision finite-size data allow for a reliable finite-size scaling even on moderate-sized lattices.

Apart from a change in the decay behavior of the superconducting correlation function, there is a second criterion describing the KT transition point: It exploits the fact that the superfluid density jumps from zero to a finite value at the KT temperature $T_{KT}^{17}$. Within QMC methods, the superfluid density can be measured quite easily by counting winding numbers$^{18−20}$. In two dimensions the superconducting density $\rho_s$ is given in terms of the mean-squared winding number $\langle W \rangle$ via $\rho_s = \frac{\mathcal{m} k_B T}{\langle W^2 \rangle}$. In finite-cluster simulations, the jump of $\rho_s$ from zero to a finite value at the KT transition point can be detected more easily than the change in the decay behavior of the correlation function itself. Fig. 5 plots

![Figure 5](image)

Figure 5. Location of the superconducting phase transition in the projected SO(5) model: The black points identify the long-distance decay behavior of the SC correlation function, the small dots corresponding to an exponential decay, while the crosses indicate a power-law decay. The connected red circles trace the transition temperature from the jump in the boson phase stiffness in the infinite-volume limit. This behavior should be compared with the typical dependence of the critical superconducting temperature versus doping (red region of Fig. 1).

the phase diagram obtained by applying both criteria independently. The figure shows that the projected SO(5) model indeed has a KT phase with quasi long-range order whose form in $\mu$-$T$ space looks like the one of the high-$T_c$ cuprates. Both criteria result in exactly the same clearly pronounced phase-separation line.

### 3.3 Spin Resonance Peak

One of the main features of SO(5) theory is that it provides an elegant explanation for the neutron resonance peak observed in some high-$T_c$ cuprates at $k = (\pi, \pi)^6$. Experiments show that the resonance energy $\omega_{res}$ is an increasing function of $T_c$, i.e. $\omega_{res}$ increases as a function of doping in the underdoped and decreases in the overdoped region$^{21}$. Here, we address the question whether the $T_c$ dependence of $\omega_{res}$ can be reproduced within the projected SO(5) model. To this purpose, we study the spin correlation function at $k = (\pi, \pi)$. 

287
Fig. 6 shows the spin correlation spectrum obtained from the projected SO(5) model in two dimensions as a function of the chemical potential.

The magnon-dominated paramagnetic region and the underdoped SC region are described correctly: spin-wave excitations are essentially massless Goldstone modes in the magnon-dominated phase at \( \mu < \mu_c \) and become massive when entering into the SC phase. The resonance energy \( \omega_{\text{res}} \) increases monotonically up to optimal doping \( \mu_{\text{opt}} \approx 1 \). On the other hand, the SO(5) theory is not expected to describe the system too far away from the AF/SC transition, i.e. in the overdoped regime. Notice that the resonance peak continuously looses weight when increasing \( \mu \), which is consistent with experimental observations\(^{21}\).

A comparison of the critical temperature \( T_c \) obtained from Fig. 5 and \( \omega_{\text{res}} \) at optimal doping yields the ratio \( T_c / \omega_{\text{res, opt}} = 0.23 \). This is again in accordance with the corresponding ratio for YBa\(_2\)Cu\(_3\)O\(_{6+x}\), for which the experimentally determined values \( T_c = 93 \) K (thus \( k_B T_c = 8.02 \) meV) and \( \omega_{\text{res, opt}} = 41 \) meV yield \( T_c / \omega_{\text{res, opt}} = 0.20 \).

3.4 How SO(5)-Symmetric is the pSO(5) Model?

So far we have presented some general properties of the pSO(5) model in two dimensions. The question “how SO(5)-symmetric is the pSO(5) model”, i.e. the question whether there exists a point for which the full SO(5)-symmetry is dynamically restored, still remains open. As one can see from Eq. (1), the excitation energy for hole pairs can be compensated by \( \mu \), in order to have equal energies for local spin and hole-pair excitations. Due to this partial compensation the mean-field ground state of this model recovers exact SO(5) invariance at \( J_c = 2 J_\sigma \) and \( \Delta_\sigma = \Delta_c \). However, this invariance is not exact, and a symmetry-breaking effect can already be seen at the Gaussian level\(^{11}\). For a classical three-dimensional SO(5)-symmetric model, numerical simulations indicate that the symmetry is asymptotically restored at a bicritical point provided the symmetry-breaking terms have the appropriate sign\(^{22,23}\). This is in contrast with the prediction from the \( \epsilon \)-expansion\(^{24}\),
which would suggest a fluctuation-induced first-order transition. The discrepancy clearly indicates that strong-coupling effects play an important role, calling for an appropriate, i.e. numerically exact treatment, as it is provided by SSE. This assumption is further supported by a recent work by A. Aharony.

Figure 7. Phase diagram of the pSO(5) model: The squares between S and the tricritical point P trace the first-order line of phase separation. The solid line from P to the right edge of the plot traces the Kosterlitz-Thouless transition between the “SC” and the normal state. The dashed line separating \( N_t \) (=magnon dominated region) and \( N_h \) (=hole pair dominated region) describes the line of equal AF and SC correlation lengths. The small inlay shows the same phase diagram on a larger \( \mu \) scale, covering the whole “SC” KT phase.

One necessary condition for an SO(5)-symmetric point is that the formation energies of hole-pair bosons and of magnons are identical. This condition is fulfilled along the line from S to the tricritical point P in Fig. 7. Another necessary condition is that hole pairs and magnons behave in the same way at long distances. This condition is fulfilled on the dashed line in Fig. 7, where the AF and SC correlation lengths \( \xi \) become equal. Interestingly, these two conditions meet (within error bar accuracy) at the tricritical point P. Although here the correlation length is still finite, we find relatively large \( \xi \) values of order 10 to 15 in the immediate vicinity of point P, demonstrating the importance of SO(5) critical fluctuations in this region.

The asymptotic restoring of SO(5) symmetry in the vicinity of the critical point cannot be conclusively answered in \( D = 2 \) dimensions, even though we have found a critical-point scenario which makes asymptotic restoring of SO(5) symmetry at a \( T \neq 0 \) bicritical point in \( D = 3 \) a good possibility. For that reason (and because there exists no AF phase at finite temperature in \( D = 2 \)) it would be interesting to analyze the model in three dimensions.

4 Summary

In summary, we have shown that the projected SO(5) model in two dimensions – or more general four-boson models of type Eq.(1), can be considered as a generic model for the high-temperature superconducting materials. It gives a semiquantitative or even quantitative description of many properties of the HTSC in a consistent way. In particular, we have identified an AF and a SC phase whose bounds look similar to the the ones of the real cuprate materials. Also, the doping dependence of the chemical potential as well as that of
the neutron-resonance peak in the underdoped regime are reproduced correctly. Furthermore, the pSO(5) model reproduces physical effects like phase separation or coexistence of antiferromagnetism and superconductivity.

Acknowledgments

The authors acknowledge financial support from BMBF (05SB8WFA1), DFG [(HA 1537/17-1) as well as from the Heisenberg program (AR 324/3-1)], KONWIHR OOPCV and the Swiss National Science Foundation. The calculations were carried out on a Cray T3E at the high-performance computing center HLRZ (Jülich) and required a total amount of about $1.5 \times 10^5$ CPU hours.

References

An efficient scheme is developed to study magnetism and structure on atomic scale. Starting by \textit{ab initio} calculations of the electronic structure in the framework of density functional theory, interaction potentials for molecular dynamics simulations of metallic nanostructures supported on metallic surfaces are carefully optimized.

The two methods are shortly explained. Examples for the application of the methods are given. Mainly electronic and structural properties of Co nanostructures on the Cu(001) surface are investigated.

1 Introduction

The essence of nanosience and technology is the ability to understand and manipulate matter at the molecular level. Structures behave differently when their dimensions are reduced to dimensions between 1 and 100 nm. Such structures show novel physical and chemical properties, due entirely to their nanoscopic size.

In the frontier field of nanomagnetism, understanding of the relationship of between magnetism and structure plays a central role. During the past few years experimental investigations of metallic nanostructures in the initial stage of heteroepitaxial growth revealed a lot of information which asks for a consistent theoretical explanation. Some important effects experimentally observed recently are:

- Surface alloying is found also for metals immiscible in bulk form (Co/Cu(001)).\textsuperscript{1,2}
- Burrowing of Co clusters into Au, Cu and Ag surfaces has been observed.\textsuperscript{3,4}
- It was observed, that the motion of adatoms on top of islands is is not the same as on a flat surface.\textsuperscript{5}

During the last years we developed a combination of methods which can be used to study such effects in detail theoretically. First we will discuss the methods used for our calculations. Especially the newly developed method of quasi-\textit{ab initio} molecular dynamics (QMD) simulations is discussed more in detail. Then we will discuss how the Korringa-Kohn-Rostocker (KKR) Green’s function (GF) method is used to study magnetic properties of magnetic nanostructures on noble metal substrates. The QMD method is used to develop a new point of view on mismatch in heteroepitaxial growth. We distinguish between macroscopic and mesoscopic mismatch. This definition allows for a more detailed discussion and understanding of the dependence of surface and island related properties.
on the size of the supported islands. Several of the above mentioned experimental results can be explained by using the \textit{ab initio} techniques and MD simulations together. This is discussed for the adatom motion on supported islands and surface alloying.

2 Theoretical Methods

2.1 Calculation of Electronic Structure

The electronic structure calculations are carried out with the KKR Green’s function method for surface defects. The method is based on density functional theory in the local-spin-density approximation (LSDA).

By removing the atomic potentials of seven monolayers from a bulk crystal we create two half-crystals which are practically decoupled. In this way the surface can be treated as a localized two-dimensional perturbation of the bulk. Multiple scattering theory is applied to obtain the Green’s function from the Dyson equation. The Green’s function of the ideal surface is used as the reference Green’s function for the calculation of the electronic properties of metallic nanostructures on the surface. The potential in a sufficiently large cluster around the supported nanostructure has to be considered as perturbed.

![Figure 1. Structure to calculate the surface Green's function for the (001) surface of the fcc-structure (blue - decoupled half-crystals, brown - vacuum layers)](image)

The full charge density is taken into account using a multipole expansion up to angular momentum $l_{\text{max}} = 6$. Coulomb and exchange correlation energies are evaluated using $l_{\text{max}} = 12$. For the latter ones the local functional of Vosko \textit{et al.} is used. Potentials are assumed to be spherically symmetric inside the Wigner-Seitz sphere.

2.2 Quasi-Ab Initio Molecular Dynamics

In the last few years we developed a method which connects the \textit{ab initio} electronic structure calculations with large scale molecular dynamic simulations. Our approach is based on fitting of the interaction parameters of potentials for molecular dynamic simulations to accurate first-principle calculations of selected cluster-substrate properties and bulk properties of the system under investigation.
To describe metallic clusters on noble metal substrates many body potentials in the second moment tight-binding approximation are used.9

The cohesive energy \( E_{coh} \) is the sum of the band energy \( E_B \) and the repulsive part \( E_R \):

\[
E_{coh} = \sum_i \left( E_R^i + E_B^i \right) \quad (1)
\]

\[
E_B^i = - \left( \sum_j \frac{\xi^2}{\alpha^2} \exp\left( -2q_{\alpha\beta} \left( \frac{r_{ij}}{r_{0\beta}} - 1 \right) \right) \right)^{1/2} \quad (2)
\]

\[
E_R^i = \sum_j \left( A_{\alpha\beta}^1 \left( \frac{r_{ij}}{r_{0\beta}} - 1 \right) + A_{\alpha\beta}^0 \right) \exp\left( -p_{\alpha\beta} \left( \frac{r_{ij}}{r_{0\beta}} - 1 \right) \right) \quad (3)
\]

where \( r_{ij} \) represents the distance between atoms \( i \) and \( j \), and \( r_{0\beta}^\alpha \) is the first-neighbor distance in the \( \alpha\beta \) lattice structure, while it is just an adjustable parameter in the case of the cross interaction. \( \xi \) is an effective hopping integral that depends on the material, and \( q_{\alpha\beta} \) and \( p_{\alpha\beta} \) describe the dependence of the interaction strength on the relative interatomic distance.

We will explain the method for the system Co/Cu(001). Co and Cu are not miscible in bulk form. Therefore the determination of the cross interaction is a problem. We solve the problem by a careful fitting to accurate first-principles calculations of selected cluster-substrate properties. The result is a manageable and inexpensive scheme able to account for structural relaxation and including implicitly magnetic effects, crucial for a realistic determination of interatomic interactions in systems having a magnetic nature. After determination of the Cu-Cu parameters which are fitted to experimental data only9 the Co-Co and Co-Cu parameters are optimized simultaneously by including in the fit the results of first-principles KKR calculations. To this purpose, we have taken the solution energy of single Co impurity in bulk Cu10 \( E_{Co \ in \ Cu} \), energies of interaction of two Co impurities in Cu bulk11 \( E_{Co-Cu} \), \( E_{Co-Cu} \), and binding energies of small supported Co clusters on Cu(001) \( E_{Co-Cu} \), \( E_{Co-Cu} \), \( E_{Co-Cu} \), \( E_{Co-Cu} \), \( E_{Co-Cu} \), \( E_{Co-Cu} \), \( E_{Co-Cu} \), \( E_{Co-Cu} \), \( E_{Co-Cu} \), \( E_{Co-Cu} \).

The set of data used to define the potential and the corresponding values calculated by means of the optimized potential are given in Table I. The bulk and surface properties are well reproduced.

The quality of the potential was tested in different ways. The \textit{ab initio} datapool contains only ideal lattice structures, i.e. relaxations are not taken into account. Recently, the KKR Green’s function method was extended to take into account the effect of lattice relaxations.12 Calculations of the lattice distortion in the vicinity of a Co impurity in Cu bulk have been performed with the KKR-GF method. Our QMD method shows a good agreement with this results. The second test of our potential is done performing calculations of interlayer distances in Co/Cu multilayers. A detailed LEED study of the Co/Cu(100) films for different Co coverages was performed in Kirschner’s group.13 Interlayer spacing was determined for different Co coverages. We use our potentials to determine interlayer distances performing energy minimization calculations. The results demonstrate that the agreement with experiment is rather good. Therefore we believe that the parametrisation developed in our paper gives a good description of the Co-Cu bonds. Applications of the potential are given later in the article.
<table>
<thead>
<tr>
<th>Quantity</th>
<th>Data</th>
<th>fitted value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cu (fcc)</td>
<td>$a_{Cu}$</td>
<td>3.615 Å</td>
</tr>
<tr>
<td></td>
<td>$E_c$</td>
<td>3.544 eV</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>1.42 Mbar</td>
</tr>
<tr>
<td></td>
<td>$C_{11}$</td>
<td>1.76 Mbar</td>
</tr>
<tr>
<td></td>
<td>$C_{12}$</td>
<td>1.25 Mbar</td>
</tr>
<tr>
<td></td>
<td>$C_{44}$</td>
<td>0.82 Mbar</td>
</tr>
<tr>
<td>Co (hcp)</td>
<td>$a_{Co}$</td>
<td>2.507 Å</td>
</tr>
<tr>
<td></td>
<td>$E_c$</td>
<td>4.386 eV</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>1.948 Mbar</td>
</tr>
<tr>
<td></td>
<td>$C_{11}$</td>
<td>3.195 Mbar</td>
</tr>
<tr>
<td></td>
<td>$C_{12}$</td>
<td>1.661 Mbar</td>
</tr>
<tr>
<td></td>
<td>$C_{13}$</td>
<td>1.021 Mbar</td>
</tr>
<tr>
<td></td>
<td>$C_{33}$</td>
<td>3.736 Mbar</td>
</tr>
<tr>
<td></td>
<td>$C_{44}$</td>
<td>0.824 Mbar</td>
</tr>
<tr>
<td>Co-Cu</td>
<td>$E_{S\text{in Cu}}$</td>
<td>0.4 eV</td>
</tr>
<tr>
<td></td>
<td>$E_{1,\text{b}}^{\text{Co-Co}}$</td>
<td>-0.12 eV</td>
</tr>
<tr>
<td></td>
<td>$E_{2,\text{b}}^{\text{Co-Co}}$</td>
<td>0.03 eV</td>
</tr>
<tr>
<td></td>
<td>$E_{\text{Co-Co}}^{\text{on Cu(100)}}$</td>
<td>-1.04 eV</td>
</tr>
<tr>
<td></td>
<td>$E_{1,\text{in Cu(100)}}^{\text{Co-Co}}$</td>
<td>-0.26 eV</td>
</tr>
<tr>
<td></td>
<td>$E_{\text{trimer}}^{\text{on Cu(100)}}$</td>
<td>-2.06 eV</td>
</tr>
<tr>
<td></td>
<td>$E_{2x2 \text{ island on Cu(100)}}^{\text{Co-Co}}$</td>
<td>-3.84 eV</td>
</tr>
</tbody>
</table>

Table 1. Data used for the fitting of the potential together with the values calculated with the optimized potential. (Cohesive energy $E_c$, bulk modulus $B$, elastic constants $C_{ij}$ from Cleri et al.\textsuperscript{9}, first and second neighbor interaction energies $E_{1,\text{b}}^{\text{Co-Co}}$, $E_{2,\text{b}}^{\text{Co-Co}}$ from Drittler et al.\textsuperscript{10}, solution energy $E_{S\text{in Cu}}^{\text{Co}}$ from Hoshino et al.\textsuperscript{11}, binding energies of small Co clusters $E_{\text{Co-Co}}^{\text{on Cu(100)}}$, $E_{1,\text{in Cu(100)}}^{\text{Co-Co}}$, $E_{\text{trimer}}^{\text{on Cu(100)}}$, $E_{2x2 \text{ island on Cu(100)}}^{\text{Co-Co}}$ are calculated using KKR Green’s function method.)

### 3 Magnetic Properties of Nanostructures on Metallic Surfaces

Using the KKR Green’s function method we have studied the properties of 3d, 4d and 5d adatoms on Ag(001), Pd(001) and Pt(001) systematically.\textsuperscript{14, 15} One central point of investigation was the study of imperfect nanostructures. We have investigated the influence of Ag impurities on the magnetism on small Rh and Ru clusters on the Ag(001) surface.\textsuperscript{16} The change of the magnetic moments could be explained in the framework of a tight-binding model. Nevertheless it was observed that the magnetism of Rh nanostructures shows some unusual effects.\textsuperscript{17} An anomalous increase in the magnetic moments of Rh adatoms on the Ag(001) surface with decreasing interatomic distance between atoms was observed, whereas for dimers of other transition metals the opposite behavior is observed.

#### 3.1 Mixed Co-Cu Clusters on Cu(001)

The magnetic properties of Co nanostructures on Cu substrate can be strongly influenced by Cu atoms. For example, Cu coverages as small as three hundredths of a monolayer...
drastically affect the magnetization of Co films. Experiments and theoretical studies demonstrated that magnetization of mixed clusters of Co and Cu depends on the relative concentration of Co and Cu in a nonobvious way. Quenching of ferromagnetism in Co clusters embedded in copper was reported. Calculations by means of the QMD method showed that surface alloying is energetically favorable for in the case of Co/Cu(001) and mixed Co-Cu clusters are formed in the early stages of heteroepitaxy. Recent experiments suggest that mixed Co-Cu clusters indeed exist.

![Figure 2. Spin polarization of Co-Cu mixed clusters on Cu(001). Magnetic moments in Bohr magnetons are given for all inequivalent site.](image)

We have studied all possible mixed configurations in $3 \times 3$-atoms islands on Cu(001) surfaces. We observe a small induced moment at the Cu atoms in the island and a decrease of the moments at the Co atoms in comparison with the $3 \times 3$ Co-island. A stronger reduction of the Co moments is achieved, if the Co$_9$ cluster is surrounded by a Cu brim and capped by a Cu cluster. A reduction of 14 % is obtained for the average moment of the Co$_9$ cluster. This effect should have a strong influence on the properties of the Co-Cu interface in the early stages of growth. Coating of Co clusters with Cu atoms has been found recently in experiments.

4 Strain and Stress on the Mesoscale

4.1 The Concept of Mesoscopic Misfit

If some material is grown on a substrate with a different bond length the lattice mismatch is at the interface leads to strain fields. Strain can be relieved through the introduction of defects in the atomic structure, such as dislocations, or by an atomic rearrangement. Usually strain relaxations are predicted on the basis of the macroscopic lattice mismatch between the two materials. However, if the deposited system is of mesoscopic size of several 100 atoms, its intrinsic bond lengths are different from the bond length in the bulk materials. For the Co/Cu(001) interface the macroscopic mismatch $m_0$ between Co and Cu defined as $m_0 = (a_{Cu} - a_{Co})/a_{Cu}$ ($a_{Cu}$ and $a_{Co}$ are the lattice constants) is only $\approx 2\%$. Several recent experiments have suggested that strain relaxations for submonolayer
coverage\textsuperscript{20} or even for a few monolayers\textsuperscript{21} cannot be explained by the macroscopic misfit between bulk materials.

In order to get a deeper insight into the local strain relaxations on an atomic scale, the equilibrium geometries of plane square Co islands of different sizes (up to 600 atoms) on Cu(001) are calculated by computing the forces at each atomic site and relaxing the geometry of islands and the substrate atoms by means of the quasi-\textit{ab initio} method.

![Figure 3. Size-dependent mismatch](image)

In Fig. 2.1 we show the change in the mismatch with the size of Co islands. It is seen, that the mesoscopic mismatch between small Co islands and the substrate is considerably larger than the mismatch calculated from the lattice constants of the two materials. Only for Co islands incorporating more than 200 atoms the local strain can be described by the macroscopic mismatch. We found that both the mesoscopic and macroscopic mismatch depend on the size of the islands and for islands larger than 60 atoms mismatch scales like $N^{-0.5}$ ($N$ - number of atoms in islands). Such scaling behavior is determined by the relaxations of the edge atoms of the islands whose number changes as $\sqrt{N}$. One very fundamental issue predicted by these results is the possible strong impact of the size dependent mismatch on the local strain field. The substrate can dynamically respond to the growth of islands and can exhibit a strong inhomogeneous strain distribution during the growth process.

### 4.2 Strain and Adatom Motion on Mesoscopic Islands

The mesoscopic mismatch or mesoscopic strain depend strongly on the size of the clusters. Therefore we expect that also the barriers for hopping diffusion depend on the size of the
clusters. The calculations show, that the barriers for hopping on the small Co islands (16-50 atoms) are found to be about $\approx 20\%$ lower than those on the large islands (100-500 atoms) (cf. Fig. 3.1). The diffusivity $D$ is related to the hopping rate of single adatoms by $D = D_0 \cdot \exp\left(-\frac{E_d}{kT}\right)$, where $E_d$ is the energy barrier for hopping, $D_0$ is the prefactor. We found that $D_0$ is nearly the same for all islands, therefore the diffusion coefficient $D$ on small Co islands at room temperature is found to be about two orders of magnitude larger than that on large Co islands.

![Graph showing energy barrier for hopping diffusion on Co square islands](image)

Figure 4. Strain dependence of energy barrier for hopping diffusion on top of Co square islands. Since the strain depends on island size, the activation barrier for diffusion depends also on island size (see upper horizontal scale).

5 Surface Alloying

In the past few years, much research has been performed on surface alloys. The most remarkable finding is that place exchange processes can result in the formation of surface alloys, even for metals immiscible in bulk form. For example it was concluded that Co atoms intermix with Cu at the Cu(100) surface by an atomic exchange process.\(^1\) Also burrowing of Co clusters into Au, Cu and Ag substrates has been observed.\(^3\)

When a material A is deposited on a surface B one needs to know whether intermixing takes place and whether A atoms are isolated or form clusters in a surface layer. Simple arguments to understand the atomic behavior on surfaces are based on such macroscopic properties as surface and interface energies of the components. However they are rather questionable when applied to individual adatoms on a metal surface, whose interaction determines the atomic picture in the early stages of thin-film growth.

The KKR Green’s function method was used to study the energetics of surface alloying of 3d adatoms on an atomic scale.\(^{22}\) Differences in total energy between the initial (A) and
the final (B) configuration of the atomic exchange process are calculated.

![Figure 5. Energetics of the exchange process. Energy difference between complex B and A is presented.](image)

Surprisingly, we find that for all 3d impurities it is energetically favorable to exchange with Au surface atoms. The energy gains are large for all impurities. For Fe, Co, and Ni impurities our calculations support the conclusions of the experimental investigations. The results suggest, that a simple model of epitaxial growth based on an abrupt boundary between components can be inadequate to describe structural and magnetic properties of 3d nanostructures on a gold surface. All 3d adatoms considered in the investigation are magnetic on the Au(100) surface. Calculations of paramagnetic Co adatoms on Au(100) show, that magnetism tends to stabilize Co on the surface and prevents site exchange.

It was shown by Hoshino et al. that the fundamental characteristic features of the phase diagrams can be qualitatively explained by the nearest-neighbor interaction of impurity pairs. The attractive interaction leads to segregation, the repulsive one to a solid solution. Thus, the calculation of the interaction energy of impurity pairs at the surface should provide information about cluster formation processes at the surface. The interaction energies of 3d impurities at nearest-neighbor sites on the surface (adatom positions) and in the first surface layer (terrace position) of Au(100) have been calculated. Negative energy means attraction and positive energy repulsion between adatoms. Only Fe and Co impurities show an attraction which is weak and not significantly different from zero.

**Acknowledgments**

The calculations by means of the KKR Green’s function method have been performed during the last years at the NIC. The projects have been supported also by Deutsche Forschungsgemeinschaft (DFG) and Deutscher Akademischer Austauschdienst (DAAD).
Figure 6. Interaction energies of $3d$ impurities on the nearest neighbor sites on the Au(100) surface and in the surface layer.

The cooperation with the groups of P.H. Dederichs (IFF, Forschungszentrum Jülich) and of J. Kirschner (MPI für Mikrostrukturphysik) is gratefully acknowledged.

References

Multi-Overlap Simulations of Spin Glasses

Wolfhard Janke, Bernd A. Berg, and Alain Billoire

1 Institut für Theoretische Physik, Universität Leipzig, 04109 Leipzig, Germany
   E-mail: wolfhard.janke@itp.uni-leipzig.de

2 Dept. of Physics, The Florida State University, Tallahassee, FL 32306, USA
   E-mail: berg@hep.fsu.edu

3 CEA/Saclay, Service de Physique Théorique, 91191 Gif-sur-Yvette, France
   E-mail: billoir@spht.saclay.cea.fr

We present results of recent high-statistics Monte Carlo simulations of the Edwards-Anderson Ising spin-glass model in three and four dimensions. The study is based on a non-Boltzmann sampling technique, the multi-overlap algorithm which is specifically tailored for sampling rare-event states. We thus concentrate on those properties which are difficult to obtain with standard canonical Boltzmann sampling such as the free-energy barriers $F_B$ of the probability density $P_J(q)$ of the Parisi overlap parameter $q$ and the behaviour of the tails of the disorder averaged density $P(q) = [P_J(q)]_{av}$.

1 Introduction

A widely studied class of spin-glass materials consists of dilute solutions of magnetic transition metal impurities in noble metal hosts, for instance Au-2.98% Mn or Cu-0.9% Mn, which is one of the best investigated metallic spin glasses. In these systems, the interaction between impurity moments is caused by the polarization of the surrounding Fermi sea of the host conduction electrons, leading to an effective interaction of the so-called RKKY form

$$J_{\text{eff}}(R) = \frac{\cos(2k_F R)}{R^3}, \quad k_F R \gg 1,$$

where $k_F$ is the Fermi wave number. For an illustration, see Fig. 1. This constitutes the two basic ingredients necessary for spin-glass behaviour, namely

- randomness – in course of the dilution process the positions of the impurity moments are randomly distributed, and

- competing interactions – due to the oscillations in (1) as a function of the distance $R$ between the spins some of the interactions are positive and some are negative.

The competition among the different interactions between the moments means that no single configuration of spins is uniquely favoured by all of the interactions, a phenomenon which is commonly called “frustration”. This leads to a rugged free-energy landscape with probable regions (low free energy) separated by rare-event states (high free energy), illustrated in many previous articles by vague sketches similar to our Fig. 2. Experimentally this may be inferred from the phenomenon of aging which is typical of measurements of the remanent magnetization in the spin-glass phase.
Figure 1. The two basic ingredients for spin-glass behavior: Randomly distributed spin moments of transition metal impurities (e.g. Mn) in a noble metal host (e.g. Cu), and the characteristic oscillatory form of their effective RKKY interactions with competing positive (ferromagnetic) and negative (antiferromagnetic) regions.

However, despite the large amount of experimental, theoretical and simulational work done in the past thirty years to elucidate the spin-glass phase, the physical mechanisms underlying its peculiar properties are not yet fully understood. To cope with the enormous complexity of the problem various levels of simplified models have been studied theoretically. A simplified lattice model which reflects the two basic ingredients for spin-glass behavior is the Edwards-Anderson\textsuperscript{9} Ising (EAI) model defined through the energy

$$E = - \sum_{\langle ik \rangle} J_{ik} s_i s_k ,$$

(2)

where the fluctuating spins $s_i$ can take the values $\pm 1$. The coupling constants $J_{ik}$ are quenched, random variables taking positive and negative signs, thereby leading to competing interactions. In our study we worked with a bimodal distribution, $J_{ik} = \pm 1$ with equal probabilities, but other choices such as the Gaussian distribution, $\mathcal{P}(J_{ik}) \propto \exp(-J_{ik}^2/2\Delta^2)$ with $\Delta$ parameterizing its width, have also been considered, in particular in analytical work. In (2), the lattice sum runs over all nearest-neighbour pairs of a $d$-dimensional (hyper-) cubic lattice of size $N = L^d$ with periodic boundary conditions.

An analytically more tractable mean-field model, commonly known as the Sherrington-Kirkpatrick\textsuperscript{10} (SK) model, emerges when each spin is allowed to interact with all others. Alternatively one may consider the mean-field treatment as an approximation which is expected to become accurate in high dimensions\textsuperscript{11}. In physical dimensions, however, its status is still unclear and an alternative droplet approximation\textsuperscript{12} has been proposed. The two treatments yield conflicting predictions which have prompted quite a controversial discussion over many years. Numerical approaches such as Monte Carlo (MC) simulations can, in principle, provide arbitrarily precise results in physical dimensions. In practice, however, the simulational approach is severely hampered by an extremely slow (pseudo-) dynamics of the stochastic process.
Figure 2. Typical sketch of the rugged free-energy landscape of spin glasses with many minima separated by rare-event barriers.

To overcome the slowing-down problem various novel simulation techniques have been devised in the past few years. While some of them only aim at improving the (pseudo-) dynamics of the MC process, others are in addition also well suited for a quantitative characterization of the free-energy barriers responsible for the slowing-down problem. Among the latter category is the multi-overlap algorithm\textsuperscript{13} which has recently been employed in quite extensive MC simulations\textsuperscript{14–17} of the EAI spin-glass model in three and four dimensions. The purpose of this note is to give an overview of the recent progress achieved with this method.

2 Model Parameters and Simulation Method

As order parameter of the EAI model one usually takes the Parisi overlap parameter\textsuperscript{11}

\[
q = \frac{1}{N} \sum_{i=1}^{N} s_i^{(1)} s_i^{(2)},
\]

where the spin superscripts label two independent (real) replicas for the same realization of randomly chosen exchange coupling constants \(\mathcal{J} = \{J_{ik}\}\). For given \(\mathcal{J}\) the probability density of \(q\) is denoted by \(P_{\mathcal{J}}(q)\), and thermodynamic expectation values are computed as \((\ldots)_{\mathcal{J}} \equiv \sum_{\{s\}} (\ldots) \exp(-\beta H[\mathcal{J}]) / \sum_{\{s\}} \exp(-\beta H[\mathcal{J}])\), where \(\beta = 1/k_B T\) is the inverse temperature in natural units. The freezing temperature is known to be at \(\beta_c = 0.90(3)\) in 3D (Ref. 18) and at \(\beta_c = 0.485(5)\) in 4D (Ref. 19), respectively.

On finite lattices the results necessarily depend on the randomly chosen quenched disorder. To get a better approximation of the infinite system (apart from special problems with so-called non-self-averaging), one performs averages over many hundreds or even
thousands of (quenched) disorder realizations denoted by

\[
P(q) = [P_\mathcal{J}(q)]_{\text{av}} = \frac{1}{\# \mathcal{J}} \sum_\mathcal{J} P_\mathcal{J}(q), \quad [\langle \ldots \rangle_\mathcal{J}]_{\text{av}} = \frac{1}{\# \mathcal{J}} \sum_\mathcal{J} \langle \ldots \rangle_\mathcal{J},
\]

where \#\mathcal{J} \to \infty is the number of realizations considered. Below the freezing temperature, in the infinite-volume limit \( N \to \infty \), a non-vanishing part of \( P(q) \) between its two delta-function peaks at \( \pm q_{\text{max}} \) characterizes the mean-field picture\(^{11}\) of spin glasses, whereas in ferromagnets as well as in the droplet picture\(^{12}\) of spin glasses \( P(q) \) exhibits only the two delta-function peaks. Most studies so far considered mainly the averaged quantities.

For a better understanding of the free-energy barriers sketched in Fig. 2, the probability densities for individual realizations \( \mathcal{J} \) play the central role. It is, of course, impossible to get complete control over the full state space, and to give a well-defined meaning to “system state” (the \( x \)-axis in Fig. 2), one has to concentrate on one or a few characteristic properties. In our work we focused on the order parameter and thus those free-energy barriers \( F^q_{\mathcal{J}} \) which are reflected by the minima of \( P_\mathcal{J}(q) \). A few typical shapes of \( P_\mathcal{J}(q) \) as obtained in our simulations are shown in Fig. 3. Conventional, canonical MC simulations are not suited for such a study because the likelihood to generate the corresponding rare-event configurations in the Gibbs canonical ensemble is very small. This problem is overcome by non-Boltzmann sampling\(^{20,21}\) with the multi-overlap weight\(^{13}\)

\[
w_\mathcal{J}(q) = \exp \left[ \beta \sum_{(ik)} J_{ik} \left( s_i^{(1)} s_k^{(1)} + s_i^{(2)} s_k^{(2)} \right) + S_\mathcal{J}(q) \right],
\]

where the two replicas are coupled by \( S_\mathcal{J}(q) \) in such a way that a broad multi-overlap histogram \( P^{\text{munq}}_\mathcal{J}(q) \) over the entire accessible range \(-1 \leq q \leq 1\) is obtained, see Fig. 4. When simulating with the multi-overlap weight (5), canonical expectation values of any quantity \( \mathcal{O} \) can be reconstructed by reweighting, \( \langle \mathcal{O} \rangle^\text{can} = \langle \mathcal{O} e^{-S_\mathcal{J}} \rangle_\mathcal{J} / \langle e^{-S_\mathcal{J}} \rangle_\mathcal{J} \).

The multi-overlap algorithm may be summarized as follows:

- An iterative construction of the weight function \( W_\mathcal{J}(q) \equiv \exp(S_\mathcal{J}(q)) \), for each of the quenched disorder realizations separately.

- An equilibration period with fixed weight function.

- The production run with fixed weight function.

Ideally \( W_\mathcal{J} \) should satisfy \( P^{\text{munq}}_\mathcal{J}(q) = P^\text{can}_\mathcal{J}(q) W_\mathcal{J}(q) = \text{const.} \), i.e., should give rise to a completely flat multi-overlap probability density \( P^{\text{munq}}_\mathcal{J}(q) \) as sketched in Fig. 4. Of course, \( P^\text{can}_\mathcal{J}(q) \) is a priori unknown and one has to proceed by iteration. Let us thus assume some approximate \( W_{\mathcal{J},n} \) is given. The simulation would then yield \( P^{\text{munq}}_{\mathcal{J},n} \) which, in general, is not yet perfectly flat. If \( P^{\text{munq}}_{\mathcal{J},n} \) was sampled with arbitrary precision, the desired weight function would be \( W_\mathcal{J} \propto W_{\mathcal{J},n} / P^{\text{munq}}_{\mathcal{J},n} \). For the update procedure we are actually only interested in relative transition amplitudes and it is therefore useful to rephrase the iteration in terms of

\[
R_\mathcal{J}(q) \equiv W_\mathcal{J}(q + \Delta q) / W_\mathcal{J}(q) = R_{\mathcal{J},n}(q) P^{\text{munq}}_{\mathcal{J},n}(q) / P^{\text{munq}}_{\mathcal{J},n}(q + \Delta q),
\]
where $\Delta q = 2/N$ is the step-width in $q$ as defined in Eq. (3). When updating the spins in the $n$th iteration, $R_{J,n}(q) = W_{J,n}(q + \Delta q)/W_{J,n}(q)$ has to be considered as a given, fixed function of $q$. The multi-overlap histogram $P_{J,n}(q)$, however, is always a noisy estimator whose statistical errors can be estimated by $\sqrt{P_{J,n}(q)}$ (neglecting auto- and cross-correlations and assuming unnormalized histograms counting the number of hits). By taking the logarithm in Eq. (6) it is then straightforward to obtain the squared error on \[ \epsilon_{\ln R_J(q)}^2 = 1/p(q) = 1/P_{J,n}(q) + 1/P_{J,n}(q + \Delta q) . \]
We now have two noisy estimators, $R_J(q)$ and $R_{J,n}(q)$ (with squared inverse errors $p(q)$ and $p_n(q)$), which may be linearly combined to yield an optimized estimator
\[ \ln R_{J,n+1}(q) = \kappa_n(q) \ln R_J(q) + [1 - \kappa_n(q)] \ln R_{J,n}(q), \]
with
\[ \kappa_n(q) = \frac{p(q)}{p(q) + p_n(q)}, \tag{8} \]
such that the statistical error of the linear combination is minimized. By exponentiating the optimized estimator and using Eq. (6), we finally arrive at the recursion
\[ R_{J,n+1}(q) = R_J(q)^{\kappa_n(q)} R_{J,n}(q)^{1-\kappa_n(q)} \]
\[ = R_{J,n}(q) \left[ \frac{P_{\mu J,n}(q)}{P_{\mu J,n}(q + \Delta q)} \right]^{\kappa_n(q)} , \tag{9} \]
\[ p_{n+1}(q) = p(q) + p_n(q). \tag{10} \]

Once $W_J(q) = \exp(S_J(q))$ is determined and kept fixed, the system is equilibrated and the data production can be performed.

We measure the (pseudo-) dynamics of the multi-overlap algorithm in terms of the autocorrelation time $\tau_{\mu J}$ which is defined by counting the average number of sweeps it takes to complete the cycle $q = 0 \rightarrow |q| = 1 \rightarrow 0$. Adopting the usual terminology\textsuperscript{20,21} for a first-order phase transition, we shall call such a cycle a “tunnelling” event. The weight iteration was stopped after at least 10 “tunnelling” events occurred, and in the production runs we collected at least 20 “tunnelling” events. To allow for standard reweighting in temperature we stored besides $P_J(q)$ and the time series of $q$ also the energies and magnetizations of the two replicas. The number of sweeps between measurements was adjusted by an adaptive data compression routine to ensure that each time series consists of $2^{16} = 65536$ measurements separated by approximately $\tau_{\mu J}$ sweeps.

Figure 4. Illustration of the relation between canonical densities $P_{\mu J}^{\text{can}}(q)$ depicted in Fig. 3 and ideally flat multi-overlap densities $P_{\mu J}^{\text{muq}}(q)$.
3 Results

Due to the large number of realizations to be simulated, the final results are relatively costly. The studied cases are summarized in Table 1, where also the simulation temperatures are given: $T = 1 \approx 0.88 T_c$ and $T = 1.14 \approx T_c$ in 3D, and $T = 1/0.6 \approx 0.85 T_c$ in 4D. The $J_{ik}$ realizations were drawn using the pseudo random number generators RANMAR\textsuperscript{22} and RANLUX\textsuperscript{23, 24} (luxury level 4). For the spin updates we always employed the faster RANMAR generator.

By fitting the averaged autocorrelation times to the power-law ansatz

$$\ln(\tau_{J_{\mu\nu}}^{\text{aut}}) = a + z \ln(N),$$

we obtained\textsuperscript{15} $z = 2.32(7)$ and $z = 1.94(2)$ in the 3D and 4D spin-glass phase, respectively. Even though the quality of the fits is quite poor and an exponential behaviour can hardly be excluded, they clearly show that the slowing down is quite off from the theoretical optimum $z = 1$ one would expect if the multi-overlap autocorrelation time $\tau_{J_{\mu\nu}}^{\text{aut}}$ is dominated by a random-walk behaviour between $q = -1$ and $+1$. In multicanonical simulations of the 3D model with broad energy histograms an even larger exponent of $z = 2.8(1)$ has been observed\textsuperscript{25}. The large values of $z$ suggest that the canonical overlap barriers are not the exclusive cause for the slowing down of spin-glass dynamics below the freezing point, i.e., the projection of the multi-dimensional state space onto the $q$-direction hides important features of the free-energy landscape of the model.

3.1 Free-Energy Barriers $F_B^q$

To define effective free-energy barriers $F_B^q$ we first construct an auxiliary 1D Metropolis-Markov chain\textsuperscript{26} which has the canonical $P_J(q)$ probability density as its equilibrium distribution. The tridiagonal transition matrix

$$T = \begin{pmatrix}
1 - w_{2,1} & w_{1,2} & 0 & \cdots \\
w_{2,1} & 1 - w_{1,2} - w_{3,2} & w_{2,3} & \cdots \\
0 & w_{3,2} & 1 - w_{2,3} - w_{4,3} & \cdots \\
\vdots & \vdots & \vdots & \ddots
\end{pmatrix}$$

is given in terms of the probabilities $w_{i,j} = \frac{1}{2} \min\left(1, P_J(q_i)/P_J(q_j)\right)$ ($i \neq j$) for jumps from state $q = q_j$ to $q = q_i = i/N$ in steps of $\Delta q = \pm 2/N$ or 0. Since $T$ fulfills the detailed balance condition (with $P_J$) it has only real eigenvalues. The largest eigenvalue $\lambda_0$
Figure 5. Distribution function $F_Q$ (14) for the 3D overlap barriers (13) (left) and the energy (right) in the spin-glass phase at $T = 1$ in units of their median values.

equals unity and is non-degenerate. The second largest eigenvalue $\lambda_1$ determines the autocorrelation time of the chain (in units of sweeps),

$$\tau_Q^q = -\frac{1}{N \ln \lambda_1} \approx \frac{1}{N(1 - \lambda_1)},$$

which we use to define the associated effective free-energy barrier in the overlap parameter $q$ as

$$F_B^q \equiv \ln(\tau_Q^q).$$

Our finite-size scaling (FSS) analyses of the thus defined overlap barriers are based on the (cumulative) distribution function $F(x)$. More precisely, we constructed a peaked distribution function $F_Q(x)$ by reflecting $F(x)$ at its median value 0.5,

$$F_Q(x) \equiv \begin{cases} F(x) & \text{for } F(x) \leq 0.5, \\ 1 - F(x) & \text{for } F(x) \geq 0.5. \end{cases}$$

For self-averaging data the function $F_Q$ collapses in the infinite-volume limit to $F_Q(x) = 0.5$ for $x = [x]_{av}$ and 0 otherwise. For non-self-averaging quantities the width of $F_Q$ stays finite. The concept carries over to quantities which diverge in the infinite-volume limit, when for each lattice size scaled variables $x/x_{med}$ are used, where $x_{med}$ denotes the median defined through $F(x_{med}) = F_Q(x_{med}) = 0.5$.

The behaviour of $F_Q(F_B^q/F_{B_{med}})$ shown on the l.h.s. of Fig. 5 for the 3D case at $T = 1$ clearly suggests that $F_B^q$ is a non-self-averaging quantity. In 4D the evidence is even stronger than in 3D. Non-self-averaging was also observed for the autocorrelation times $\tau_{muq}$ of our algorithm while the energy is an example for a self-averaging quantity; cf. the r.h.s. of Fig. 5. For non-self-averaging quantities one has to investigate many samples and should report the FSS behaviour for fixed values of the cumulative distribution function $F$.

In Ref. 15 we performed FSS fits for $F = i/16$, $i = 1, \ldots, 15$, assuming an ansatz suggested by mean-field theory,

$$F_B^q = a_1 + a_2 N^{1/3},$$

corresponding to $\tau_B^q \propto e^{a_2 N^{1/3}}$. In both dimensions the goodness-of-fit parameter $Q$ turned out to be unacceptably small. The 3D fits are depicted on the l.h.s. of Fig. 6. We
therefore also tried fits to the ansatz

$$F_B^q = \ln(c) + \alpha \ln(N) ,$$

(16)

corresponding to $\tau_B^q \propto N^\alpha$; cf. the r.h.s. of Fig. 6. Since in 3D as well as in 4D the average $Q$-value is now within the statistical expectation, the latter ansatz (16) is strongly favoured over the mean-field prediction (15). As a function of $F (= 1/16 - 15/16)$ the exponent $\alpha = \alpha(F)$ in the power law (16) varies smoothly from 0.8 to 1.1 in 3D and from 0.8 to 1.3 in 4D. A similar analysis\(^{15}\) for the autocorrelation times $\tau_{\mu q}^J$ of the multi-overlap algorithm gives exponents $\alpha(F)$ which are larger, $\alpha_{\mu q}^{\text{mult}}(F) \approx \alpha_B^q(F) + 1$. This is in agreement with our observation that other relevant barriers exist, which cannot be detected in the overlap parameter $q$.

### 3.2 Averaged Probability Densities $P(q)$

The averaged canonical densities $P(q)$ of the 3D model are shown in Fig. 7 for both $T = 1 \approx 0.88T_c$ and $T = 1.14 \approx T_c$. At least close to $T_c$ one expects that, up to finite-size corrections, the probability densities scale with system size. A method to confirm this visually is to plot $P'(q) \equiv \sigma P(q)$ versus $q' = q/\sigma$, where $\sigma$ is the standard deviation. By fitting the standard deviation to the expected FSS form $\sigma = c_1 L^{-\beta/\nu}$ we obtained\(^{17}\),

$$\frac{\beta}{\nu} = 0.312(4) , \quad Q = 0.32 \quad \text{for} \quad T = 1.14 , \quad \text{and} \quad (17)$$

$$\frac{\beta}{\nu} = 0.230(4) , \quad Q = 0.99 \quad \text{for} \quad T = 1 . \quad (18)$$

On the l.h.s. of Fig. 8 we show the scaling plot\(^{17}\) for $T = 1.14$ which demonstrates that the five probability densities collapse onto a single master curve.

### 3.3 Tails of $P(q)$

The multi-overlap algorithm becomes particularly powerful when studying the tails of the probability densities which are highly suppressed compared to the peak values; see the
r.h.s. of Fig. 8 which shows \( P(q) \) at \( T = 1.14 \) over more than 150 orders of magnitude. Based on the replica mean-field approach, theoretical predictions for the scaling behaviour of the tails have been derived by Parisi and collaborators\(^{31}\). They showed that

\[
P(q) = P_{\text{max}} f(N(q - q_{\text{max}})^x) \quad \text{for} \quad q > q_{\text{max}}^\infty
\]

and concluded more quantitatively that

\[
P(q) \sim \exp \left[ -c_1 N (q - q_{\text{max}}^\infty)^x \right] \quad \text{for} \quad N (q - q_{\text{max}}^\infty)^x \quad \text{large},
\]

(19)

with a mean-field exponent of \( x = 3 \). By allowing for an overall normalization factor \( c_0^{(N)} \) and taking the logarithm twice we have performed fits of the form\(^{32,33}\)

\[
Y \equiv \ln \left[ -\ln \left( P/c_0^{(N)} \right) \right] - \ln N = \ln c_1 + x \ln (q - q_{\text{max}}^\infty).
\]

(20)

Leaving the exponent \( x \) as a free parameter, we arrived at the estimates \( x = 12(2) \) in 3D at \( T = 1 \) and \( x = 5.3(3) \) in 4D at \( T = 1/0.6 \) which are both much larger than the mean-field value of \( x = 3 \).

By looking for reasonable alternatives we realized that for many other systems the statistics of extremes as pioneered by Fréchet, Fisher and Tippert, and von Mises\(^{34,35}\) has led to a good ansatz with universal properties\(^{36,37}\). It is based on a standard result\(^{34,35}\), due to Fisher and Tippert, Kawata, and Smirnow, for the universal distribution of the first, second, third, \ldots smallest of a set of \( N \) independent identically distributed random numbers. For an appropriate, exponential decay of the random number distribution their probability densities are given by the Gumbel form

\[
f_\alpha(x) = C_\alpha \exp \left[ a \left( x - e^x \right) \right],
\]

(21)

in the limit of large \( N \). The exponent \( a \) takes the values \( a = 1, 2, 3, \ldots \), corresponding, respectively, to the first, second, third, \ldots smallest random number of the set, \( x \) is a scaling variable which shifts the maximum value of the probability density to zero, and \( C_\alpha \) is a normalization constant. For certain spin-glass systems the possible relevance of this universal distribution has been pointed out by Bouchaud and Mézard\(^{38}\), and for instance also in applications to the 2D XY model in the spin-wave approximation\(^{36,37}\) the Gumbel ansatz (21) fitted well, albeit with a modified value of \( a = \pi/2 \).
In our case we set $x = b(q' - q_{\text{max}}')$ and modified the first $x$ on the r.h.s. of (21) to $c \tanh(x/c)$, where $c > 0$ is a constant, in order to reproduce the flattening of the densities towards $q' = 0$. Notice that in the tails of the densities, i.e. for large, positive $x$, this term is anyway subleading. A symmetric expression for $P'(q')$ reflecting its $q' \rightarrow -q'$ invariance is obtained by multiplying the above construction with its reflection about the $q' = 0$ axis,

$$P'(q') = C \exp \left\{ a \left[ c \tanh \left( \frac{b}{c} (q' - q_{\text{max}}') \right) - e^{+b(q' - q_{\text{max}}')/c} \right] \right\} \times (q' \rightarrow -q'). \quad (22)$$

Of course, the important large-$x$ behaviour of Eq. (21) is not at all affected by our manipulations.

By fitting this ansatz to our 3D data we obtained final estimates of $a = 0.448(40)$ for $T = 1.14$ and $a = 0.446(37)$ for $T = 1$, respectively. For $T = 1.14$ our best fit is already included on the l.h.s. of Fig. 8. We see a good consistency between the data and the fit over a remarkably wide range of $q'$. Even more impressive is the excellent agreement in the tails of the densities. Taking the $T = 1.14$, $L = 16$ result at face value, we find a very good fit over a remarkable range of $200/\ln(10) \approx 87$ orders of magnitude!

4 Summary and Conclusions

Employing non-Boltzmann sampling with the multi-overlap MC algorithm we have investigated the probability densities $P_J(q)$ of the Parisi order parameter $q$. The free-energy barriers $F_B(q)$ as defined in Eq. (13) turn out to be non-self-averaging. The logarithmic scaling ansatz (16) for the barriers at fixed values $F$ of their cumulative distribution function is found to be favoured over the mean-field ansatz (15). Further, relevant barriers are still reflected in the autocorrelations of the multi-overlap algorithm.

The averaged densities exhibit a pronounced FSS collapse onto a common master curve even in the spin-glass phase. For the scaling of their tails towards $q = \pm 1$ we find qualitative agreement with the decay law predicted by mean-field theory, but with an exponent
that is, in particular in 3D, much larger than theoretically expected. A much better fit over more than 80 orders of magnitude is obtained in 3D by using a modified Gumbel ansatz, rooted in extreme order statistics. The detailed relationship between the EAI spin-glass model and extreme order statistics remains to be investigated, and it is certainly a challenge to extend the work of Bouchaud and Mézard to the more involved scenarios of the replica theory.

Acknowledgements

We would like to thank A. Aharony, K. Binder, F. David, E. Domyany and A. Morel for useful discussions. The project was partially supported by the German-Israel-Foundation (GIF-I-653-181.14/1999), the US Department of Energy (DOE-DE-FG02-97ER41022), and the T3E computer-time grants hmz091 (NIC, Jülich), p526 (CEA, Grenoble) and (in an early stage of the project) bvpl01 (ZIB, Berlin).

References


Many-Body Correlation Effects in Photoexcited
Semiconductor Heterostructures

Torsten Meier, Christian Sieh, Stefan Weiser, Matthias Reichelt,
Christoph Schlichenmaier, Sven Siggelkow, Peter Thomas, and Stephan W. Koch

Department of Physics and Material Sciences Center, Philipps University
Renthof 5, 35032 Marburg, Germany
E-mail: {Torsten.Meier, Christian.Sieh, Stefan.Weiser, Matthias.Reichelt,
Christoph.Schlichenmaier, Sven.Siggelkow, Peter.Thomas,
Stephan.W.Koch} @physik.uni-marburg.de

A microscopic many-body theory describing the optical properties of semiconductors and semiconductor heterostructures is briefly reviewed. The optical response is described by the Semiconductor Bloch Equations, which include many-body effects arising from the Coulomb interaction among the photoexcited carriers. It is demonstrated that, in particular, many-body correlation contributions beyond the Hartree-Fock level strongly influence the optical response and thus have to be treated properly. The microscopic theory is able to predict and reproduce a number of experimentally observable effects. Many of the numerical solutions could only be obtained by using massively parallel computer programs which were run on the Cray T3E system in Jülich.

1 Introduction

To properly describe the optical properties of semiconductors one has to treat the dynamics of the light field, the material excitations, and their interactions. Since it is known from classical electrodynamics that the electric field couples to the polarization of a medium, the description of the material has to include this polarization as a key quantity.

In dipole approximation, the light-matter interaction is proportional to the scalar product of electric field and polarization. For a semiconductor that is excited with optical light fields the induced polarization of the material is due to interband transitions from the valence to the conduction bands. For the linear optical absorption spectrum of an unexcited semiconductor, which consists of completely filled valence and empty conduction bands, it is sufficient to consider only these interband transitions and no further material excitations are required. Already in the linear spectra the Coulomb interaction leads to characteristic signatures. These so called excitonic effects are a consequence of the Coulomb attraction between photoexcited electrons and holes.

If the semiconductor is not in the ground state before the photoexcitation, or if the exciting light fields are not in the low-intensity limit, one has to go beyond the linear regime and additional material quantities are required to consistently compute the nonlinear optical response. The Coulomb interaction among the optically excited carriers introduces a many-body problem and thus usually no fully exact treatment of the material response is possible. During the last decade, however, a number of approximation schemes have been developed and applied successfully to different excitation regimes.

One way to consistently include the Coulomb interaction in a theory for the optical response of semiconductors is to treat the many-body terms on the level of the Hartree-Fock approximation. On this level the dynamic material response is described in terms of
the diagonal and off-diagonal components of the reduced single-particle density matrix. Whereas the latter terms are given by interband transitions which determine the optical polarization the former terms are the carrier occupation probabilities in the different bands. On this level the dynamics of optically induced excitations can be described by coupled equations of motion for the reduced single-particle density matrix. These are well-known as the Hartree-Fock Semiconductor Bloch Equations which include Coulombic effects via bandgap and field renormalization\textsuperscript{1,2}.

In many-particle physics the term correlation energy refers to ground state energy contributions beyond the Hartree-Fock level. Similarly, in semiconductor optics all aspects of the nonlinear optical semiconductor response that cannot be described on the basis of a Hartree-Fock treatment are called correlation effects. To theoretically describe such many-body correlations different schemes have been developed.

The second Born approximation\textsuperscript{2–5} includes terms which are of second-order in the Coulomb interaction and keeps the reduced single-particle density matrix as the dynamic object. It is applicable to the regime of strong fields and is well suited to describe exciton saturation and excitation-induced dephasing processes due to a plasma of free carriers\textsuperscript{7,9,10}.

A number of experimentally observable signatures like transitions from excitons to bound bieexcitons which are made of two attractively interacting electron-hole pairs, can, however, not be described just on the basis of the reduced single-particle density matrix. One way to treat such processes is to include higher-order, i.e. four- (and more) particle correlations, in the theoretical analysis and to classify the nonlinear optical response in powers of the optical fields\textsuperscript{11} using the so-called dynamics-controlled truncation scheme\textsuperscript{12,13}. Restricting the analysis to a finite order in the interaction with the optical field closes the many-body hierarchy at a particular level. Since this approach is based on an expansion in powers of the fields, it is limited to the regime of not too strong fields. Nevertheless this scheme has been successfully used to analyze various nonlinear optical experiments performed on semiconductor and semiconductor heterostructures. The results obtained in this context within our project have been described in some reviews\textsuperscript{6–8,14} and a number of original publications\textsuperscript{15–25}.

In Sect. 2 we give a brief review of the microscopic theory that is able to describe the optical response of semiconductors. Some numerical results that are obtained using the dynamics-controlled truncation scheme, which in particular highlight the role of excitonic correlations, are presented in Sect. 3. The paper closes with a short summary in Sect. 4.

2 The Semiconductor Bloch Equations

2.1 Hartree-Fock Approximation

The dynamic optical response of the material is described by the macroscopic optical polarization $P$. For ordered systems it is convenient to expand the $P$ into a Bloch basis\textsuperscript{1}

\[ P = \sum_{k,e,h} (d_{eh}^c)^* P_{eh}^c + \text{c.c.}, \quad (1) \]

where $d_{eh}^c$ is the electron-hole interband dipole matrix element between conduction band $e$ and valence band $h$. In terms of electron ($a_k^e, a_k^c$) and hole ($b_k^h, b_k^c$) creation and destruction
operators the microscopic polarization $P_{k}^{ch}$ and the carrier occupation probabilities $f_{k}^{ee,hh}$ are the diagonal and off-diagonal elements of the reduced single-particle density matrix:

$$
\begin{pmatrix}
\langle a_{k}^{\dagger} a_{k} \rangle \\
\langle b_{-k}^{\dagger} b_{-k} \rangle
\end{pmatrix}
=
\begin{pmatrix}
f_{k}^{ee} \\
f_{k}^{hh}
\end{pmatrix}
\begin{pmatrix}
P_{k}^{ch} \\
(P_{k}^{ch})^{*}
\end{pmatrix}.
$$

(2)

The standard many-body Hamiltonian considered in semiconductor optics is\textsuperscript{1}

$$
H = H_{0} + H_{C} + H_{f}.
$$

(3)

Here,

$$
H_{0} = \sum_{k,e} \varepsilon_{k}^{e} a_{k}^{\dagger} a_{k} + \sum_{k,h} \varepsilon_{k}^{h} b_{-k}^{\dagger} b_{-k}
$$

(4)

is the single-particle Hamiltonian containing the bandstructure ($\varepsilon_{k}^{e}$ and $\varepsilon_{k}^{h}$),

$$
H_{C} = \frac{1}{2} \sum_{k,k',q\neq 0} V_{q} a_{k+q}^{\dagger} a_{k'}^{\dagger} a_{k'} a_{k} + \frac{1}{2} \sum_{k,k',q\neq 0} V_{q} b_{k+q} b_{k'}^{\dagger} b_{k'} b_{k}
$$

(5)

describes the Coulomb interaction among the photoexcited carriers, and

$$
H_{f} = -E(t) \sum_{k,e,h} (d_{k}^{e} a_{k}^{\dagger} b_{-k}^{\dagger} + (d_{k}^{h})^{*} b_{-k} a_{k})
$$

(6)

denotes the dipole interaction with a classical electromagnetic field\textsuperscript{1}.

The Coulomb interaction, Eq. (5), is given by the sum of three terms describing the repulsion among electrons and holes and the attraction between electrons and holes. $V_{q}$ is the Fourier transform of the Coulomb interaction potential. The light field, Eq. (6), creates or destroys pairs of electrons and holes. In all equations the operators $a_{k}$ and $b_{k}$ as well as the corresponding creation operators refer to bands $e$ and $h$, respectively, and similarly $a_{k}'$ and $b_{k}'$ belong to $e'$ and $h'$.

The Heisenberg equation of motion for $P_{k}^{ch}$ reads\textsuperscript{1}

$$
i\hbar \frac{\partial}{\partial t} P_{k}^{ch} = [ P_{k}^{ch}, H ] = -(\varepsilon_{k}^{e} + \varepsilon_{k}^{h}) P_{k}^{ch} + (d_{k}^{e} - f_{k}^{ee} d_{k}^{e h} - f_{k}^{hh} d_{k}^{h h}) E
$$

+ \sum_{q \neq 0} V_{q} P_{k-q}^{ch}
- \sum_{q \neq 0, k'} V_{q} \left[ \langle a_{k+q}^{\dagger} a_{k'}^{\dagger} b_{k+q} b_{k'} a_{k} a_{k}^{\dagger} b_{k} b_{k}^{\dagger} a_{k} b_{k}^{\dagger} a_{k} \rangle
- \langle a_{k+q}^{\dagger} a_{k}^{\dagger} b_{k+q}^{\dagger} b_{k} a_{k}^{\dagger} b_{k} b_{k}^{\dagger} a_{k} b_{k}^{\dagger} a_{k} \rangle
+ \langle a_{k+q}^{\dagger} b_{k+q}^{\dagger} b_{k} b_{k}^{\dagger} a_{k} b_{k}^{\dagger} a_{k} \rangle
- \langle a_{k+q}^{\dagger} b_{k+q} b_{k}^{\dagger} b_{k} a_{k} b_{k}^{\dagger} a_{k} \rangle \right],
$$

(7)

where $e = e'$ and $h = h'$ in the two-band case with only one relevant conduction and valence band. In the more general multiband configuration the primed superscripts correspond to summations over all the relevant bands.

Besides the coupling of the components of the density matrix among themselves (via $H_{0} + H_{f}$) Eq. (7) and the corresponding equations for the carrier occupation probabilities contain terms that couple the two-operator terms to four-operator terms (via $H_{C}$). This is the beginning of the well known many-body hierarchy: If one writes equations of motion
for the four-operator expectation values, one obtains a coupling to six-operator terms, and so on. To close the coupled set of equations, one has to truncate this hierarchy at some level in a self consistent fashion.

The dynamical Hartree-Fock approximation uses a decoupling scheme like

\[
\langle a_k^+ a_{k'} a_{k-q} a_{k'} \rangle \simeq \langle a_k^+ a_k \rangle \langle a_{k-q} a_{k-q} \rangle \delta_{k,k'} ,
\]

where no other contribution appears since \( q \neq 0 \). Applying this procedure to all terms one obtains the well-known Hartree-Fock Semiconductor Bloch Equations. For a two-band system these are given by

\[
\begin{align*}
\frac{i \hbar}{\partial t} - \epsilon_k^e(t) - \epsilon_h^h(t) & = \left[ 1 - f_{k}^{ee}(t) - f_{k}^{hh}(t) \right] \Omega_k(t) , \\
\frac{\partial}{\partial t} f_{k}^{aa}(t) & = - \frac{2}{\hbar} \text{Im} \left[ \Omega_k(t)(P_{k}^{eh}(t))^\dagger \right] ,
\end{align*}
\]

where

\[
\Omega_k(t) = \epsilon^{eh} E(t) + \sum_{k' \neq k} V_{k-k'} P_{k'}^{eh}(t) , \quad \epsilon_k^e(t) = \epsilon_k^a - \sum_{k' \neq k} V_{k-k'} f_{k'}^{aa}(t)
\]

are the renormalized field and transition energy, respectively. These Hartree-Fock or exchange renormalizations couple all \( k \)-states of the semiconductor and furthermore introduce optical nonlinearities in the Hartree-Fock Semiconductor Bloch Equations. The renormalization contributions are no small corrections, with the field renormalization which includes the excitonic effects observable in linear spectra being the leading-order Coulomb effect. The nonlinearities in the Hartree-Fock Semiconductor Bloch Equations, Eq. (9), arise due to phase-space filling (terms proportional to the occupation probabilities times the field \( E_f \)), as well as energy and field renormalization. These nonlinearities are a consequence of the Fermionic nature of the electrons and holes (Pauli blocking and Fermionic exchange).

### 2.2 Dynamics Controlled Truncation

If the analysis of the nonlinear optical response is restricted to a finite order in the interaction with the light field\(^{11}\) the many-body hierarchy of equations of motion closes at a certain stage\(^{12,13}\). Such a procedure establishes a systematic truncation scheme of the Coulombic many-body correlations for purely coherent optical excitation configurations. In the following we present the dynamic equations as obtained in the coherent \( \chi^{(3)} \) limit\(^{12,13,26}\).

In order to be able to distinguish between the Hartree-Fock and the correlation contributions it is convenient to define the correlation function

\[
B_{k,k',k'',k^{\prime\prime}}^{eh',eh} = \langle a_k^+ b_{k'}^+ a_{k''}^+ b_{k^{\prime\prime}} \rangle - \langle a_k^+ b_{k'}^+ \rangle \langle a_{k''}^+ b_{k^{\prime\prime}} \rangle - \langle a_k^+ b_{k'}^{\dagger} \rangle \langle a_{k''}^+ b_{k^{\prime\prime}}^{\dagger} \rangle - \langle a_k^+ b_{k'}^{\dagger} \rangle \langle a_{k''}^+ b_{k^{\prime\prime}} \rangle
\]

such that the polarization equation in third order can be written as\(^{26}\)

\[
\frac{\partial}{\partial t} P_{k}^{eh} = \frac{\partial}{\partial t} P_{k}^{eh}|_{\text{hom}} + \sum_{n=1}^{3} \frac{\partial}{\partial t} P_{k}^{eh}|_{\text{inhom},n} ,
\]

318
contributions to the polarization equation, see Eq. (16), consist of four terms of the structure is coupled to higher-order correlation functions. see that due to the many-body Coulomb interaction the two-particle electron-hole ampli-
Figure 1. Polarization-dependent differential absorption in optical Stark effect configuration pumping 4.5 meV below the 1s exciton resonance. The origin of the energy scale coincides with the spectral position of the 1s exciton (from Ref. [15]).

consist of the Coulomb interaction potential times products of two polarizations, $VPP$, see Eq. (19).

3 Numerical Results

On the basis of Eqs. (12) and (17) and its extensions accounting for multiple valence bands and higher-order interactions with the field a variety of optical excitation configurations such as pump-probe or four-wave-mixing experiments have been analyzed. In the following we describe three examples where it has been shown that the dynamics of many-body correlations are of particular importance\textsuperscript{15, 18, 20}.

3.1 Excitonic Optical Stark Effect

A pump-probe experiment is performed using two laser pulses, the pump and the probe. One then monitors the changes of the absorption of the probe pulse induced by the excitation with the pump. This differential absorption $\delta \alpha(\omega)$ is proportional to the Fourier-transform of the pump-induced differential polarization $\delta P(\omega)$.\textsuperscript{1}

Theoretical results for the light-polarization dependent excitonic optical Stark effect, i.e. the induced spectral changes around the exciton for detuned optical pumping below the resonance, are displayed in Fig. 1. For co-circularly and for linearly polarized pump and probe beams one obtains a spectral blue shift of the exciton reflected in the dispersive-type absorption changes that are positive above and negative below the exciton. This blue shift is the “classical” result well-known from a simple two-level system that is pumped spectrally below its resonance frequency and can simply be understood in terms of level repulsion between the light field and the optical resonance. However, for cross-circularly polarized pulses ($\sigma^+\sigma^-$) the sign of the shift is reversed, i.e. a spectral shift of the exciton
Pauli-blocking

\[XX=XY\]

\[\tau = 2.5\text{ps}\]

\[4\text{ps}\]

\[6\text{ps}\]

Figure 2. Polarization-dependent time-resolved four-wave-mixing signals for a disordered semiconductor nano-structure using pulse delays \(\tau = 2.5, 4,\) and \(6\text{ps}\). The disorder strength is much weaker than the exciton binding energy. For the upper panel many-body correlations are neglected, whereas for the lower panels both disorder and Coulomb correlations have been considered on a microscopic level (from Ref. [20]).

towards the pump (red shift) is found. The calculated polarization-dependent shifts are in agreement with experiments performed on semiconductor quantum wells\(^{15}\).

For the case of cross-circularly polarized pump and probe pulses it has been shown that the signal is solely induced by many-particle correlations, since the Hartree-Fock contributions vanish as long as only heavy-hole and no light-hole transitions are relevant. This is due to the fact that the two degenerate heavy-hole excitons that can be excited with \(\sigma^+\) and \(\sigma^-\) polarized light share no common electronic state and are thus uncoupled if many-particle correlations are not taken into account\(^{15, 16, 19, 23}\). Whereas the Hartree-Fock contributions which correspond to a blue shift dominate the signal for all other configurations, for \(\sigma^+\sigma^-\) one selectively probes only the correlation terms which result in the red shift. The occurrence of this red shift is not directly related to the existence of a bound biexciton that can be excited with \(\sigma^+\sigma^-\) polarized pulses. Instead, as shown in Ref.\(^{15}\) the red shift can be identified to be a genuine consequence of the memory character (non Markovian dynamics) of the many-body correlations dynamics.

### 3.2 Disorder-Induced Dephasing

A four-wave-mixing experiment is performed using two laser pulses from directions \(k_1\) and \(k_2\) which are temporally delayed by the time delay \(\tau\). One then detects the field that is diffracted in the direction \(2k_2 - k_1\). For optically thin samples this signal is proportional to \(|P_{2k_2-k_1}(t, \tau)|^2\), see Ref.\(^1\).

In Ref.\(^{20}\) we have presented numerical solutions of the real-space version of Eqs. (12) and (17) including energetic disorder in the valence and conduction bands. The analysis of disorder effects is important since almost all epitaxy-grown semiconductor heterostructures contain some type of imperfections. For example, the disorder relevant in two-dimensional
3.3 Coherent Excitation Spectroscopy

Coherent excitation spectroscopy is an example for the so called partially non-degenerate four-wave-mixing. In coherent excitation spectroscopy pulse $k_1$ is temporally long (spec-
trally narrow), whereas pulse $k_2$ is short (spectrally broad). Since the long pulse $k_1$ excites only transitions in a narrow spectral region around its center frequency $\omega_1$, the coherent excitation spectroscopy scans obtained from the coherent signal diffracted in direction $2k_2 - k_1$ are ideally suited to investigate coherent couplings among different subsystems and also within disorder broadened lines\textsuperscript{18}. The coherent excitation spectroscopy signal is proportional to the Fourier-transform of the four-wave-mixing polarization $|P_{2k_2 - k_1}(\omega, \omega_1, \tau)|^2$ and depends on both the detection and excitation frequencies, $\omega$ and $\omega_1$, respectively.

In Fig. 3 we display the polarization dependence of the coherent excitation spectroscopy scans for linearly polarized pulses. As in Fig. 2 the differences between between $XX$ and $XY$ excitation are solely induced by many-body Coulomb correlations. For $XX$ excitation, Fig. 3(a), one finds only a single peak which occurs at the position of the exciton ($\omega = 0$) when it is resonantly excited by pulse $k_1$ ($\omega_1 = 0$). For $XY$, Fig. 3(b), however, three peaks are present. The two additional peaks occur at $\omega \approx -1.5\text{meV}$ when $\omega_1 = 0$ and at $\omega = 0$ when $\omega_1 \approx -1.5\text{meV}$. $1.5\text{meV}$ is just the biexciton binding energy as has been confirmed by additional independently performed calculations. Thus the former one of the two additional peaks is simply due to the exciton-biexciton transition which shows up in four-wave-mixing after the exciton has been excited \textit{resonantly}. The peak obtained for $\omega_1 \approx -1.5\text{meV}$, however, results from an exciton-biexciton transition starting from an \textit{off-resonantly} driven exciton transition\textsuperscript{27}. More detailed investigations of the correlation signatures contained in the coherent excitation spectroscopy scans including multiple valence bands have recently been performed\textsuperscript{27}.

4 Summary

A microscopic many-body theory describing the optical and electronic properties of semiconductors and semiconductor heterostructures has been briefly reviewed. The main steps involved in the derivation of the Semiconductor Bloch Equations with many-body correlation contributions beyond the Hartree-Fock level have been discussed. The importance of these correlations for polarization-dependent pump-probe and four-wave-mixing experiments has been demonstrated by showing numerical results for a few selected examples. Many of our results, in particular those on the disorder-induced dephasing and the coherent excitation spectroscopy, could only be obtained using massively parallel computer programs which were run on the Cray T3E system in Jülich. Presently, our grant on this machine allows us to use 27500 KE (Kontingenteinheiten) per month, which corresponds to a monthly usage of 10000 hours of computer time on a single processor.

Acknowledgments

This work is supported by the Deutsche Forschungsgemeinschaft (DFG) through the Leibniz prize, the Schwerpunkt Quantenkohärenz, and the project No. KO816/8-1. We wish to thank the John von Neumann - Institut für Computing (NIC), Forschungszentrum Jülich, Germany, for grants for extended CPU time on their supercomputer systems.
References

Computer Simulation of Chiral Liquid Crystal Phases

Reiner Memmer

Fachbereich Chemie
Universität Kaiserslautern, 67663 Kaiserslautern, Germany
E-mail: memmer@rhrk.uni-kl.de

Computer simulations, e.g. the Monte Carlo (MC) and molecular dynamics (MD) method, provide an important tool to aid the understanding of the macroscopic properties of condensed matter systems and of their phase transitions. In the field of liquid crystals, a state of matter with long range orientational order, many applications e.g. in display technology are due to material properties linked to the existence of helical superstructures originating from molecular chirality. Phase diagrams of chiral calamitic liquid crystals were investigated by MC in the canonical (NVT) and the isobaric-isothermal (NpT) ensemble in order to deduce a link between macroscopic and microscopic properties, e.g. the molecular structure, starting with interactions between the molecules. A rich polymorphism of chiral liquid crystal phases was observed in dependence on temperature, pressure and a chirality parameter. In addition to the cholesteric phase, for the first time several blue phases and a phase with characteristics of a twist grain boundary phase have been observed. Recently, the discovery of chiral superstructures in liquid crystal phases of achiral so-called banana-shaped molecules attracted considerable interest. On cooling an isotropic system of banana-shaped molecules, at low temperature a biaxial smectic phase with a local polar arrangement of the steric dipoles arranged in an antiferroelectric-like superstructure could be proven. Additionally, at intermediate temperature a nematic-like phase has been formed whereas hints for a spontaneous achiral symmetry breaking have been determined. The structures of all phases were characterised by order parameters, scalar and pseudoscalar orientational correlation functions, and by visual representations of selected configurations.

1 Introduction

Liquid crystal phases, formed by molecules with a characteristic shape anisotropy, are a state of matter with long range orientational order. Many of their applications e.g. in display technology are due to material properties linked to the existence of helical superstructures originating from the chirality of mesogenic molecules which leads to a variety of chiral liquid crystal phases, see Fig. 1 for a typical phase diagram in the temperature-chirality parameter plane. Characteristic for these phases are helical superstructures in the arrangement of the molecules, i.e. the manifestation of suprastructural chirality, basing on a spontaneous twist of the molecular orientation. Well known examples for chiral liquid crystal phases without positional order are the cholesteric phase (N*) and the blue phases (BP I, BP II, and BP III), which are formed by chiral calamitic and discotic molecules, in single component systems and induced by chiral dopants dissolved in achiral liquid crystal host phases. There exist also many chiral liquid crystal phases with positional order, where the formation of both a helical superstructure and a layered structure can be realised free of defects, as e.g. in smectic S_{C_{*}} phases of calamitic molecules or yields frustrated structures, as in the case of the recently discovered twist grain boundary phases (TGB) with a helical arrangement of smectic regions. In spite of the enormous success in the characterisation of new structures of chiral liquid crystal phases, relatively little is known about the connection between the molecular chirality and the suprastructural chirality of the phases, i.e. about
the link between chirality at the microscopic and the macroscopic level, as recently pointed out by Lubensky et al.\(^4\) The determination of phase diagrams in the temperature-chirality parameter plane and of the phase structures has been subject of many experimental and theoretical investigations already since the discovery of liquid crystals, but in general it is very difficult to make predictions about the chiral liquid crystal phases or even the phase sequence appearing for a selected chiral mesogenic molecule. Recently, new types of liquid crystal phases formed by so-called banana-shaped molecules were discovered which show interesting material properties, e.g., the appearance of ferroelectricity\(^5,6\). In these systems of achiral molecules the detected formation of chiral superstructures, i.e., a spontaneous achiral symmetry breaking, is even more remarkable and not yet fully understood. Here, many possible structures are still under discussion for the chiral domains.

Investigations by means of molecular modeling allow to deduce a link between microscopic and macroscopic properties starting with model interactions between the molecules and have been successfully applied in the field of liquid crystals\(^7\), but mainly restricted to achiral liquid crystals, e.g., nematic and smectic phases. Recently, a variety of model systems for cholesteric liquid crystal phases have been studied by means of computer simulation, e.g., systems of chiral atropisomeric molecules\(^8\) and lattice systems of chiral molecules\(^9\). In the following, computer simulation results will be presented for chiral calamitic molecules described by the chiral Gay-Berne fluid\(^10,11\), the first model system
where a rich polymorphism of chiral liquid crystal phases could be proven in dependence on temperature, pressure and a chirality parameter. Additionally, a MC study of phase structures of systems composed of model Gay-Berne banana-shaped molecules will be presented\(^\text{12}\).

2 Chiral Calamitic Molecules

In order to describe chiral calamitic molecules the chiral Gay-Berne fluid has been defined as an extension of the Gay-Berne fluid, a potential taking into account the molecular anisotropy in both the shape and the attractive forces which has been successfully used to describe achiral calamitic liquid crystals\(^\text{13}\). An additive chiral interaction term is taken into account in order to include the specific feature of chiral molecules, the energetic discrimination between mirror image arrangements. The total intermolecular interaction energy

\[ U(\hat{u}_i, \hat{u}_j, r_{ij}) = aU_a(\hat{u}_i, \hat{u}_j, r_{ij}) + cU_c(\hat{u}_i, \hat{u}_j, r_{ij}), \]

where \(U_a(\hat{u}_i, \hat{u}_j, r_{ij})\) denotes the energy of the achiral interaction taken to be the Gay-Berne potential\(^\text{13}\)

\[ U_a = 4\epsilon (\hat{u}_i, \hat{u}_j, r_{ij}) \left( \frac{\sigma_0}{r_{ij} - \sigma(\hat{u}_i, \hat{u}_j, r_{ij}) + \sigma_0} \right)^{12} - \left( \frac{\sigma_0}{r_{ij} - \sigma(\hat{u}_i, \hat{u}_j, r_{ij}) + \sigma_0} \right)^{6}, \]

and \(U_c(\hat{u}_i, \hat{u}_j, r_{ij})\) is the energy of the chiral interaction given by

\[ U_c = 4\epsilon (\hat{u}_i, \hat{u}_j, r_{ij}) \left( \frac{\sigma_0}{r_{ij} - \sigma(\hat{u}_i, \hat{u}_j, r_{ij}) + \sigma_0} \right)^{7} (\hat{u}_i \times \hat{u}_j) \cdot r_{ij} (\hat{u}_i \cdot \hat{u}_j). \]

The strengths of the corresponding interaction energies are given in terms of dimensionless scalar and pseudoscalar parameters \(a\) and \(c\), respectively, i.e. the latter one, the chirality parameter, changes sign if enantiomers are considered.

In order to investigate the influence of chirality on the structure of liquid crystal phases, the phase diagram in the temperature-chirality parameter plane of the chiral Gay-Berne fluid was studied as a function of the chirality parameter \(c\) along an isotherm in the stability range of the nematic phase (N) of the achiral Gay-Berne fluid. Systems of \(N=2048\) molecules were investigated under periodic boundary conditions in a cubic box by MC in the canonical ensemble. Discontinuities of thermodynamic quantities, i.e. the energy and the second rank order parameter \(S^*\), allowed the approximate localisation of appearing phase transitions and a first characterisation of the different phases. For low values of \(c\) the observed positive values of \(S^*=0.8\) indicate a favoured orientation of the long molecular axes \(\hat{u}_i\) parallel to a space-fixed axis, the director, a behaviour characteristic for nematic phases (Fig. 2a). The formation of nematic phases is maintained until a threshold value of \(c\) is reached. Then, values of \(S^*=-0.4\) indicate a favoured orientation of the long molecular axes \(\hat{u}_i\) perpendicular to a selected space-fixed axis, i.e. a behaviour typical in a cholesteric phase with respect to the helical axis (Fig. 2b). Here, the continuous rotation of the favoured orientation of the long molecular axes along the helical axis is obvious. Finally, for higher values of \(c\) values of \(S^*=0.0\) are obtained, indicating an isotropic
character of the phase. Here, for many chiral mesogens often blue phases appear in the sequence BP I - BP II - BP III both with increasing temperature and chirality parameter (see Fig. 3a for a schematic phase diagram). The essential part in the theories of blue phases are so-called double twist cylinders characterised by a director field as sketched in Fig. 3b. Whereas in the cholesteric phase the twisting occurs only in one direction, i.e. along the helical axis, in a double twist cylinder a helical superstructure exists over a selected distance from the cylinder centre along every axis perpendicular to the cylinder axis. Different arrangements of such double twist cylinders taking into account unavoidable defects and isotropic regions have been considered in the framework of Landau-de Gennes theory. The network of interwoven double twist cylinders in a body-centred cubic structure, consistent with experimental results for BP I, is shown in Fig. 4a,b as elementary cell and in form of the resulting periodic structure. The corresponding characteristic features are obvious in the snapshot shown in Fig. 4c, where in spite of order parameters of zero regions with orientational correlations can be seen, especially if the angle between the long molecular axes and the box face normal is colour coded. There exist regions where the molecules are preferred parallel to the box face normal, whereas the preferred molecular orientation twists along all directions perpendicular to this axis moving away from these regions, i.e. helical superstructures appear but only locally and limited about a selected distance. The chiral Gay-Berne fluid is the first many particle system based on intermolecular interactions where such double twist structures, arranged here according to the structure of BP I, could be proved by computer simulation. At higher chirality parameter there still exist local double twist regions, but these are no longer infinite under the limitations of the system size nor arranged in a lattice. As in the so-called spaghetti model of blue phase III (Fig. 5a) the double twist regions are now worm-like and seem to be randomly oriented (Fig. 5b).

In general, the clear evidence should be emphasised that a chiral interaction potential proportional to the first pseudoscalar term of the expansion in rotational invariants is able
Figure 3. Blue phases: (a) Schematic phase diagram in the temperature-chirality parameter plane, (b) director field of a double twist cylinder.

Figure 4. Blue phase I (BP I): Theoretical model, characterised by (a) arrangement of double twist cylinders in the unit cell, (b) three-dimensional arrangement, (c) visualisation of a configuration (MC-NVT) of BP I observed in dependence on chirality parameter along an isotherm.

to give the phase sequence nematic, cholesteric and blue phases as characteristic for many chiral mesogenic molecules in dependence on one pseudoscalar chirality parameter.

In canonical MC simulations of chiral liquid crystals generally cholesteric phases with non-equilibrium pitch were formed under periodic boundary conditions, which require that only selected orientations of the helical axis combined with an integral number of half turns of a cholesteric phase are possible in order to commensurate with the periodic images. In
order to study the temperature influence on helical superstructures and to overcome system-size effects the phase diagram was additionally studied in the \( NpT \) ensemble. Here, the ability of the simulation box to change the dimensions during the simulation enabled the determination of the equilibrium pitch of cholesteric phases. Under isobaric conditions several phase transitions characterized by phase transition enthalpy, entropy and relative volume change have been identified in dependence on temperature along an isochiral. On heating a cholesteric phase, for the first time a temperature driven phase transition to a blue phase (BP II) could be proven by computer simulation, a characteristic feature of many experimental and theoretical phase diagrams of chiral liquid crystals. The network of interwoven double twist cylinders in a simple cubic structure, consistent with experimental results for BP II, is shown in Fig. 6a,b as elementary cell and in form of the resulting periodic structure. The corresponding characteristic features are obvious in the snapshot shown in Fig. 6c. Additionally, on cooling a cholesteric system a phase with a helical superstructure and smectic layers formed perpendicular to the helical axis has been observed. Shown in Fig. 7a-d are equidistant intersections chosen perpendicular to the helical axis. The molecular orientations are preferred perpendicular to the helical axis. Locally, the molecules have a favored orientation parallel to a common direction. In contrast to the absence of long-range positional order in a cholesteric phase, now additionally a layered structure has been formed along the local director with layer normals spiraling along the helical axis, one of the main characteristics of the molecular arrangement in a twist grain boundary (TGB) phase, shown for comparison in Fig. 7e.

The observed phase sequence TGB-like phase, cholesteric phase, blue phase and isotropic phase obtained in dependence on a scalar parameter, the temperature, corresponds well to the characteristics of many experimentally observed phase diagrams, e.g. the phase behavior of a binary non-racemic mixture of (S) and (R) enantiomers in dependence on temperature (Fig. 1).
Figure 6. Blue phase II (BP II): Theoretical model, characterised by (a) arrangement of double twist cylinders in the unit cell, (b) three-dimensional arrangement, (c) visualisation of a configuration (MC-$NpT$) of BP II observed in dependence on temperature along an isochoiral.

Figure 7. Characterization of a twist grain boundary phase (TGB): (a-d) Snapshot of the molecular organization in the TGB-like phase of the chiral Gay-Berne fluid visualized by equidistant intersections chosen perpendicular to the helical axis (MC-$NpT$); (e) molecular arrangement in a TGB $A^*$ phase (Fig. from$^{14}$).
3 Achiral Banana-Shaped Molecules

The phase behaviour of achiral banana-shaped molecules described by model intermolecular interactions based on the Gay-Berne potential was studied by computer simulation. The characteristic molecular structure was considered by joining two calamitic Gay-Berne particles through a bond to form a biaxial molecule of point symmetry group $C_{2v}$ (Fig. 8a) which has a steric dipole along the polar axis intersecting the bending angle. Systems of $N=1024$ rigid banana-shaped molecules with a bending angle of $140^\circ$ have been studied in dependence on temperature by MC simulations in the $NpT$ ensemble.

![Figure 8. (a) Banana-shaped molecules: (above) chemical structure of a typical molecule, (below) a banana-shaped Gay-Berne molecule of point symmetry group $C_{2v}$ with a steric dipole along the polar axis, defined by linking two equivalent calamitic Gay-Berne particles, (b) snapshot of the molecular organisation in the isotropic phase formed at high temperature.](image)

On cooling down an isotropic system (Fig. 8b), two phase transitions characterised by transition enthalpy, entropy and relative volume change have been observed. For the first time, at low temperature an untilted smectic phase showing a global phase biaxiality and a spontaneous local polarisation in the layers, i.e., a local polar arrangement of the steric dipoles, with an antiferroelectric-like superstructure could be proven by computer simulation (Fig. 9b). The clear evidence should be emphasised that the spontaneous local polarisation and the antiparallel polarisation of neighbouring smectic layers are due to the anisotropy in shape and attractive interactions of the banana-shaped molecules with a steric dipole only without considering dipole-dipole interactions. Except the in-plane positional order the layered phase shows the characteristic structural features of the $C_{P\Lambda}$-phase, as recently observed experimentally for the first time for an achiral banana-shaped molecule having an achiral orthogonal biaxial smectic phase exhibiting antiferroelectric switching behaviour\textsuperscript{15}. Additionally, at intermediate temperature a nematic-like phase has been proved (Fig. 9a), a feature which has been observed in systems of banana-shaped molecules only in exceptional cases. Here, close to the phase transition to the smectic phase hints for a spontaneous achiral symmetry breaking have been determined. A helical superstructure has been formed, with a structure (Fig. 9c) as recently predicted for achiral banana-shaped molecules considering the effects of a negative bend elastic constant with respect to the formation of nematic phases with spatially modulated director by means of a Landau-like model\textsuperscript{16}.  

332
Figure 9. Snapshot of the molecular organisation (a) in the nematic phase and (b) in the biaxial smectic phase with a local polar arrangement of the steric dipoles arranged in an antiferroelectric-like superstructure, (c) theoretical model for nematic phases of banana-shaped molecules with spatially modulated director (Fig. from 16).

4 Conclusions

Molecular systems of anisotropic molecules have been studied by computer simulation in order to figure out relationships between molecular structure and phase properties focusing especially to the formation of liquid crystal phases with suprastructural chirality. A rich polymorphism of liquid crystal phases was characterised in dependence on temperature, pressure, chirality parameter and molecular structure. A future task is now to extend the study to dynamical properties, especially taking into account the effects of external fields and surfaces.

Acknowledgments

Generous allocation of computer time by the John von Neumann-Institut für Computing, Jülich, where the research project “Computer simulation of chiral liquid crystal phases” has been funded by about 6000 CPU-hours on the CRAY T90 during the period 1996-2001, and the Regionales Hochschulrechenzentrum Kaiserslautern and financial support from the Deutsche Forschungsgemeinschaft and the Fonds der Chemischen Industrie are gratefully acknowledged.
References


Simulations of Liquid Crystals: Bulk Structure and Interfacial Properties

N. Akino¹, G. Germano¹, N. H. Phuong¹, F. Schmid¹, and M. P. Allen²

¹ Theoretische Physik, Universität Bielefeld
33501 Bielefeld, Germany
E-mail: {germano, phuong, schmid}@physik.uni-bielefeld.de
hiroshi.akino@nifty.ne.jp

² Centre for Scientific Computing, University of Warwick
Coventry CV4 7AL, United Kingdom
E-mail: M.P.Allen@warwick.ac.uk

We present large scale molecular dynamics simulations of liquid crystals, which are modeled as fluids of soft repulsive ellipsoidal molecules. In the first part of the paper, we discuss the bulk structure of nematic liquid crystals. The direct correlation function (DCF) has been determined for the first time in a nematic fluid without any approximations. We demonstrate that it can be used to calculate the Frank elastic constants, which are important phenomenological parameters in the continuum theory of liquid crystals. In the second part, we consider an interface between a nematic and an isotropic phase. The interplay between the surface tension and the elastic interactions in the nematic phase leads to an unusual fluctuation spectrum.

1 Introduction

Liquid crystals are beautiful examples of materials that are fascinating from a fundamental point of view – as complex fluids with intriguing symmetries and unusual elastic properties – and highly interesting for technological applications¹,². Like in most materials, many important properties depend on the structure of inhomogeneities and interfaces in the material. For example, the alignment of nematic liquid crystal on surfaces plays a key role in the domain of liquid crystal display devices³⁻⁵. For the theoretical physicist, interfaces in liquid crystals are appealing for yet another reason: Surfaces and interfaces break two continuous symmetries, the isotropy of space and the translational invariance. The interplay of this symmetry breaking with the broken symmetries in the various liquid crystal phases leads to a wealth of new intriguing phenomena⁶⁻¹⁰. Apart from discussing these effects from a phenomenological point of view, it is also interesting to investigate how they come about on a microscopic level, and to study the relation between the local structure of liquid crystals and relevant “mesoscopic” phenomenological parameters.

In this paper, we review some computer simulation work that we have done in this direction. We shall discuss the local structure and interfacial properties in two of the most important liquid crystal phases: The isotropic phase (I), which is an ordinary fully symmetric fluid phase, and the nematic phase (N), where the fluid has translational symmetry in all directions, but long range orientational order. Since we wish to restrict ourselves to these two phases, it is convenient to study a model which does not exhibit any other liquid
crystal phases – in particular, no smectic phases. We chose a simple idealized model of soft ellipsoids, which is particularly suited for large scale molecular dynamics simulations.

Our work pursues two goals: First, we seek to improve our understanding of local and large scale interfacial properties. Second, we wish to explore ways which allow to bridge between the microscopic structure – which is to some extent accessible to computer simulations of particle based models – and phenomenological theories, which operate on larger scales. Our paper will highlight some aspects of these two issues.

The paper is organized as follows: The model and the simulation method are presented in the next section. Then we discuss the local liquid structure in isotropic and nematic fluids and show how that knowledge can be used to evaluate elastic constants in a nematic liquid crystal. The procedure is based on a density functional result, which is originally due to Poniewiersky and Stecki\textsuperscript{11}. It is an example of a procedure that bridges between local properties and phenomenological theories. In the third section, we describe results from large scale molecular dynamics simulations of a nematic-isotropic interface. Among other, we find an intriguing capillary wave spectrum, which reflects the complex interplay between the bare surface tension and the elastic interactions in the nematic phase. Finally, we summarize and discuss briefly the prospective future work.

2 Model, Method and Technical Details

Our model liquid crystal is defined as follows: We study soft ellipsoidal particles of mass $m_0$ with elongation $\kappa = \sigma_{\text{end-end}}/\sigma_{\text{side-side}}$. Two particles $i$ and $j$ with orientations $u_i$ and $u_j$ separated by the center-center vector $r_{ij}$ interact via a purely repulsive pair potential,

$$V_{ij} = \begin{cases} 
4\epsilon_0 \left(X_{ij}^{12} - X_{ij}^6\right) + \epsilon_0 : X_{ij}^6 > 1/2 \\
0 : \text{otherwise}
\end{cases}$$

where $X_{ij} = \sigma_0/(r_{ij} - \sigma_{ij} + \sigma_0)$ and

$$\sigma_{ij} = \sigma_0 \left\{ 1 - \frac{\chi}{2} \left[ \frac{(u_i \cdot \hat{r}_{ij} + u_j \cdot \hat{r}_{ij})^2}{1 + \chi u_i \cdot u_j} + \frac{(u_i \cdot \hat{r}_{ij} - u_j \cdot \hat{r}_{ij})^2}{1 - \chi u_i \cdot u_j} \right] \right\}^{-1/2}$$

with $\chi = (\kappa^2 - 1)/(\kappa^2 + 1)$. The function $\sigma_{ij}$ approximates the contact distance between the two ellipsoids in the direction $\hat{r}_{ij} = r_{ij}/r_{ij}^{12}$.

The systems were studied using a domain decomposition molecular dynamics program (GBMEGA). The simulations and parts of the data analysis (extensive evaluations of correlation functions) were carried out on the CRAYs T3E at NIC, using 128 and 256 processors in a typical run. The simulations were performed in the microcanonical ensemble using the RATTLE integrator\textsuperscript{13}, in rectangular simulation boxes with periodic boundary conditions in all directions. The time step $\Delta t$ and the moment of inertia $I$ depended on the elongation $\kappa$ of the ellipsoids in the systems under consideration: For systems with $\kappa = 3$, we chose $\Delta t = 0.003 \sqrt{m_0/\epsilon_0} \sigma_0$ and $I = 2.5 m_0 \sigma_0^2$, and for systems with $\kappa = 15$, $\Delta t = 0.002 \sqrt{m_0/\epsilon_0} \sigma_0$ and $I = 50 m_0 \sigma_0^2$. Further simulation details will be given in the appropriate section.
We describe the orientation of an ellipsoid \( i \) by a unit vector \( u_i \). The nematic phase is then characterized by a nonzero order tensor

\[
Q = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{2} (3u_i \otimes u_i - I),
\]

(3)

where the sum \( i \) runs over all \( n \) particles of the system, \( I \) is the unit matrix and \( \otimes \) the dyadic product of two vectors. The largest Eigenvalue of this matrix is the nematic order parameter \( S \). The corresponding Eigenvector points in the direction of alignment. It is called the director \( n \).

In the following, we shall use scaled units defined in terms of \( \sigma_0, \epsilon_0, m_0 \) and the Boltzmann constant \( k_B \).

3 Bulk Structure and Elastic Constants

Before studying interfacial and surface properties of fluids, it is important to understand the local bulk structure of these fluids. A good knowledge of the local equilibrium structure allows one to calculate effective parameters, which in turn can be used in phenomenological theories to predict material properties on a larger scale.

On a phenomenological level, nematic liquid crystals are often described by the Frank free energy functional\(^{14, 15}\)

\[
\mathcal{F}(n(r)) = \frac{1}{2} \int d\mathbf{r} \left\{ K_{11} [\nabla \cdot n]^2 + K_{22} [n \cdot (\nabla \times n)]^2 + K_{33} [n \times (\nabla \times n)]^2 \right\}.
\]

(4)

Here it is assumed that the absolute value of the order parameter \( S \) is roughly constant throughout the system. Nevertheless, long wavelength fluctuations of the director \( n(r) \) must exist, because the nematic order breaks the isotropy of space, which is a continuous symmetry, and the Goldstone theorem applies\(^{16}\). The three contributions to the free energy functional – the splay mode \((K_{11})\), the twist mode \((K_{22})\), and the bend mode \((K_{33})\) – are illustrated in Figure 1. They are controlled by the Frank elastic constants \( K_{ii} \) and determine almost exclusively the structure and the properties of nematic liquid crystals on mesoscopic length scales.

On a microscopic level, the structure of fluids is described by \( N \)-particle distribution functions\(^{17, 18}\) and in particular by pair correlations. The central quantity in many theories of liquid matter is the so-called direct correlation function (DCF). It is defined as follows:

\[ \]
The starting point are the two-particle probability density \( \rho^{(2)}(u_1, r_1, u_2, r_2) \) of finding simultaneously one particle of orientation \( u_1 \) at the position \( r_1 \), and another with orientation \( u_2 \) at the position \( r_2 \), and the one-particle probability density \( \rho^{(1)}(u) \) of finding a particle with the orientation \( u \) at the position \( r \). In the bulk, \( \rho^{(1)} \) is independent of \( r \), and \( \rho^{(2)} \) depends only on the relative coordinates \( r_{12} = r_1 - r_2 \). The correlations between two particles are described by the total correlation function

\[
h(u_1, u_2, r_{12}) = \frac{\rho^{(2)}(u_1, u_2, r_{12})}{\rho^{(1)}(u_1)\rho^{(1)}(u_2)} - 1. \tag{5}\]

This function subsumes the direct effect of a particle 1 on a particle 2 and all the “indirect” effects mediated by the bulk of surrounding fluid. Due to the elasticity of the nematic fluid, it decays only slowly with an algebraic power law \( 1/r \). The long range correlations result from indirect contributions. In order to characterize the local structure, it thus seems desirable to separate “indirect” effects from “direct” effects. This is done by the Ornstein-Zernike equation\(^{17,18}\)

\[
h(u_1, u_2, r_{12}) = c(u_1, u_2, r_{12}) + \int c(u_1, u_3, r_{13}) \rho^{(1)}(u_3) h(u_3, u_2, r_{32}) du_3 dr_3. \tag{6}\]

The DCF is the function \( c(u_1, u_2, r_{12}) \). As it turns out, \( c \) is indeed short ranged even in the nematic phase.

The DCF is the starting point for several liquid state theories\(^{17}\). Many material constants can be calculated from the DCF. In nematic fluids, one is specially interested in the parameters which control the director profiles. Poniewierski and Stecki\(^{11}\) have derived a set of expressions which relate the director profiles. These equations have been used a few times in the past to calculate elastic constants from simulations\(^{19-22}\). However, approximations were used for the form of the DCF, and the results therefore differed from values obtained by other methods for the same systems\(^{23}\).

We have determined for the first time the exact direct correlation functions in a nematic fluid\(^{24-27}\). To this end, we have studied systems of up to 8000 soft ellipsoids at the number density \( \rho = 0.3/\sigma_0^3 \) and temperature \( T = 0.5 \varepsilon_0/k_B \). This corresponds to a nematic state with the order parameter \( \langle S \rangle = 0.69 \). The phase transition to the isotropic phase occurs at \( \rho = 0.29/\sigma_0^3 \). For comparison, we have also calculated the DCF at the number density \( \rho = 0.24/\sigma_0^3 \), i.e., in an isotropic fluid. The run lengths were 5-10 million molecular dynamics steps.

All densities and correlation functions were expanded in spherical harmonics \( Y_{lm}(u) \), in a frame where the \( z \)-axis points along the director:

\[
\rho^{(1)}(u) = \varrho \sum_{l,m} f_l Y_{l0}(u) \tag{7} \\
F(u_1, u_2, r_{12}) = \sum_{l_1, m_1, l_2, m_2, l m} F_{l_1 m_1 l_2 m_2 l m}(r) Y_{l_1 m_1}(u_1) Y_{l_2 m_2}(u_2) Y_{l m}(\hat{r}_{12}). \tag{8}
\]
where $F = \rho^{(2)}$, $h$ or $c$, $r = |r_{12}|$, and $\hat{r}_{12} = r_{12}/r_{12}$. For symmetry reasons, all coefficients $f_l$ and $F_{ll'}(r)$ are real, and only coefficients with $m_1 + m_2 = -m$ and even $l$ enter the expansion.

The expansion coefficients of $\rho^{(2)}$ were computed from simulation configurations using

$$
\rho^{(2)}_{l_1 m_1, l_2 m_2, l m} (r) = 4\pi g^2 g(r) \left< Y_{l_1 m_1}^* (\mathbf{u}_1) Y_{l_2 m_2}^* (\mathbf{u}_2) Y_{l m}^* (\hat{r}) \right> \delta r,
$$

where $g(r)$ is the radial distribution function, i.e. the number of molecular pairs at distances between $r$ and $r + \delta r$, divided by $4\pi \sigma^2 r \delta r$. These time-consuming averages were calculated on a Cray T3E. We have determined coefficients for $l_i, l$ up to $l_{\text{max}} = 6$ in all systems, and up to $l_{\text{max}} = 8$ in the smallest nematic system. A calculation with $l_{\text{max}} = 8$ requires the evaluation of 1447 different expansion coefficients of $\rho^{(2)}$ from the configuration data. Our computational resources did not permit to perform such a time consuming analysis in the larger systems. Choosing $l_{\text{max}} = 6$, we still had to calculate 469 different coefficients, which we did mostly on the CRAYS at NIC. The bin size was $\delta r = 0.04 \sigma_0$.

Results for selected coefficients of $\rho^{(2)}$ and $c$ are presented in Figures 2 and 3. Figure 2 shows orientational averages (over $\mathbf{u}_i, \mathbf{u}_j$, and $r_{ij}$) of the pair distribution function and the DCF. Note that these curves are proportional to the coefficients $\rho^{(2)}_{000000}$ and $c_{000000}(r)$. The pair distribution does not look very different in the isotropic and the nematic phase. The DCF however reveals that the nematic phase has more hidden structure. Since we consider orientational averages, the long range algebraic correlations are not yet apparent: Both $\left< \rho^{(2)} \right>$ and $\left< c \right>$ are short ranged.

The effects of the elasticity become evident when looking at orientation dependent correlations. Figure 3 shows results for the expansion coefficients with $l_1 = l_2 = l = 2$, $m_1 = -m_2 = 1, m = 0$. In the case of $\rho^{(2)}$, this coefficient has a particularly pronounced long range tail. It disappears entirely in the corresponding curve for the DCF.

Figure 2. Pair distribution function $\rho^{(2)}$ (left) and direct correlation function $c$ (right) vs. molecular distance $r$, averaged over all orientations of $\mathbf{u}_i, \mathbf{u}_j$, and $r_{ij}$, for different system sizes $N$, at the density $\rho = 0.3/\sigma_0^3$ (nematic phase). Solid black line shows corresponding curve for $\rho = 0.24/\sigma_0^3$ (isotropic phase) for comparison.
Based on these results, we can establish the connection with the elastic constants in the nematic fluid. The Poniewierski-Stecki equations\(^{11}\) are most conveniently evaluated in Fourier space. They can be written as\(^{24}\)

\[
K_{ii} = \frac{1}{2} \frac{d^2}{dk^2} C_{ii}(k) \bigg|_{k=0} \quad \text{for} \quad i = 1, 2, 3, \quad (10)
\]

where the \(C_{ii}(k)\) are defined as

\[
C_{ii}(k) = \frac{k_B T \phi^2}{8 \sqrt{\pi}} \sum_{l_1,l_2} \sqrt{l_1(l_1+1)l_2(l_2+1)} f_{l_1} f_{l_2} \left\{ c_{l_1 l_2 - 100}(k) + c_{l_1 - l_2 100}(k) \right\} \\
+ \sqrt{5} \frac{v_i}{2} \left[ c_{l_1 l_2 - 120}(k) + c_{l_1 - l_2 120}(k) \right] + \sqrt{15} \frac{w_i}{8} \left[ c_{l_1 l_2 - 122}(k) + c_{l_1 - l_2 - 122}(k) \right]
\]

with \((v_1, v_2, v_3) = (-1, -1, 2), (w_1, w_2, w_3) = (-1, 1, 0)\). The elastic constants \(K_{ii}\) can thus be determined from the initial slope of a plot of \(C_{ii}(k)\) vs. \(k^2\). Such a plot is shown in Figure 4. For comparison and as a check of our method, we have also performed an analysis of order tensor fluctuations following a method proposed by Allen et al\(^{29, 23}\). To this end, we have performed additional molecular dynamics simulations of 4000 and 16000 particles in an ensemble where the director was constrained to be aligned along one side of the simulation box.

The results are summarized in Table 1. We have evaluated the pair distribution functions and calculated the DCF in systems of 1000, 4000, and 8000 particles. The upper cutoff \(l_{\text{max}}\) for the spherical harmonics expansion was chosen \(l_{\text{max}} = 8\) in the smallest systems, and \(l_{\text{max}} = 6\) in the others. Results for \(l_{\text{max}} = 6\) and \(l_{\text{max}} = 8\) were compared in the smallest system. They differed most critically in the value of \(K_{33}\), but the difference is small (a few percent at \(N = 1000\)). As explained above, we were not able to carry out an analysis with \(l_{\text{max}} = 8\) for all system sizes.
Comparing the results for the elastic constants with those obtained from the analysis of order tensor fluctuations, we find that the values of $K_{11}$ and $K_{22}$ are identical for both methods. The DCF method slightly underestimates $K_{33}$, but since the value increases with $l_{\text{max}}$, this is probably an artefact of the cutoff in the spherical harmonics expansion.

Despite these systematic errors, the agreement between the values obtained from the two methods is reasonable. Hence we have established a practical way to apply the Poniewierski-Stecki equations to simulation data. It allows to calculate the elastic constants, which determine the mesoscopic structure of a nematic fluid, from the DCF, which characterizes its local liquid structure. Moreover, our findings give confidence in our results for the DCF. As a central quantity in density functional theory, the possible applications of the DCF go of course far beyond the calculations of elastic constants\textsuperscript{17,30}. We plan to explore some of these in the future.

<table>
<thead>
<tr>
<th>$N$</th>
<th>method</th>
<th>$l_{\text{max}}$</th>
<th>$\langle K_{11}\rangle$</th>
<th>$\langle K_{22}\rangle$</th>
<th>$\langle K_{33}\rangle$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1000</td>
<td>PS</td>
<td>8</td>
<td>0.55 (2)</td>
<td>0.35 (3)</td>
<td>1.56 (4)</td>
</tr>
<tr>
<td>1000</td>
<td>PS</td>
<td>6</td>
<td>0.51 (2)</td>
<td>0.34 (3)</td>
<td>1.52 (4)</td>
</tr>
<tr>
<td>4000</td>
<td>PS</td>
<td>6</td>
<td>0.52 (2)</td>
<td>0.31 (1)</td>
<td>1.51 (3)</td>
</tr>
<tr>
<td></td>
<td>OF</td>
<td>6</td>
<td>0.53 (1)</td>
<td>0.30 (1)</td>
<td>1.60 (1)</td>
</tr>
<tr>
<td>8000</td>
<td>PS</td>
<td>6</td>
<td>0.51 (2)</td>
<td>0.33 (2)</td>
<td>1.48 (3)</td>
</tr>
<tr>
<td>16000</td>
<td>OF</td>
<td>6</td>
<td>0.53 (1)</td>
<td>0.30 (1)</td>
<td>1.59 (1)</td>
</tr>
</tbody>
</table>

Table 1. Elastic constants as obtained in systems with $N$ particles with two methods. PS: Poniewierski-Stecki equations with DCF, calculated with a spherical harmonics expansion with an upper cutoff $l_{\text{max}}$; the results for $l_{\text{max}} < 6$ were about 20\% worse. OF: analysis of order tensor fluctuations. The statistical error on the last digit is given within parentheses.
4 The Nematic/Isotropic Interface

We turn to discuss interfacial properties in nematic liquid crystals. From a macroscopic point of view, the presence of surfaces or interfaces introduces three new effective parameters: the interfacial tension, the anchoring angle, and the anchoring energy. The anchoring angle is the director angle (with respect to the surface normal) favored by the surface. The anchoring energy is related to the force needed to twist the director out of the anchoring angle.

Our work at NIC currently concentrates on the study of interfaces between nematic and isotropic states (NI-interface). These interfaces are of fundamental interest, because they connect two fluid phases with a different symmetry. They are essential for wetting phenomena, which can be exploited to manipulate the anchoring at surfaces. They also play a key role for important nonequilibrium phenomena such as shear banding\(^{31,32}\). In all of these cases, the physics depends crucially on the structure and in particular on the fluctuations of the associated interfaces.

We have started with studying a free NI-interface. There have been a few simulations of NI-interfaces previously\(^{34-36}\). However, the systems were small, and fluctuation effects could not be studied. We considered a system with 115200 soft ellipsoids of elongation \(\kappa = 15\). The high value of \(\kappa\) ensured a broad coexistence region: At the temperature \(T = 1\epsilon/k_B\), the coexistence densities of the nematic and the isotropic phase were \(\rho_N = 0.018/\sigma_0^3\) and \(\rho_I = 0.016/\sigma_0^3\), respectively. The simulations were done in the microcanonical ensemble in a box geometry \((L_x : L_y : L_z) = (1 : 1 : 2)\) with periodic boundary conditions. The average density \(\rho = 0.17/\sigma_0^3\) (i.e. \(L_x = 150.2\sigma_0\)) was chosen in the coexistence region. Thus the system phase separated into a nematic slab and an isotropic slab, with two interfaces in between. After a long equilibration procedure (see Reference [33] for details), the systems were sampled over 3 million molecular dynamics steps.
A snapshot of part of a configuration is shown in Figure 5. One recognizes the nematic side (bottom), the isotropic side (top), and the interface between the two. The anchoring on the nematic side is planar, i.e., the particles are aligned parallel to the interface on average. One also recognizes strong fluctuations of the position of the interface (capillary waves). These clearly have to be taken into account when studying the interfacial structure.

The simplest theory of interfacial fluctuations, the capillary wave theory\textsuperscript{37}, presumes that the local position of the interface can be parametrized by a unique function $h(x,y)$, and that the fluctuations of $h$ are controlled solely by the interfacial tension $\gamma$. Under that assumption, the (two-dimensional) Fourier modes of $h$ are distributed like

$$\langle |h(q)|^2 \rangle = k_B T / (\gamma q^2).$$  \hfill (11)

The capillary waves broaden the interfacial region. The apparent interfacial width $\omega$ is predicted to depend on the lateral size $L_\parallel$ like

$$\omega^2 = \omega_0^2 + \frac{k_B T}{4 \gamma} \ln(L_\parallel/b_0),$$  \hfill (12)

where $b_0$ is a microscopic reference length, and $\omega_0$ is the interfacial width on that length scale.

In order to study these effects, we have divided the simulation box in columns of size $B \times B \times L_z$ and calculated the two local interface positions $h$ in each block. Profile averages were then computed with respect to the distance to the closest interface. Figure 6 (left) shows the resulting order parameter profiles for different block sizes $B$. The apparent width of the interface increases with the block size $B$. We have determined the interfacial width $\omega$ by fitting the profiles to a tanh-function. In a plot of $\omega^2$ as a function of the block size $B$ (Figure 6, right), we recover the logarithmic growth predicted by the capillary wave

![Figure 6. Left: Order parameter profiles for different block sizes as indicated. $S_N$ and $S_I$ denotes average values of the order parameter in the nematic and the isotropic phase, respectively. Right: Squared interfacial width $\omega^2$ vs. block size $B$.](image-url)
Figure 7. Inverse of mean-squared Fourier components of the interface position, $1/\langle |h(q)|^2 \rangle$, vs. square of the wave vector $q^2$, for wavevectors pointing parallel (circles) and perpendicular (squares) to the director. Dashed lines are guides to the eye. Solid lines are predictions of the capillary wave theory (11) with $\gamma = 0.016 \sigma_0^2/k_B T$ (upper, blue line) and $\gamma = 0.0093 \sigma_0^2/k_B T$ (lower, red line).

The capillary wave spectrum $h(q)$ was analyzed from the landscape of interfacial positions $h$ at a fixed block size $B = L_z/8$. Figure 7 plots the results for $1/\langle |h(q)|^2 \rangle$ versus $q^2$. Comparing these data with the equation (11), one notices that the prediction of the simple capillary wave theory does not describe the interface quite as well as the results for the interfacial broadening made us believe. First, the fluctuations of the interfacial position are anisotropic: They are much larger in the direction perpendicular to the director than in the parallel direction. Second, the amplitude of the fluctuations is larger than expected on small length scales (large $q$), and smaller than expected on large length scales ($q \to 0$). The discrepancy with the capillary wave prediction on large length scales becomes even worse if we use an independent estimate of the interfacial tension, $\gamma = 0.0093 \sigma_0^2/k_B T$, which was calculated from the anisotropy of the pressure tensor.$^{36}$

Hence the simulations reveal an unexpectedly complex capillary wave spectrum. In fluids with purely short-range interactions, the capillary wave theory (11) usually describes fluid-fluid interfaces quite well on large length scales. The discrepancies observed here suggest that the effectively long-range elastic interactions in the nematic phase influence the interfacial fluctuations significantly. One can speculate that they might even suppress them entirely in the limit of infinitely long particles$^{33}$ (Onsager limit$^{38,39}$). These relations between elastic properties and fluctuations shall be explored in more detail in the future. Obviously, they are not only important for the nematic-isotropic interface, but more generally for all surfaces of nematic fluids.
5 Conclusions and Outlook

To summarize, we study the structure and properties of liquid crystals by molecular dynamics simulations. As for all materials, a global understanding of the properties of liquid crystals requires simulations on many different length scales and levels of coarse graining. Our simulations operate on a coarse grained microscopic level: The length scale is microscopic, yet details of the molecular structure are disregarded. The purpose of our studies is twofold: On the one hand, we aim at a better understanding of the relationship between the microscopic, local structure, and the mesoscopic properties of inhomogeneous liquid crystals: For example, we have demonstrated that the DCF can be used to bridge between the microscopic length scale and higher levels of coarse graining, in which only effective parameters like the elastic constants matter. On the other hand, we explore physical phenomena which are characteristic for our length scale – such as the specific properties of interfacial fluctuations on the length scale of a few molecular lengths.

Our future work shall follow these lines. We plan to study the relationship between the local structure at surfaces and the phenomenological parameters which describe surface anchoring. Here again, density functional approaches which make use of DCF information seem promising. Furthermore, we believe that the unusual fluctuation spectrum which we have observed in equilibrium interfaces will lead to a wealth of new phenomena in nonequilibrium interfaces, i.e., interfaces under shear.

Acknowledgements

We have benefitted from discussions and interactions with K. Binder, A. McDonald, and H. Lange. The work presented here was carried out using extensive computer time on the CRAY's T3E of the NIC (Jülich). Altogether, we have used roughly 200,000 processor hours (400,000 billing units). The parallel program GBMEGA was originally developed by the EPSRC Complex Fluids Consortium, UK. Our work was supported by the German Science Foundation (DFG).

References

Metal-Insulator Transitions and Realistic Modelling of Correlated Electron Systems

N. Blümer¹, K. Held², G. Keller, and D. Vollhardt

Theoretische Physik III, Elektronische Korrelationen und Magnetismus, Institut für Physik
Universität Augsburg, 86135 Augsburg, Germany
¹Present address: Institut für Physik, Universität Mainz, 65099 Mainz, Germany
²Present address: Physics Department, Princeton University, Princeton, NJ 08544, USA

The correlation-induced metal-insulator transition found experimentally in Cr-doped V₂O₃ is investigated within the one-band Hubbard-type model and a realistic multi-band model, using dynamical mean-field theory (DMFT). The DMFT equations are solved numerically with an auxiliary-field quantum Monte-Carlo algorithm. For the one-band model a detailed phase diagram is obtained which resolves a long-standing controversy. The calculated spectra, orbital occupations and the spin state obtained for the realistic model are found to agree with recent polarization dependent X-ray-absorption experiments.

1 Introduction

The calculation of physical properties of electronic systems by controlled approximations is one of the most important challenges of modern theoretical solid state physics. In particular, the physics of transition metal oxides – a singularly important group of materials both from the point of view of fundamental research and technological applications – may only be understood by explicit consideration of the strong effective interaction between the conduction electrons in these systems. The investigation of electronic many-particle systems is made especially complicated by quantum statistics, and by the fact that the phenomena of interest (e.g., metal insulator transitions and ferromagnetism) usually require the application of non-perturbative theoretical techniques.

One of the most famous examples of a cooperative electronic phenomenon of this type is the transition between a paramagnetic metal and a paramagnetic insulator induced by the Coulomb interaction between the electrons, referred to as Mott-Hubbard metal-insulator transition. The question concerning the nature of this transition poses one of the fundamental theoretical problems in condensed matter physics.¹,² Correlation-induced metal-insulator transitions (MIT) are found, for example, in transition metal oxides with partially filled bands near the Fermi level. For such systems band theory typically predicts metallic behavior. The most famous example is V₂O₃ doped with Cr;³⁻⁵ see Fig. 1. While at low temperatures V₂O₃ is an antiferromagnetic insulator with monoclinic crystal symmetry, it has a corundum structure with a small trigonal distortion in the high-temperature paramagnetic phase. All transitions shown in the phase diagram (Fig. 1) are of first order. In the case of the transitions from the high-temperature paramagnetic phases into the low-temperature antiferromagnetic phase this is naturally explained by the fact that the transition is accompanied by a change in crystal symmetry. By contrast, the crystal symmetry across the MIT in the paramagnetic phase remains intact, since only the ratio of the c/a axes changes discontinuously. This was usually taken as an indication for the predominantly electronic origin of this MIT.
2 One-Band Hubbard Model

To explain a MIT induced by electronic correlations one may choose to start with the investigation of an electronic many-body model to understand, at least, some of the basic features of the MIT, or employ more material-specific approaches. Concerning the former approach,\textsuperscript{5,1} the spin $S = 1/2$, half-filled, single-band Hubbard model,\textsuperscript{6–8}

$$\hat{H}_{\text{one-band}} = -t \sum_{\langle i,j \rangle, \sigma} \left( \hat{c}_{i\sigma}^\dagger \hat{c}_{j\sigma} + \hat{c}_{j\sigma}^\dagger \hat{c}_{i\sigma} \right) + U \sum_i \hat{n}_{i\uparrow} \hat{n}_{i\downarrow},$$

is certainly the simplest possible model to be investigated. Here, the operators $\hat{c}_{i\sigma}^\dagger$ and $\hat{c}_{i\sigma}$ create and annihilate electrons of spin $\sigma$ on lattice site $i$, respectively, while $\hat{n}_{i\sigma} = \hat{c}_{i\sigma}^\dagger \hat{c}_{i\sigma}$ counts the electrons of spin $\sigma$ on site $i$. It should be noted that due to the Pauli principle a lattice site can be occupied by at most one electron per spin species $\sigma \in \{\uparrow, \downarrow\}$. While the on-site interaction energy parametrized by $U$ is diagonal in the real-space representation (1), the kinetic energy, i.e., the hopping processes parametrized by $t$, is diagonal in momentum space. Thus the two terms in the Hamiltonian are maximally incompatible concerning their commutation.

The existence of a MIT in the paramagnetic phase\textsuperscript{9} of the half-filled Hubbard model had been investigated already in the early work of Hubbard.\textsuperscript{10} The details of the transition remained unclear for a long time. During the last few years, however, our understanding of the MIT in the one-band Hubbard model has considerably improved due to the application of the dynamical mean-field theory (DMFT). Within the DMFT, the electronic lattice problem is mapped onto a single-impurity Anderson model (SIAM) with a self-consistency condition. The mapping becomes exact in the limit of infinite coordination number\textsuperscript{11,12} and allows one to investigate the dynamics of correlated lattice electrons non-perturbatively at all interaction strengths. This is of essential importance for a problem like the MIT which
occurs at a Coulomb interaction comparable to the electronic band-width. The Green function for the SIAM can be written as a functional integral over Grassmann variables $\psi, \psi^*$,

$$G_{\sigma n} = -\langle \psi_{\sigma n}^\dagger \psi_{\sigma n} \rangle_{\mathcal{A}} = -\frac{1}{Z} \int \mathcal{D}[\psi^*] \mathcal{D}[\psi] \psi_{\sigma n}^\dagger \psi_{\sigma n} e^{\mathcal{A}[\psi, \psi^*, \mathcal{G}]}.$$  \hspace{1cm} (2)

Here, the index $n$ denotes the fermionic Matsubara frequency $\omega_n = (2n + 1)\pi T$ for temperature $T$ and $Z$ is the partition function

$$Z = \int \mathcal{D}[\psi^*] \mathcal{D}[\psi] e^{\mathcal{A}[\psi, \psi^*, \mathcal{G}]}.$$  \hspace{1cm} (3)

The action $\mathcal{A}$ can be expressed as $\mathcal{A}[\psi, \psi^*, \mathcal{G}] = \lim_{\Lambda \to \infty} \mathcal{A}_\Lambda[\psi, \psi^*, \mathcal{G}]$ in terms of a discretized imaginary time $\tau_l = l\Delta \tau$ with a finite number $\Lambda = \beta/\Delta \tau$ of time slices, where $\beta = 1/T$ is the inverse temperature and

$$\mathcal{A}_\Lambda[\psi, \psi^*, \mathcal{G}] = (\Delta \tau)^2 \sum_{\sigma} \sum_{l, l'} \psi_{\sigma l}^*(G_{\sigma}^{-1})_{ll'} \psi_{\sigma l'} - \Delta \tau U \sum_{l=0}^{\Lambda-1} \psi_{\uparrow l}^* \psi_{\uparrow l} \psi_{\downarrow l}^* \psi_{\downarrow l}.$$  \hspace{1cm} (4)

The self consistency requires that the impurity Green function, which is related to the (local) self energy $\Sigma$ by the impurity Dyson equation

$$G_{\sigma n} = \int_{-\infty}^{\infty} d\varepsilon \frac{\rho_0(\varepsilon)}{i\omega_n + \mu - \Sigma_{\sigma n} - \varepsilon}.$$  \hspace{1cm} (5)

This equation is essential since it introduces the lattice nature of the problem via the non-interacting density of states $\rho_0(\varepsilon)$. A numerical solution of the DMFT problem is achieved in an iterative process as illustrated in Fig. 2. Starting with some initial self energy $\Sigma_0$, (5) is used to calculate the bath Green function $\mathcal{G}$ which defines the impurity problem. Its solution is a difficult problem which is in this work achieved using the auxiliary-field quantum Monte Carlo (QMC) algorithm by Hirsch and Fye.\textsuperscript{13} After a Trotter decomposition of $e^{\mathcal{A}}$, a discrete Hubbard-Stratonovich transformation is employed,

$$\exp \left( \frac{\Delta \tau U}{2} (\psi_{\uparrow l}^* \psi_{\uparrow l} - \psi_{\downarrow l}^* \psi_{\downarrow l})^2 \right) = \frac{1}{2} \sum_{s = \pm 1} \exp \left( \lambda s (\psi_{\uparrow l}^* \psi_{\uparrow l} - \psi_{\downarrow l}^* \psi_{\downarrow l}) \right) \hspace{1cm} (6)$$
with $\cosh \lambda = \exp(\Delta \tau U/2)$. Thus, the electron-electron interaction term which is quartic in the Grassmann variables is replaced by the interaction with an external auxiliary Ising field $\{s_l\}$ with $s_l = \pm 1$. For a given Ising configuration, the functional integral becomes

$$G_{\sigma ll'} = \frac{1}{Z} \sum_{\{s_i\}} (M^{-1}_{\sigma} M_{\sigma})^{-1}_{ll'} \det M_{\downarrow}^{\{s_i\}} \det M_{\uparrow}^{\{s_i\}}, \quad (7)$$

where $M_{\downarrow}^{\{s_i\}}$ is the matrix with elements $M_{\downarrow}^{s_l} = (\Delta \tau)^2 (G^{-1}_{\sigma} \sigma)_{ll'} - \lambda \sigma \delta_{ll'} s_l$, and the partition function has the value

$$Z = \sum_{\{s_i\}} \det M_{\downarrow}^{\{s_i\}} \det M_{\uparrow}^{\{s_i\}}. \quad (8)$$

Computing one of the $2^\Lambda$ terms in (7) directly is an operation of order $O(\Lambda^3)$. However, if the terms are ordered in a way so that successive configurations $\{s_i\}$ and $\{s_i\}'$ only differ by one flipped spin $s_l \rightarrow -s_l$ then all matrices and determinants can be updated at a cost of $O(\Lambda^2).$ Only for $\Lambda \lesssim 24$ can all terms be summed up exactly. Computations at larger $\Lambda$ are made possible by Monte Carlo importance sampling which drastically reduces the number of terms that have to be calculated explicitly from $2^\Lambda$ to order $O(\Lambda)$. Here, the absolute values of the products of determinants in (7) and (8) are used as unnormalized probabilities in a Metropolis single-spin update Markov chain. Numerical errors arise at several stages of the algorithm: The finite Monte Carlo sampling length introduces a statistical error (and possibly a systematic error from incomplete warm-up); systematic errors incur from the finite number of iteration cycles and, most importantly, the finite discretization $\Delta \tau$ of the imaginary time. For constant accuracy, i.e., for a constant value of $\Delta \tau$ and constant numbers of iterations and sweeps, the total numerical effort scales with $T^{-3}$, i.e., rapidly increases with decreasing temperature. In the one-band calculations discussed in this section, very high accuracy could be obtained even at low temperatures by using up to $10^6$ sweeps per iteration, about 10-100 iterations, and up to $\Lambda_{\text{max}} = 400$ time slices.

The DMFT provides a framework for deriving a coherent picture of the electronic spectrum at all energy scales, i.e., of the incoherent features at high energies (Hubbard bands), and the coherent quasi-particle behavior at low energies. At $T = 0$ the transition from the metallic to the insulating state is then signaled by the collapse of the Fermi liquid quasi-particle peak at the Fermi energy when the interaction is increased. Qualitatively, this behavior is still seen at finite, not too high temperatures as shown in Fig. 3 for $T = 0.05$. When the interaction $U$ becomes of the order of the noninteracting bandwidth $W = 4$, weight is shifted from the quasi-particle peak at $\omega \approx 0$ to the upper and lower Hubbard bands. Within the Fermi liquid regime ($U \lesssim 4.4$), the DOS at the Fermi energy ($\omega = 0$) is essentially pinned to its noninteracting value. The deviation for $U = 4.7$ signals the breakdown of the Fermi liquid before the MIT occurs at $U \approx 4.8$. The insulating gap becomes more pronounced and widens when the interaction is further increased.

Spectra as shown in Fig. 3 require analytic continuation to the real axis via a maximum-entropy method which is quantitatively less reliable than the QMC procedure itself. Therefore, the phase diagram can be much more accurately determined from the analysis of static properties. In Fig. 4a, the total energy $E$ per lattice site is shown for a relatively high temperature $T = 0.067$ as a function of $U$. For each value of the discretization $\Delta \tau$, a
Figure 3. Spectra of the half-filled one-band Hubbard model for discretization $\Delta \tau = 0.1$.

Figure 4. Energy $E$ per lattice site as a function of on-site interaction $U$ for (a) $T = 0.067$, and (b) $T = 0.031$. Both figures show results for finite discretization $\Delta \tau$ and extrapolations to the physical limit $\Delta \tau = 0$.

kink clearly signals the metal-insulator transition. Due to the regular shape, it is also possible to extrapolate to the physical limit $\Delta \tau \rightarrow 0$. At the lower temperature $T = 0.031$, one finds hysteresis between metallic and insulating solutions which coexist for $4.8 \leq U \leq 5.0$ (depending on $\Delta \tau$) as seen in Fig. 4b. This coexistence points to a first-order transition at low temperatures in qualitative agreement with the respective transition observed in $V_2O_3$.

The crosses in the phase diagram (Fig. 5) summarize the QMC measurements of crossover lines, the transition point, and the coexistence region above, at, and below the critical temperature $T^* \approx 0.067$, respectively. These results show excellent agreement with recent finite-temperature calculations\textsuperscript{18} using the numerical renormalization group (NRG) to solve the DMFT equations, but deviate considerably from early estimates\textsuperscript{12} obtained within iterated perturbation theory (IPT). Obviously, the first-order phase transition line has to be pinpointed within the coexistence region for obtaining the full phase diagram. While a direct comparison of free energies is numerically unstable, we succeeded in deter-
mining the first-order transition line from a Clausius-Clapeyron differential equation using very precise measurements of energy and double occupancy plus low-temperature Fermi liquid properties. The corresponding thick curve in Fig. 5 is drawn as a solid line where it is completely determined from QMC, and as a dashed line at low temperatures where it is dominated by additional (reliable) input from ground state methods such as NRG and the projective self-consistent theory (PSCT). The precise determination of the phase diagram Fig. 5 does not only resolve a long-standing controversy about the physics of the half-filled frustrated one-band Hubbard model within the DMFT, but may also serve as a benchmark for alternative numerical methods of solving the DMFT equations.

3 Realistic Model for V$_2$O$_3$

Although the Hubbard model is able to explain certain basic features of the Mott-Hubbard MIT and the phase diagram of V$_2$O$_3$, it cannot explain the physics of that material in any detail. Clearly, a realistic theory of V$_2$O$_3$ must take into account the complicated electronic structure of this system. In the high-temperature paramagnetic phase V$_2$O$_3$ has an electronic structure with a $3d^2$ $V^{3+}$ state, where the two $e_g$-orbitals are empty and the three $t_{2g}$-orbitals are filled with two electrons. A small trigonal distortion lifts the triple degeneracy of the $t_{2g}$-orbitals, resulting in one non-degenerate $a_{1g}$-orbital and two degenerate $e_g^\pi$ orbitals. Starting from this orbital structure Castellani et al. proposed a widely accepted model with a strong covalent $a_{1g}$-bond between two V ions along the c-axis. This bonding state is occupied by a singlet pair (one electron per V) and hence does not contribute to the local magnetic moment. The remaining electron per V has a twofold orbital degeneracy within the $e_g^\pi$ orbitals and a spin $S = 1/2$. This $S = 1/2$ model, derived
within a material-specific approach, strongly supported the hope that the half filled, one-band Hubbard model was not only the simplest possible, but, in fact, a valid starting point for the investigation of V$_2$O$_3$.

However, recent experimental results by Park et al.\textsuperscript{20} seem to require an interpretation in terms of a $S = 1$ spin state, and an $e^g_\pi e^g_\sigma$ orbital state with an admixture of $e^g_\delta a^1_\delta$ configurations. Subsequently, Ezhov et al.\textsuperscript{21} and Mila et al.\textsuperscript{22} argued for $S = 1$ models without and with orbital degeneracy, respectively, for the antiferromagnetic insulating phase of V$_2$O$_3$. LDA+U calculations indicate that the atomic Hund’s rule is responsible for the high-spin ground state of the V ions.\textsuperscript{21} While LDA+U may be used to describe the antiferromagnetic insulating phase of V$_2$O$_3$, the metal-insulator transition within the correlated paramagnetic phase is beyond the scope of this approach since the Coulomb interaction is treated within Hartree-Fock. Here the computational scheme LDA+DMFT,\textsuperscript{23–26} obtained by combining electronic band structure theory (LDA) with the many-body technique DMFT, is the best available method for the investigation of real systems close to a Mott-Hubbard MIT.\textsuperscript{27} To solve the DMFT-equations we\textsuperscript{28} employ a multiband-version of the quantum Monte Carlo method\textsuperscript{29, 30} which yields a numerically exact solution;\textsuperscript{31} the resulting calculational scheme is referred to as LDA+DMFT(QMC).\textsuperscript{24, 26, 27} Note that in multiband QMC calculations (with $M = 3$ bands), the numerical cost increases by an additional factor of $2M^2 - M = 15$ compared to the single-band case, so that for a comparable computer allocation the lowest attainable temperature increases by about a factor of 2.5.

In a first step, LDA calculations were performed for paramagnetic metallic V$_2$O$_3$ and paramagnetic insulating (V$_{0.962}$Cr$_{0.038}$)$_2$O$_3$, respectively.\textsuperscript{28} The LDA results for corundum V$_2$O$_3$ and (V$_{0.962}$Cr$_{0.038}$)$_2$O$_3$ are very similar. In particular, the changes in crystal and electronic structure occurring at the transition are insufficiently reflected by the LDA calculations and the experimentally observed insulating gap is missing in the LDA DOS. It is generally believed that this insulating gap is due to strong Coulomb interactions which are not adequately accounted for by the LDA. This is where our LDA+DMFT(QMC) scheme sets in. Using this approach we can show explicitly that the insulating gap is indeed caused by electronic correlations.

In the LDA+DMFT approach\textsuperscript{23–27} the LDA band structure, expressed by a one-particle Hamiltonian $H^0_{\text{LDA}}$, is supplemented with the local Coulomb repulsion $U$ and Hund’s rule exchange $J$:

$$
\hat{H} = \hat{H}^0_{\text{LDA}} + U \sum_m \sum_i \hat{n}_{im\uparrow} \hat{n}_{im\downarrow} + \sum_{i,m \neq m',\sigma,\delta} (V - \delta_{\sigma\delta} J) \hat{n}_{im\sigma} \hat{n}_{i'm'\delta}.
$$

(9)

Here, $i$ denotes the lattice site; $m$ and $\tilde{m}$ enumerate the three interacting $t_{2g}$ orbitals. The interaction parameters are related by $V = U - 2J$ which holds exactly for degenerate orbitals and is a good approximation for the $t_{2g}$ orbitals. Furthermore, since the $t_{2g}$ bands at the Fermi energy are rather well separated from all other bands we restrict the calculation to these bands (for details of the computational scheme see Refs. 26, 27). With this restriction only the LDA DOS of the three $t_{2g}$ bands enter the LDA+DMFT calculation.\textsuperscript{27} The LDA-calculated value of the Coulomb repulsion $U$ has a typical uncertainty of at least 0.5 eV.\textsuperscript{26} For this reason, we adjust $U$ to yield a metal-insulator transition with Cr doping. A posteriori, we can compare whether the adjusted value is in the range of values obtained from a constrained LDA calculation.
The spectra obtained by LDA+DMFT(QMC) imply that the critical value of $U$ for the MIT is about 5 eV. Indeed, at $U = 4.5$ eV one observes pronounced quasiparticle peaks at the Fermi energy, i.e., characteristic metallic behavior, even for the crystal structure of \((V_{0.962}Cr_{0.038})_2O_3\), while at $U = 5.5$ eV the form of the calculated spectral function is typical for an insulator for both sets of crystal structure parameters. (We note that at $T \approx 0.1$ eV one only observes metallic-like and insulating-like behavior, with a rapid but smooth crossover between these two phases, since a sharp MIT occurs only at lower temperatures.)

To compare with the photoemission spectrum of $V_2O_3$ by Schramme et al., the LDA+DMFT(QMC) spectra are multiplied with the Fermi function at $T = 0.1$ eV and Gauss-broadened by 0.05 eV to account for the experimental resolution. The theoretical results for $U = 5$ eV are seen to be in good agreement with experiment (Fig. 6), in contrast with the LDA results. We also note that the DOS is highly asymmetric with respect to the Fermi energy due to the orbital degrees of freedom. This is in striking contrast to the result obtained with a one-band model. The comparison between our results, the data of Müller et al. obtained by X-ray absorption measurements, and LDA in Fig. 7 shows that, in contrast with LDA, our results not only describe the different bandwidths above and below the Fermi energy ($\approx 6$ eV and $\approx 2 - 3$ eV, respectively) correctly, but even resolve the two-peak structure above the Fermi energy.

Particularly interesting are the spin and the orbital degrees of freedom in $V_2O_3$. We find (not shown) that for $U \gtrsim 3$ eV the squared local magnetic moment $\langle m_z^2 \rangle$ saturates at a value of 4, i.e., there are two electrons with the same spin direction in the \((a_{1g}, e_{\pi1}, e_{\pi2})\) orbitals. Thus, we conclude that the spin state of $V_2O_3$ is $S = 1$ throughout the Mott-Hubbard transition region. Our $S = 1$ result agrees with the measurements of Park et al. and also with the data for the high-temperature susceptibility. Thus LDA+DMFT(QMC) provides a remarkably accurate microscopic theory of the strongly correlated electrons in $V_2O_3$. 

Figure 6. a) LDA+DMFT(QMC) spectrum for $U = 4.5, 5$ and $5.5$ eV at $T = 0.1$ eV $\approx 1000$ K; b) Comparison with the LDA spectrum and the photoemission experiment by Schramme et al. for two different $V_2O_3$ single-crystal surfaces at $T = 300$ K. Note, that the (10\(\bar{1}2\)) surface has the same coordination number as the bulk.
Figure 7. Comparison of the LDA and LDA+DMFT(QMC) spectra at $T = 0.1$ eV (Gaussian broadened with 0.2 eV) with the X-ray absorption data of Müller et al.\textsuperscript{36}.

the paramagnetic phase of V$_2$O$_3$.\textsuperscript{28}

The MIT will eventually become first order at lower temperatures;\textsuperscript{38,16} QMC simulations at $T \approx 300$ K are under way, but are very computer-expensive. Furthermore, future investigations will have to clarify the origin of the discontinuous lattice distortion at the first-order MIT which leaves the lattice symmetry unchanged. In particular, the MIT might be the driving force behind the lattice distortion by causing a thermodynamic instability with respect to changes of the lattice volume and distortions.

4 Conclusions

The Mott-Hubbard metal insulator transition is a phenomenon occurring in solids where the effective strength of the Coulomb interaction between the conduction electrons is comparable to the electronic band-width. Thus it belongs to the class of notoriously difficult intermediate coupling problems whose investigation requires non-perturbative techniques. We presented detailed results for the transition scenario occurring in the one-band Hubbard model and in a realistic multi-band model, obtained by the numerical solution of the equations of the dynamical mean-field theory (DMFT) using an auxiliary-field quantum Monte-Carlo algorithm (QMC). Our QMC results provide accurate insight into the details of the Mott-Hubbard metal insulator transition in the one-band model and thereby resolve a long-standing controversy about the nature of this transition. Furthermore, applying the newly developed LDA+DMFT(QMC) scheme, which merges the conventional local density approximation (LDA) with DMFT in combination with QMC, to investigate the paramagnetic phase of V$_2$O$_3$ we find remarkable agreement with recent experiments. Indeed, LDA+DMFT(QMC) turns out to be a workable computational scheme which provides, at last, a powerful tool for future ab initio investigations of real materials with strong electronic correlations.

The novel computational scheme for the investigation of strongly interacting matter described in this report can only be implemented on vector supercomputers like the CRAY-T90 and massively parallel machines with large caches like the CRAY-T3E of the NIC. In fact, using a total of about 2500 CPU hours on the T90 and 35000 CPU hours on the T3E, the computer simulations reported here already tested the limits of these supercomputers. Eventually a genuine fusion of LDA with DMFT will make it necessary to go to
considerable lower temperatures, i.e., room temperature and below. Furthermore, truly realistic investigations of transition metals and their oxides will require the implementation of a scheme where all relevant electronic bands are included. As discussed above a larger number of bands and lower temperatures imply an enormous increase of numerical effort. This shows that substantial progress in the modelling of correlated electronic materials will only be possible with much faster computers. In view of the eminent importance of these materials for fundamental research and technological applications the construction and availability of faster supercomputers is thus highly desirable.

Acknowledgments

We thank A. Sandvik for making his maximum entropy code available to us. This work was supported by the John v. Neumann-Institut für Computing, Jülich, and by the Deutsche Forschungsgemeinschaft through SFB 484.

References

9. At low temperatures the paramagnetic phase can be stabilized by appropriately frustrating the incipient antiferromagnetic order in the Hubbard model at half filling.
31. For a comparison of LDA+DMFT results for the photoemission spectra of $La_{1-x}Sr_xTiO_3$ obtained by different numerical techniques see Ref. 26.
32. Use of the crystal structure of Cr-doped $V_2O_3$ for the insulating phase of pure $V_2O_3$ is justified by the observation that Cr-doping is equivalent to the application of (negative) pressure.
33. All QMC results presented were obtained for $T = 0.1$ eV. Simulations for $V_2O_3$ at $U = 5$ eV, $T = 0.143$ eV, and $T = 0.067$ eV suggest, however, only a minor smoothing of the spectrum with increasing temperature.
35. M. Schramme, Ph.D. thesis, Universität Augsburg, 2000; M. Schramme et al. (unpublished). Recent experiments by H.-D. Kim et al. (cond-mat/0108044) which are more sensitive to bulk properties show an even better agreement.
In our daily life we are inevitably constantly confronted with the properties of polymer materials, or more general ’soft matter materials’. On the one hand we ourselves largely consist of soft matter (bio polymers, such as DNA, proteins, and lipids which form the cell membranes ...) plus some minerals and a lot of water and, on the other hand, our daily life is not longer conceivable without goods containing polymers. The technical and commercial applications range from rather simple commodity polymer products (“plastics”) such as polyethylene for plastic bags through polymers for high tech applications such as polycarbonate for CDs or DVDs, all the way to electronic applications (laser materials, circuit boards for the semiconductor industry). While most synthetic polymers are not water soluble, some and most bio polymers certainly are water soluble. Such systems find widespread application in the environment and hygiene sector (water treatment, super absorbers) or as food additives and/or pharmaceutical products, not to speak of the natural abundance and relevance for all living organisms (e. g. cellulose). This demonstrates the relevance of ”soft matter” and with no doubt gives rise to intense technological investigations.

However, why are such systems of scientific interest? Above all applications polymers are intriguing molecules by themselves which pose many intellectual challenges in their own right. The fact that many such problems are close to biological questions and/or technical problems just adds another momentum. The characteristic feature of soft matter is their molecular structure. Unlike conventional solids or liquids, polymers are extended molecules (e. g. long chain molecules), where the overall diameter easily reaches hundreds of times of that of a typical small molecule such as water. The properties are governed by a combination of interaction energy (local submolecule-submolecule interaction) and entropy resulting from the many possible conformations such a huge molecule can obtain. Consequently, deformations of a polymer system typically only weakly change the energies but significantly the entropy per chain. Since the chains are huge, the effective particle density (polymer density) is small allowing for a relatively weak response: the systems are, compared to conventional solids, ”soft”. The typical energy scale is of the order of the thermal energy $k_B T$. This makes them unique in many respects, but also difficult to understand with conventional, though very advanced, paper and pencil theory. Therefore, already for many years attempts have been undertaken to employ computer simulations in order to study soft matter. First, dating back to the ’50s of last century, polymer chains were modeled by self repelling paths on lattices. This approach can be viewed to be of a similar importance to polymer simulations as the Ising model turned out to be for magnetism. However, as the examples of this volume show, meanwhile one goes very much beyond that approach. The different approaches employed nowadays are illustrated in the figure (courtesy F. Müller-Plathe).
The present contributions concentrate on the mesoscopic model regime. They span from test cases for finite size effects in phase transitions via morphology development and dynamics in melts and mixtures, polyelectrolytes in bulk solution and near surfaces to properties of membranes close to applications in biophysics:

Due to the extended size of the whole molecules each chain encounters the presence of many other chains, the effective interaction becomes a long range interaction. Such systems display mean field properties up to a regime very close to the critical point, where the range of the correlation functions, which diverge at the critical point, clearly exceeds the size of the chains. This crossover from mean field to the Ising critical behaviour is the main subject of the contribution by Binder, Luijten, Houdayer and Müller. In another contribution by Binder, Baschnagel and Müller chain dynamics as well as collective dynamics of melts far away and in proximity to the glass transition is studied. This also includes the phase separation kinetics and morphology developments of mixtures in bulk and especially thin films.

A second group of two papers deals with polyelectrolytes in solution (Holm, Kremer) and polyelectrolyte brushes (Seidel). Most polyelectrolytes are only water soluble due to the dissociation of small ions. The typical hydrocarbon backbone is highly hydrophobic. This leads to a competition of the ion induced solubility of the chains and an effectively attractive interaction of the chain beads. The consequences of this competition are subject of the first contribution. There, also another topic is addressed. If the electrostatic interaction becomes strong enough, an overall neutral system tries to collapse, for polyelectrolytes this means they are falling out of solution. For semi-flexible systems that leads to aggregation
phenomena, as e. g. found in DNA. This is the second topic of that paper. Ch. Seidel in his contribution addresses a similar problem however for polyelectrolyte brushes. He studies as a function of grafting density, chain length and charge density the chain conformation and determines a structure diagram.

A rather different class of systems is studied by Shillock and Lipowsky. They determine by a rather novel approach, namely dissipative particle dynamics (DPD) properties of liquid model membranes, as they are the prime constituents of biological membranes. This new method allows to study much bigger model systems compared to the so far employed standard models.

All these studies require a joint effort on the hardware side, as it is provided through the computer time grants of the NIC, and on the software side. The second is, where all the efforts of the different research groups goes in order to perform competitive simulations. The CPU time needed for the different project varied between about 50000 processor hours per year and 50000 processor hours (T3E) total for the whole project, while the number of used CPU hours on the vector machines of course is much smaller.
Phase Transitions in Macromolecular Systems: Monte Carlo and Molecular Dynamics Simulations of Coarse Grained Models

Kurt Binder\textsuperscript{1}, Jörg Baschnagel\textsuperscript{2}, Marcus Müller\textsuperscript{1}, and Wolfgang Paul\textsuperscript{1}

\textsuperscript{1} Institut für Physik, Johannes Gutenberg Universität Mainz
Staudinger Weg 7, 55099 Mainz, Germany

\textsuperscript{2} Institut Charles Sadron, Université Louis Pasteur
6 rue Boussingault, 67087 Strasbourg, France

The study of phase transitions that occur in dense polymer systems (glass transition; unmixing and wetting transitions in polymer mixtures; mesophase ordering in block copolymer melts) is very difficult, because due to the flexibility of large macromolecules there is a complicated interplay between entropic and enthalpic effects. Due to the many length scales involved in these problems, and the slow relaxation of long chain molecules, computer simulations cannot study these phenomena in full atomistic detail, but need to use coarse-grained models. Two such models, the bond fluctuations model on the lattice, and the bead-spring model in the continuum, will be described, and it will be shown that simulations of these models can help to better understand the phase behavior of polymeric materials. As characteristic examples, we shall mention static and dynamic properties of the glass transition, capillary waves of interfaces and between coexisting unmixed phases in polymer blends, and block copolymers confined between walls and their transition from parallel to perpendicular orientation of the lamellae.

1 Introduction

Macromolecules are not only the constituents of all living bodies, but are also the ingredients of the plastic materials which now are ubiquitously used. The theoretical understanding of such polymeric systems hence is of utmost importance, but also is very difficult since even the simplest polymers, linear homopolymers being formed by a succession of identical repeat units ("monomers"), exhibit structure from the scale of a chemical covalent bond ($\approx 1 \text{ rÅ}$) up to the size of the random-walk like coil ($\approx 100 \text{ rÅ}$, cf. Fig. 1). These polymer coils (in solution or melt) are very soft objects: Note that the end-to-end distance

![Figure 1. Length scales describing the structure of a long polymer coil. Note that the "persistence length" describes the scale on which correlations between bond orientations along the backbone of the polymer decay. The example shown here refers to polyethylene, $C_{N_p}H_{2N_p+2}$, where $N_0 = 2$ repeat units $CH_2$ are linked together (with CH$_3$ groups at the chain ends).](image-url)
of a random walk of \( N \) steps of step length \( b \) is \( R = bN^{1/2} \), so the volume taken by a coil is \( V = 4\pi R^3/3 = 4\pi b^3/3N^{3/2} \), but the density of the effective monomeric units represented by the individual steps of the walk inside this volume is quite small, \( \rho = N/V \propto b^{-3}N^{-1/2} \to 0 \) for \( N \to \infty \). Therefore in a dense melt each polymer coil is interpenetrated by many other coils taking to some extent the same volume, and interactions between the monomers of a coil with other polymers or with external fields (e.g. due to walls or interfaces) may lead to a significant deformation of the configuration of such a chain molecule. Therefore the interplay between configurational entropy and enthalpy is a subtle matter for polymeric systems and analytical theories can provide only a limited understanding. Although thus the use of computer simulation to understand such systems is very desirable, the complicated geometrical structure (Fig. 1) and the associated very slow relaxation of the configuration of such macromolecules are severe obstacles for a viable simulation, too! However, some aspects of this difficulty can be overcome by introducing very simplified, coarse grained models which disregard chemical detail but focus on “universal” properties of polymer chains, such as chain connectivity plus “excluded volume” interactions (no two monomers can sit precisely on top of each other). The idea behind such models is to integrate several (e.g. \( n = 5 \)) successive monomers along the backbone of a chain into one “effective bond” connecting “effective monomers”. Thus the length \( b \) mentioned above is not the length of a covalent chemical bond, but rather a “mesoscopic length” (e.g. of the order of the “persistence length”), and the number \( N \) of such effective bonds is not the degree of polymerization \( N_p \), but only \( N = N_p/n \). Also the interaction potentials are very much simplified – while the force field for a chemically realistic description of a polymer is very complicated, potentials depending on bond lengths, bond angles, torsional angles etc., for the coarse grained model one may choose a simple anharmonic spring potential along the chain and a Lennard-Jones potential among non-bonded monomers (Fig. 2). This model is not only very well suited to investigate the glass transition of polymer melts by Molecular Dynamics simulations (Sec. 2), but it can also be used to model phase separation between polymer rich and polymer poor phases in bad solvents, and to study the structure of wetting layers adsorbed on surfaces etc.

Figure 2. Effective bond potential (broken curve) and Lennard Jones (LJ) potential (full curve) plotted versus distance between effective monomers (left part). The LJ potential depends on the distance \( r \) between effective monomers as \( U_{LJ}(r) = 4\epsilon (\sigma/r)^{12} - (\sigma/r)^{6} \), and one chooses its strength \( \epsilon \) as the unit of energy scale (\( \epsilon \equiv 1 \)) and its range \( \sigma \) as unit of length scale (\( \sigma \equiv 1 \)). The bond potential (see Ref. for its precise definition) is chosen such that its minimum position (\( \ell_{\text{min}} \approx 0.96 \)) is smaller than that of the LJ potential (\( \approx 1.13 \)). This incompatibility prevents the formation of a regular crystalline structure at low temperatures, even for short chains (mostly we use \( N = 10 \)).
An alternative model, well suited for Monte Carlo simulations is the bond fluctuation model on the simple cubic lattice\(^{11–20}\). In this model, an effective monomer blocks all eight sites of an elementary cube on the lattice, and effective bonds between subsequent monomers are allowed to take the lengths \(b = 2, \sqrt{5}, \sqrt{6}, 3\) and \(\sqrt{10}\) (for this model the lattice spacing is taken as the unit of length). If one wishes to model the glass transition of polymers, one chooses an energy function \(\mathcal{H}(b)\) depending on the length \(b\):\(^{8,12–14}\) the choice \(\mathcal{H}(b = 3) = 0\) while \(\mathcal{H}(b \neq 3) = \epsilon = 1\) has the effect that bonds taking their minimum energy length \((b = 3, \text{which should happen at low temperature, since the system should approach its ground state then})\) waste “free volume”, 4 lattice sites per bond being then not accessible for occupation by any monomers. In a dense system, not enough free volume is available that all bonds can reach their ground state easily, some bonds remain blocked in unfavorable states, and this “geometric frustration” is enough to cause a glass transition.

Alternatively, when one wishes to model a binary mixture of two types of polymers (which we denote as A and B here), we may take energy parameters \(\epsilon_{AA}, \epsilon_{AB}\) and \(\epsilon_{BB}\) between different pairs of monomers, to simulate non-bonded interactions, and the unmixing behavior that results\(^{15–20}\) (Sec. 3a). In this case, we arbitrarily choose \(\mathcal{H}(b) \equiv 0\), to avoid complications due to glass-like behavior at low temperatures (unlike real polymers, where the complication of the glass transition or crystallization can never be avoided!). It is a big advantage of simulations, however, that one can focus on one physical phenomenon at a time, “switching off” interaction energies that would complicate matters, and reach a step-by-step understanding of the very complicated materials. In the conclusions (Sec. 5) we shall give some further comments on this strategy.

### 2 Simulation of the Glass Transition of Polymers

As has been emphasized in the introduction, many complicated features of real polymers have been disregarded from the start by the choice of a very simplified model. Therefore it is important to first validate the model, showing that it still reproduces all the important phenomena known from experiment. Thus, Fig. 3a shows that the specific volume of the supercooled polymer melt shows a kink at a temperature \(T_g \approx 0.41\) (remember that \(\epsilon \equiv 1\) for the model of Fig. 2, and we also choose Boltzmann’s constant \(k_B \equiv 1\), so our temperature variable \(T\) is dimensionless), while in that regime the structure factor \(S(q)\) of the melt changes only very little. As discussed in more detail in the recent review\(^8\), the simulation “data” shown in Fig. 3b exhibit a striking similarity to corresponding experiments. Therefore, it makes sense to examine the model more closely and study the slowing down of various properties as the glass transition is approached.

Due to this slowing down, very long Molecular Dynamics runs (extending up to 8 decades in time) are required, and since the configurations generated are highly correlated with each other, many equivalent “replicas” of the same system must be run in parallel and their properties must be averaged over, in order to obtain sufficiently accurate and statistically significant information. It is this high computational effort which necessitates the use of large scale supercomputer facilities such as the CRAY-T3E of the NIC Jülich.

Fig. 4 shows a typical example of the type of simulation results that can be obtained: Since many quantities (various mean square displacements, coherent and incoherent scattering functions, etc.) can be obtained for the same model simultaneously and with very
good precision, a more stringent test of theories is possible with such simulations than with corresponding experiments. In particular, simulations such as shown in Fig. 4 could provide a compelling test of the so-called “mode-coupling theory” of the glass transition which describes the blocking of diffusive motions of an atom in a dense fluid by the cage formed by its neighbors.

Of course, the question to what extent results from a model calculation are universal or reflect particular properties of a model is always a matter of concern. Therefore it is gratifying that very similar results on the glass transition of polymer melts could in fact be deduced from a rather different model, namely the bond fluctuation model on the simple cubic lattice. The diffusive motion of effective monomers there is modeled in the framework of a Monte Carlo simulation by choosing effective monomers at random and attempting to displace them by a lattice unit in randomly chosen directions. Moreover, for this model the notion of a configurational entropy of the melt has a well-defined meaning, thus allowing a stringent test of the “entropy theory” of the glass transition. According to this theory, the transition from the undercooled fluid to an underlying “ideal glass” (which is not realized experimentally, because the system falls out of thermodynamic equilibrium before it reaches this ideal phase transition) occurs when the configurational entropy of the fluid vanishes (thus the “entropy catastrophe” of a negative configurational entropy is avoided). Fig. 5 shows now the configurational entropy of the bond fluctuation model in comparison with the theoretical predictions of Gibbs and di Marzio, Flory and Milchev. One sees that these theories are rather unreliable, and while there is indeed
Figure 4. Incoherent intermediate scattering function \( \phi_q(t) = (1/N) \sum \{ \exp[i\vec{q} \cdot [\vec{r}_j(t) - \vec{r}_j(0)]] \} T \) of the model of Fig. 2 plotted vs. time (in the Molecular Dynamics time units of \( t_0 = (m\sigma^2/\epsilon)^{1/2} \), where the mass of effective monomeric units is set to unity as well). Here the sum over \( j \) is extended over all \( N \) monomeric units in the simulation box (typically \( N = 1200 \), and periodic boundary conditions are used), \( \vec{r}_j(t) \) being the position of the \( j \)-th monomer unit at time \( t \). The wave number \( q = 6.9 \) is chosen, corresponding to the peak position in Fig. 3b, in order to study the dynamics of a particle associated with the “cage” of its nearest neighbors surrounding it. Several theoretical approximations are included: the approximation that all density fluctuations are Gaussian distributed amounts to \( \phi_q(t) = \exp[-q^2 g_0(t)/6] \), with \( g_0(t) \) the mean square displacement of monomers after time \( t \). This approximation describes only the early stages of relaxation, where a particle is still well confined in its “cage” (broken curve). The curve marked “\( \beta \)-correlator” describes the result of a mode coupling analysis, for the relaxation towards a plateau (described by the “nonergodicity parameter” \( f_q \)) and its initial decay with the asymptotic decay is described by the “\( \alpha \)-relaxation”, namely a decay with the so-called KWW stretched exponential relaxation, \( \phi_q(t) \propto \exp[-(t/\tau)^{\beta K}] \), with \( \beta K \approx 0.75 \). Repeating this analysis at a range of temperature (the present figure refers to \( T = 0.48 \)), one finds that the relaxation time \( \tau \) scales as \( \tau \propto (T/T_c - 1)^{-\gamma} \), with \( \gamma \approx 2.1, T_c \approx 0.45 \). This behavior is well compatible with predictions of the mode coupling theory (MCT). From Baschnagel et al.\(^6\).

3 Interfaces between Coexisting Phases in Unmixed Polymer Blends

Polymer mixtures are technologically very important since mixing several polymers together often improves the physical properties of the resulting material. However, due to chain connectivity the configurational entropy is smaller by a factor of \( 1/N \), in comparison with a small molecule mixture. Therefore many polymer mixtures are not homogeneous on atomistic length scales, but rather phase separation in a binary (A,B) mixture has occurred,

...
such that mesoscopic domains of A-rich and B-rich phases coexist, separated by interfaces (phase separation on a macroscopic scale can be avoided by quenching the system fast enough underneath the glass transition temperature, or by introducing chemical cross links which limit the length scale over which unmixing can occur, etc.). Since most real polymer blends contain lots of interfaces, the latter control the properties of the material, and hence a theoretical understanding of such interfaces is called for.

This problem can be addressed with the bond fluctuation model, where we allow for two types of chains (A,B) with identical chain lengths \( N_A, N_B \) and interactions \( \epsilon_{AA} = \epsilon_{BB} = -\epsilon_{AB} = \epsilon \), if the two monomers are within a range of \( \sqrt{6} \) lattice spacings. For this model, the bulk phase diagram could be obtained with very good precision\(^{25}\). In order to prepare a system which contains a single interface, it is convenient to use the so-called “antiperiodic boundary condition”: In one lattice direction one introduces
an identity switch ($A \rightarrow B$ or $B \rightarrow A$), if a monomer crosses the system boundary in that direction. Fig. 6 shows a snapshot picture of a system prepared in this way, for the case $N_A = N_B = N = 32$ and a temperature far below the critical temperature $T_c$ of unmixing, $T/T_c = 0.48$. One sees that at this temperature within the coexisting domains phase separation is essentially complete, i.e. a pure A-phase coexists with a pure B-phase, intermixing occurs only in the interfacial region. Therefore it is advisable to focus on the interface behavior, and disregard the structure of the bulk phases completely. It is one of the major advantages of computer simulations, that one can extract selectively the pertinent information.

Figure 7. Snapshot picture of an instantaneous configuration of a well-equilibrated interface between an A-rich and a B-rich phase of polymers of chain lengths $N_A = N_B = N = 32$ at $T/T_c = 0.48$. Here a $L \times L \times D$ geometry was chosen, with $L = 64$, $D = 64$, periodic boundary conditions in $xy$-directions, but antiperiodic boundary conditions in the $z$-direction, so that precisely one interface is stabilized in the systems. The $xy$-plane at $z = 0$ is divided into $B \times B$-blocks with $B = 8$, and for each block the local height $h(x, y)$ of the interface defined from the condition that the relative concentration of A and B is locally 50% is shown. From Werner et al.\textsuperscript{16}.

Fig. 7 shows a corresponding snapshot picture where just the surface representing the instantaneous position of the center of the interfacial region is shown\textsuperscript{16}. One recognizes a “hill and valley”-type structure, i.e. there occur pronounced long wavelength fluctuations that displace the local position of the interface up and down. These interfacial fluctuations are the so-called “capillary waves”, which turn out to be very pronounced for polymer mixtures, since the interfacial tension is rather small for these systems. When one averages interfacial profiles over these fluctuations, allowing for a large lateral linear dimension, the interfacial profile would seem to be much broader rather than when one takes a local average only. Thus the judgment what is attributable to the “intrinsic” interfacial profile and what is due to capillary wave broadening is rather subtle\textsuperscript{16–18}.

Nevertheless the simulations yield a wealth of illuminating information on the structure of such interfaces in polymer blends: in the interface, polymer coils get oriented parallel to the interface, and in the strong segregation limit even individual effective bonds have a preferential orientation parallel to the interface\textsuperscript{17}. The density profile shows a dip in the center of the interface, i.e. vacancies get enriched there, and also chain ends are enriched. Knowledge of such properties is important if one wishes to discuss the mechanical strength of interfaces, chemical reactivity at interfaces, etc.

A very interesting aspect occurs when interfaces are confined in a thin film geometry (e.g. by external walls which create a binding potential acting on the interface): It then may happen that the interface at low temperatures is bound to the wall (as shown in Fig. 8),

\[369\]
Figure 8. Snapshot picture of an instantaneous configuration of an interface between an A-rich and a B-rich phase of polymers in a thin film with “antisymmetric” walls (i.e., one wall preferentially attracts the A-monomers and the other wall attracts B-monomers with the same strength). The system geometry chosen is $128 \times 128 \times 32$, with periodic boundary conditions in $x$ and $y$ directions, while the two walls are located at $z = -16$ and $z = 17$, respectively. The chain length is again $N = 32$, $\epsilon_w = 0$, and the temperature $T = 0.48T_c$.

while at higher temperatures an “interface unbinding” transition occurs. In a semi-infinite system, this is nothing but a wetting transition where a macroscopically thick enrichment layer of the phase preferred by the wall forms. In a thin film geometry, the analog of wetting is the interface bound to confining walls in a thin film. When the temperature is raised, there occurs a discontinuous (= first order) transition to the situation of an unbound interface, freely fluctuating in the center of the thin film, qualitatively very similar to an interface between macroscopic coexisting phases (as shown in Fig. 7).

4 Thin Films of Block Copolymer Melts

Figure 9. Snapshot pictures of the configurations of thin block copolymer films, choosing $N = 32$, $f = 1/2$, and an inverse temperature $\epsilon/k_B T = 0.1769$ for the geometry $30 \times 96 \times 96$, $46 \times 93 \times 93$ and $56 \times 96 \times 96$, choosing the wall energy $\epsilon_w = 0.1k_B T$ and periodic boundary conditions parallel to the walls. From Geisinger et al.\textsuperscript{20}.

In order to synthesize a block copolymer of composition $f$ and chain length $N$, one covalently links a chain of type A (chain length $N_A$) and another chain of type B (chain length $N_B$) at one end together, to form a chain of total length $N = N_A + N_B$ and $f \equiv N_A / N$. As in the case of the polymer mixture, energies $\epsilon_{AA}$, $\epsilon_{BB}$, $\epsilon_{AB}$ between the effective monomers are present, which would lead to macroscopic phase separation between species A and B if there were not the covalent bond keeping the A-block and the B-block together: The “compromise” chosen by the system then is “microphase separation”.

370
i.e. a periodic arrangement of A-rich and B-rich domains is formed. For compositions in the vicinity of \( f = 1/2 \), the arrangement of A-rich layers alternating with B-rich ones, the wavelength \( \lambda \) of this structure being of the same order as the gyration radius of the polymers, in the vicinity of the order-disorder transition of the block copolymer melt. Note that long range periodic order forms here in a system that is fluid, but again one can produce a solid material by performing a quench underneath the glass transition temperature.

A particularly interesting situation again occurs when one considers block copolymers in thin film geometry, confined between two parallel walls at a distance \( D \). Normally, there will be a preferential attraction of one of the blocks to the walls, causing hence an arrangement of the lamellae parallel to the confining walls. However, the lamellar arrangement will nicely fit into the thin film only if the wavelength \( \lambda \) and the film thickness \( D \) are commensurate, and there is no reason that this holds true in general (note that \( \lambda \) depends both on \( N \) and on temperature\(^{26}\)). In the case of strong incommensurability, it becomes energetically preferable to orient the lamellae perpendicular to the walls rather than parallel, and such a perpendicular orientation has been observed both experimentally and by computer simulations\(^{20}\), see Fig. 9. The perpendicular oriented morphology has attracted abiding interest as a template for laterally structured devices on the nanometer scale. Simulations as shown in Fig. 9 are very important, since they allow a better understanding of the conditions for which this perpendicular ordering is stable.

Of course, due to the low temperatures necessary for this ordering the relaxation times are very long, and runs of a length of several million Monte Carlo steps are required. For the rather large lattices used in Figs. 6-9 a use of CRAY-T3E supercomputers again is mandatory, making use of a strip-like domain decomposition parallel to the walls for an efficient use of up to 64 processors in parallel.

5 Concluding Remarks

Three examples have been discussed in order to show that the large scale simulation of simple models for polymer chains yields valuable insight in the phase transitions that occur in polymeric materials, namely the glass transition, unmixing of symmetric polymer blends, particular in confined geometry, and mesophase order-disorder transitions of thin block copolymer films. All these phase transitions at the same time are challenging problems of soft matter physics and important from the point of view of materials research. Thus, the results (described in more detail in the original papers\(^ {2-20} \)) have both elucidated controversial theoretical questions and stimulated experiments. Nevertheless it is clear that such work is a first step only – we still lack the possibility to predict the properties of specific polymeric materials (e.g. where the glass transition of a particular polymer occurs, and how we can change it when we modify the chemical structure), and also on a qualitative level fascinating problems are still out of reach, such as the interplay of phase separation and glassification, or the more complex ordering of multiblock copolymers and the interplay between phase separation and ordering in mixtures of different block copolymers, etc. Such problems will become accessible only with the next generation of even more powerful supercomputers.
Acknowledgments

We thank M. Aichele, C. Bennemann, B. Dünweg, T. Geisinger, L. G. MacDowell, F. Schmid, A. Werner, and M. Wolfgardt for their fruitful collaboration on the problems described here. This research was supported by the Deutsche Forschungsgemeinschaft (DFG) under grants No SFB262/D2, Bi314/3, Bi314/17 and the Bundesministerium für Bildung und Forschung (BMBF) under grants No 03N8008C and 03N6015. We are particularly grateful to the NIC Jülich and the HLRS Stuttgart for generous allocations of computer time on CRAY-T3E supercomputers.

References

Critical Phenomena and “Crossover Scaling” in Model Systems for Soft Condensed Matter

Kurt Binder, Jérôme Houdayer, Erik Luijten, and Marcus Müller

1 Institut für Physik
Johannes Gutenberg-Universität
Staudinger Weg 7, 55099 Mainz, Germany

2 Max-Planck-Institut für Polymerforschung
Postfach, 55128 Mainz, Germany

3 Department of Materials Science and Engineering
University of Illinois at Urbana-Champaign
Urbana, IL 61801, USA

Critical points where phase separation between coexisting phases sets in are quite common in systems such as micellar solutions, polymer solutions, mixtures of homopolymers and copolymers, and related soft matter systems. While the asymptotic critical behavior associated with this phase separation should be the same as for simple Ising ferromagnets, the mesoscopic scale of the molecular constituents of these systems often causes a broad regime of parameters where simple mean-field type behavior is observed. Determination of the conditions for which a crossover between these two types of critical phenomena occurs constitutes a crucial test of our understanding of the effective interactions in these systems. In this report, two representative examples are described: (i) a ternary system of solvent plus polymer A plus polymer B, where solvent–polymer unmixing competes with polymer–polymer phase separation; (ii) a melt of random copolymers, where macroscopic phase separation and microphase separation compete. As a prelude, the Ising to mean-field crossover is studied by Monte Carlo simulation of an Ising magnet with a large but finite interaction range, providing accurate reference data for a crossover scaling description. It also is shown that new and more efficient simulation algorithms need to be (and have been) developed in order to be able to tackle such problems.

1 Introduction

“Soft condensed matter” is a synonym for a class of systems that are somewhere in between simple fluids and “hard” solids: Colloidal dispersions, gels, microemulsions, membranes, polymeric materials, etc. These systems have very broad industrial applications (production of lubricants, paints, perfumes, food industry, biotechnology and pharmaceutical industry, etc.), and much empirical knowledge has been accumulated about them; however, only rather recently these systems have moved into the mainstream of scientific basic research, and many aspects of their behavior are still not completely understood. In particular, the “soft” character of these materials can, to a large extent, be attributed to a delicate interplay between energy and entropy on the atomistic scale, giving rise to non-trivial mesoscopic structure, and the interactions between effective units on the mesoscopic scale are often entropic in origin. Understanding how to control the mesoscale structure and the effective interactions is a key aspect in order to systematically improve the properties of these materials and to develop novel applications.

One phenomenon that is decisively influenced by the nature and range of effective interactions on the mesoscale is the critical behavior associated with the onset of phase sepa-
ration into two distinct phases that coexist with each other. For instance, if the quality of a solvent used to dissolve flexible polymers is deteriorated by lowering its temperature $T$ below the so-called “Theta temperature” $\Theta$, phase separation may occur into a polymer-rich solution and a phase consisting essentially of pure solvent, with very few isolated and collapsed chains remaining in this solution. For very high molecular weight (or, equivalently, chain length $N$) of the polymer, the phase diagram is very asymmetric, the critical concentration $\phi_{\text{crit}}$ of the polymer being rather small, and for $N \to \infty$ one even has $\phi_{\text{crit}} \to 0$, and the critical temperature $T_c(N) \to \Theta$. This particular limit causes a special crossover of critical properties from one class of behavior to another one, which has been intensively studied by experiments, theory, and simulation, and is now rather well understood. However, when we consider the extension to the case where two kinds of polymers ($A, B$) are dissolved in the same solution, the situation becomes much more intricate: In the polymer-rich solution, one must expect that a further phase separation occurs, into a phase rich in polymer $A$ and another phase rich in polymer $B$. Thus, in the same polymer solution, two rather different types of phase separation occur and compete with each other, and the subtle interplay of these phenomena is clearly of interest. This problem will be described in more detail in Section 3.

In problems of the type described above, a generic characteristic length scale occurs, such as the gyration radius of a polymer coil, which in a sense is analogous to the range of interactions in a system of small molecules, as far as the character of the critical phenomena is concerned. In order to understand the basic notions of how one can “move” from one type of critical behavior to another, it is therefore useful to study the model problem of a simple Ising lattice system, where the range of exchange interactions between the magnetic moments can be varied to a great extent (Section 2).

Instead of considering competition between two types of regular phase separation, where ultimately the spatial inhomogeneity develops on macroscopic scales, it is also possible to have a competition between macroscopic phase separation and mesophase ordering. While periodically modulated mesophase structures are rather well understood for systems such as liquid crystals, monodisperse block copolymers, etc., a more irregular form of mesophase or “microphase separation” is expected for random copolymers, where monomers of type $A$ and monomers of type $B$ alternate along the chain backbone in an irregular fashion. Such random heteropolymers can be viewed as a crude, qualitative model of biopolymers, which makes their study of particular interest as well. First steps to simulate the phase transitions of random copolymer melts will be described in Section 4.

For all these problems, critical points can be located by extensions of the well-known “finite-size scaling” technique, which requires very accurate computer simulation results over a broad range of lattice sizes for the models studied. “Brute-force”-type application of supercomputers would not lead very far—due to the mesoscopic length scales involved (medium interaction range or coil size, etc.) very large systems are needed, and near critical points the problem of strongly correlated system configurations (“critical slowing down”) presents an additional complication. For the heteropolymers, the low temperatures of interest imply that glass-like relaxation is an additional problem, and very efficient sampling techniques are needed to cope with this problem. Thus, a key feature of the present project has been the development of new and more efficient algorithms, as well as their careful testing and demonstration of viability. Consequently, physical applications are still in their early stages.
2 A Prelude: The Ising to Mean-Field Crossover in the Medium-Range Ising Model

The Ising model is a lattice model for a uniaxial ferromagnet, where variables (“spins”) \( S_i = \pm 1 \) designate whether the magnetic moment at a lattice site \( i \) is oriented up or down along the preferred axis. Ferromagnetic order is induced by an exchange-type interaction energy \( E = -J_{ij} S_i S_j \) with a positive “exchange constant” \( J_{ij} \) between spins at sites \( i \) and \( j \).

In statistical mechanics, this is the prototype model to study phase transitions, and it is also the “yard stick” against which the efficiency of computer simulation programs, algorithms, analysis techniques, etc. is measured (for a detailed account, see the recent review Ref. 1, which actually grew out of the earlier stages of this NIC (HLRZ) project). The “order parameter”, namely the magnetization per spin, is large but finite. This problem is difficult because of “crossover”: For small enough distance \( R \) for the case where \( r < R_m \) and \( J_{ij} \equiv 0 \) for \( r > R_m \) the problem becomes trivial: Then \( T_c = z J/k_B \) (\( k_B = \text{Boltzmann’s constant} \) and \( \beta = 1/2 \) i.e., mean-field theory (or, equivalently, the Weiss molecular-field theory) holds exactly.

We now wish to find \( T_c \) and the critical exponents (of the magnetization and of other quantities, such as the magnetic susceptibility, i.e., the response function to an external magnetic field \( H \)). \( \chi = (\partial m/\partial H)_{H=0} \propto |1 - T/T_c|^{-\gamma} \), \( \gamma \) being another critical exponent) for the case where \( R \) is large but finite. This problem is difficult because of “crossover”: For small enough distance \( t \equiv T/T_c - 1 \) from the critical point, the exponents will take the universal values of the “Ising universality class”, but the region where they can be observed will decrease with increasing \( R \). Outside this region, one finds “effective exponents”, e.g., \( \gamma_{\text{eff}}(t) \equiv -d \ln \chi/d \ln |t| \). Since the power law \( \chi = \hat{\chi} |t|^{-\gamma} \) implies a straight line with slope \( \gamma \) on a log–log plot of \( \chi \) vs. \( |t| \), the effective exponent is simply the “local slope” of the corresponding function on the log–log plot. Figure 1 gives an example for the susceptibility exponent \( \gamma_{\text{eff}} \) (the superscript minus stands for the sign of \( t \)) for \( T < T_c \) and \( d = 2 \) dimensions (where one knows that \( \gamma = 7/4 \) exactly while the mean-field result is always \( \gamma_{\text{MF}} = 1 \)).

One can see that there is a nontrivial (and in this case even nonmonotonic) variation of the effective exponent with the scaling variable \( t R^2 \), and there is no analytic theory available that would yield this information! Also for a Monte Carlo simulation such an estimation of effective exponents is a tour de force; therefore, significant error bars as shown in the figure are still inevitable, and the full crossover (which spreads out over 5 decades in \( t R^2 \)) cannot be displayed for a single choice of \( R \), but only by combining many values of \( R \) — that probe different parts of the crossover scaling function — in a single plot.

Why is such a study so difficult? To answer this question, we recall that Monte Carlo
Figure 1. Effective susceptibility exponent $\gamma_{\text{eff}}$ of the two-dimensional Ising square lattice for various choices of the interaction range $R$ (denoted by different symbols, as given by the key in the figure), plotted vs. the crossover scaling variable $tR^2$ (note the logarithmic scale of the abscissa).

studies amount to simulating a finite $L \times L$ lattice (in $d = 2$) with periodic boundary conditions, and for finite lattices critical singularities are in fact rounded: Asymptotically, for $L \to \infty$, this rounding of singularities is described by finite-size scaling theory (see, e.g., Ref. 3 for a review), and actually this theory is needed for a precise location of $T_c$ and estimation of critical exponents from the simulation.1,3–5 However, when $R$ is large, the asymptotic region of finite-size scaling is only reached when $L/R^2 \gg 1$ (in $d = 2$) or even $L/R^4 \gg 1$ (in $d = 3$), and thus it follows at once that huge lattice sizes $L$ are needed. For the standard single-spin-flip Monte Carlo algorithm,3,4 which suffers from “critical slowing down” (i.e., subsequently generated system configurations in the critical region are correlated over a “time” $\tau \propto L^z$ Monte Carlo steps (MCS) per lattice site, with a “dynamic exponent” $z \approx 2$1,4,5), this task would be hopeless even for the most powerful supercomputers available today! However, this difficulty could be overcome by an elegant extension of the single cluster algorithm6 to systems with long-range interactions.7 The cluster algorithms of Refs. 6, 7 reduce the critical slowing down dramatically ($\tau \propto \ln L$ in $d = 2$). The important point of the algorithm of Luijten and Blöte7 is that, despite the long-range interaction, the computational effort scales only linearly (and not quadratically) with the number of sites in the generated cluster.

Even with the availability of this novel algorithm, the need to study very large linear lattice dimensions with very high statistical accuracy, for many choices of the three parameters $L$, $t$ and $R$, made the use of a very powerful computer such as the CRAY-T3E indispensable: Due to the relatively small correlation time $\tau$, the task could be split over many processors collecting the data within a reasonable total time.

Why is such an Ising to mean-field crossover relevant for soft matter systems? This
Figure 2. Plot of the scaled order parameter $\langle |m| \rangle N^{1/2}$ of symmetrical polymer mixtures plotted vs. the scaled temperature distance $N|t|$ from the critical point. Different symbols denote the different chain lengths investigated, as indicated in the figure. Broken straight line shows the mean-field result. All data refer to simulations of the bond fluctuation model of polymer chains on the simple cubic lattice, for a density $\phi = 1/2$ of occupied sites.

The question is most easily answered for a binary polymer mixture $(A, B)$, with symmetrical chain lengths $N_A = N_B = N$ (unlike the following section, we do not consider the additional effects of a solvent here). Given the fact that a flexible polymer chain in a melt has a random-walk-like structure, i.e., the gyration radius $R_g$ scales like $R_g = bN^{1/2}$, $b$ being the size of an effective segment, we conclude that each chain takes a volume $V = (4\pi/3)R_g^3 \propto b^3 N^{3/2}$. However, the density of segments of the chain inside this volume is rather small, $\rho = N/V \propto b^{-3} N^{-1/2}$; much smaller than the total segment density in the melt ($\rho \propto b^{-3}$). As a consequence, one concludes that $N^{1/2}$ chains must share the same volume and that each chain interacts with $z = N^{1/2}$ "neighbors". Upon noting that the order parameter of a polymer mixture $[m = (n_A - n_B)/(n_A + n_B), n_A, n_B$ being the number of chains of type $A, B$, respectively] corresponds to the magnetization of the Ising problem, one can essentially provide a mapping between the medium-range Ising problem and symmetrical polymer mixtures. In fact, these arguments are supported both by experiments and by direct simulations of symmetrical polymer mixtures—although some problems do remain. So far, the existing direct simulations of critical properties of polymer mixtures (e.g., Fig. 2) do give some evidence for crossover because of the curvature seen on the log–log plot, but the data do not allow to obtain effective exponents via logarithmic derivatives yet. Extending studies such as shown in Fig. 2 belongs to the goals that are still under consideration in the present project.
Critical Phenomena in Ternary Systems: Polymer A + Polymer B + Solvent

As for Fig. 2, we use the bond fluctuation model on the three-dimensional simple cubic lattice, but now the density $\rho$ of monomers is not fixed at the melt density, but rather a dynamical variable controlled by the appropriate chemical potential. In this model, effective monomers are modeled by elementary cubes on the lattice (all the 8 corners of the cube being blocked from further occupation), and are connected by effective bonds which may have lengths $b = 2, \sqrt{5}, \sqrt{6}, 3$ or $\sqrt{10}$ lattice spacings (on the simple cubic lattice). The idea behind this model is to perform, for a chemically realistic model of a chain, a coarse-graining along the backbone of the chain: about $n \approx 5$ chemical monomers correspond to one “effective bond” of the bond fluctuation model. The spread of effective bond lengths $b$ is thought to come from the various conformational structures that a group of $n$ monomers may take, due to the various choices of torsional angles. From such a mapping, one may conclude that a lattice spacing physically corresponds to about $2r_A$, and the physical degree of polymerization $N_p$ corresponds to the chain length $N$ via $N = N_p/n$. In order to deal with two types of chains ($A, B$), we choose chain lengths $N_A, N_B$, and interaction parameters $\varepsilon_{AA}(\vec{r}), \varepsilon_{BB}(\vec{r}), \varepsilon_{AB}(\vec{r})$. For simplicity, all these interactions are taken of the square-well type, i.e., constant and nonzero up to the range $R_m$ (we choose $R_m = \sqrt{6}$) and zero elsewhere. Only the most symmetric case is considered, namely $N_A = N_B = N$ and $\varepsilon_{AA} = \varepsilon_{BB} = -\varepsilon_{AB}$. Thus, the internal structure and length of the two types of chains are identical; they differ only by their label ($A$ or $B$). The solvent is not accounted for explicitly (it is just represented by the vacant sites). If we would consider a dense melt (volume fraction of occupied sites $\phi = 1/2$ or larger), only the combination $\varepsilon_{AB} - (\varepsilon_{AA} + \varepsilon_{BB})/2$ of the interaction constants would matter, and the only phase transition to consider is a phase separation between an $A$-rich phase and a $B$-rich phase (this was the situation treated in the model of Fig. 2). However, for small $\phi$ (or small monomer density $\rho = \phi/8$, remembering that each monomer blocks 8 lattice sites) this phase separation between two polymer phases competes with the phase separation between polymer and solvent.

As an example, the phase diagram of such a system is shown for the case of $N = 20$ in Fig. 3. Even for relatively short chains (such as $N = 20$) the construction of this phase diagram becomes very difficult, since in the grand-canonical ensemble polymer chains have to be inserted in a very dense system (note that $\rho_{\text{coex}}(\tau) \geq 0.5$ for $T^* = T/\varepsilon_{AB} < 1.70$). Even the use of the well-know “configurational bias” algorithm would not allow to carry out such simulations with the necessary accuracy. Therefore the recently invented “recoil growth” algorithm, which is a generalization of the “configurational bias” algorithm suitable for very dense polymer systems, has been adapted to the present problem, and the data shown in Fig. 3 are actually generated with this method. The idea behind all these techniques is that a chain is not inserted into a dense system in one step—which would never be an accepted Monte Carlo move, since the excluded volume condition would always be violated—but one inserts only one end of the chain, and from there on it is “grown” segment by segment. This growth is not done blindly, but rather occurs in a biased way, so that the chain “searches” for empty regions of the system into which it can still grow. Of course, in the simulation one must correct for the applied bias, and great care is required in order to ensure that the algorithm respects the “detailed balance” condition—we refer the
Figure 3. Phase diagram in the plane of variables $T^* = T/\varepsilon_{AB}$ and polymer concentration (or volume fraction of occupied sites) $\phi$, for a binary polymer solution with $N_A = N_B = N = 20$. The curves were obtained from simulations of the bond fluctuation model for three lattice linear dimensions ($L = 40, L = 60$, and $L = 80$, respectively). The “gap” near $T^* = 1.78$ corresponds to the vicinity of the critical point, where the numerical data are strongly affected by finite-size rounding. The left part of the coexistence curve ($\rho_{\text{coex}}^{(1)}$) separates the solvent-rich phase from the two-phase region, the right part ($\rho_{\text{coex}}^{(2)}$) separates the two-phase region from the polymer-rich phase. The curve in the center of the two-phase region is the rectilinear diameter, $\rho_{\text{rd}}(T) = [\rho_{\text{coex}}^{(1)}(T) + \rho_{\text{coex}}^{(2)}(T)]/2$. Note that the phase separation between $A$-rich and $B$-rich polymer phases (curve not shown) hits the coexistence curve $\rho_{\text{coex}}^{(2)}(T)$ at a “critical end point” (CEP), causing a singularity both in $\rho_{\text{coex}}^{(2)}(T)$ and in $\rho_{\text{rd}}(T)$, just above $T^* = 1.74$.}

Figure 3 highlights the effect of the polymer $A$–polymer $B$ phase separation (that occurs in the polymer-rich phase) on the polymer–solvent phase diagram: When the critical line of the $A$–$B$-separation hits the coexistence curve $\rho_{\text{coex}}^{(2)}(T)$ at the “critical end point” (CEP) the coexistence curve shows a rather pronounced wiggle. This anomaly in the phase diagram, however, does not lead to a particular crossover for the critical exponents. Only when the CEP would coincide with the polymer–solvent critical point, the two coincident critical phenomena would change their character into a “tricritical” phenomenon, and interesting crossover scaling would occur. To realize this tricritical point, other choices than $\varepsilon_{AA} = \varepsilon_{BB} = -\varepsilon_{AB}$ have to be considered.

A special behavior is also predicted for the good solvent case, $\varepsilon_{AA} = \varepsilon_{BB} = 0$, $\varepsilon_{AB} \neq 0$. Now, there is no polymer–solvent phase separation, and polymer $A$–polymer $B$ phase separation can occur only if $\phi$ exceeds some critical value $\phi_p$, of the same order as the overlap concentration $\phi^*$. For $\phi$ slightly less than $\phi_p$, instead of bulk phase separation one expects large aggregates of slightly overlapping $A$-chains, separated from large aggregates of slightly overlapping $B$-chains (Fig. 4). The phase transition at $\phi \rightarrow \phi_p$ then is a kind of continuum percolation of soft spheres, each sphere representing a whole poly-
Figure 4. Qualitative picture of a system consisting of polymer $A$ + polymer $B$ + solvent, under good solvent conditions, but with large repulsive interaction $\varepsilon_{AB}$, for a volume fraction $\phi$ slightly less than a "percolation concentration" $\phi_p$ where macroscopic aggregates of chains of one type appear.

mer chain. Due to the huge lattice linear dimensions that would be required to study this transition, somewhat more powerful computers than presently available are clearly needed.

4 Phase Separation in Melts of Random Copolymers

A random copolymer is composed of two types of monomers $A$ and $B$, which form a random sequence with average composition $f$ (only $f = 1/2$ is studied here) and a parameter $\lambda$ that describes the correlation of composition along the chain. For $\lambda = 0$, the sequence is completely uncorrelated, while for $\lambda = -1$ we have a regularly alternating copolymer $(ABABAB\ldots)$ and for $\lambda = +1$ we have homopolymers (the ratio of $A$ and $B$-homopolymers is determined by $f$). Note that the random sequence, once created, is fixed, since this sequence is determined by the chemical synthesis of the chain.

Such random copolymers are often studied with a motivation to understand biopolymers; remember that a single strand of DNA is a (quasi-random) sequence of 4 types of amino acids (the detailed sequence then contains the genetic information). Here, however, we do not have biologically motivated questions, such as protein folding, in mind, but rather ask what is the phase behavior of a melt of short copolymer chains (of chain lengths $N = 5, 10, 20$), as a function of the blockiness parameter $\lambda$ and the (repulsive) interaction $\varepsilon_{AB}$ between monomers of different kind. In the block-copolymer case, one knows that
a microphase separation occurs where an alternating arrangement of A-rich and B-rich domains is stable (“lamellar mesophase”), but it is clear that this structure will get more and more disrupted as the randomness increases. Note also that the randomness has the effect that individual (short) chains have a notable surplus in A or B—the composition $f = 1/2$ holds only on average. As a consequence, macroscopic phase separation into A-rich and B-rich phases is possible, and has actually been found in our simulations (Fig. 5, see Ref. 20 for details). However, in the region where both $\lambda$ and $\varepsilon_{AB}/k_BT$ are large, an additional phase occurs which could be a kind of “microemulsion” (see the configuration snapshot, Fig. 6), or a lamellar mesophase somewhat distorted by randomness, or a two-phase region between a microemulsion and the lamellar phases. Note that microemulsions normally form in ternary systems $A, B$ plus surfactant (which stabilizes $A$–$B$ interfaces), and depending on the compositions of these constituents also regions of lamellar order or two-phase-coexistence regions occur in the phase diagram, in addition to a sponge-like phase. It could be that our melt of random copolymers is a kind of disordered analog of such a simple microemulsion.

In order to be able to obtain the results shown in Figs. 5 and 6, it was first of all necessary to develop a new, more efficient version of the “slithering snake” algorithm in order to be able to equilibrate the model system, and long runs at the CRAY-T3E were indispensable to establish the large-scale structures seen in Fig. 6. Very recently, a new generalization of the configurational-bias algorithm (where chains are removed in one part of the lattice and “regrown” in another part) was developed,\textsuperscript{21} which hopefully will allow to clarify the open questions about the phase diagram and critical behavior of this system.
Figure 6. Snapshot of a microemulsion-like state of a random copolymer melt, for system parameters $N = 10$, $\lambda = 0.75$, $\varepsilon_{AB}/k_BT = 0.8$, and a $120^3$ lattice.

5 Concluding Remarks

In this paper, it has been shown that various types of phase separation in soft matter systems (polymer–solvent, polymer $A$–polymer $B$, random copolymers with excess of one kind of monomer versus excess of the other type of monomer) give rise to very interesting phase diagrams and interesting critical phenomena. However, the large scale provided by the large characteristic lengths, together with the common difficulties of polymer simulations (slow configurational relaxation of long chains, difficult chain insertions in dense systems, etc.), render the study of such phenomena by means of large-scale computer simulations particularly difficult. Developing new algorithms targeted to ease these problems, together with the availability of huge computer resources on very efficient parallel supercomputers, has proven crucial in order to achieve progress. Two of us (E.L., J.H.) were supported by the Max-Planck Society via Max-Planck fellowships. We are grateful to Prof. H.W.J Blöte, Dr. H.-P. Deutsch, Dr. A. Sariban, and Dr. N.B. Wilding for fruitful collaboration on partial aspects of the research described here. We are particularly grateful to the NIC Jülich for a generous allocation of computer time on the CRAY-T3E.
Acknowledgements

Two of us (E.L., J.H.) were supported by the Max-Planck Society via Max-Planck fellowships. We are grateful to Prof. H.W.J. Blöte, Dr. H.-P. Deutsch, Dr. A. Sariban, and Dr. N.B. Wilding for fruitful collaboration on partial aspects of the research described here. We are particularly grateful to the NIC Jülich for a generous allocation of computer time on the CRAY-T3E.

References

Computer Modeling of Charged Polymers

Christian Holm\textsuperscript{1}, Kurt Kremer\textsuperscript{1}, Markus Deserno\textsuperscript{2}, and Hans Jörg Limbach\textsuperscript{1}

\textsuperscript{1} Max-Planck-Institut für Polymerforschung
Ackermannweg 10, 55128 Mainz, Germany
E-mail: \{holm, kremer, limbach\}@mpip-mainz.mpg.de

\textsuperscript{2} Department of Chemistry and Biochemistry
UCLA, Los Angeles, USA
E-mail: markus@chem.ucla.edu

We give a short introduction into polyelectrolytes and show some of the problems of modeling charged polymer systems. We concentrate on two model systems. The first one consists of flexible polyelectrolytes exposed to a poor solvent. These polymers show interesting conformational behavior. We will discuss them and hint to the difficulties of observing conformational transitions in light scattering and AFM single molecule experiments. The second example is relevant for stiff charged systems like DNA solutions. Here we demonstrate that two like-charged macromolecules can attract each other under certain circumstances.

1 Introduction

Polyelectrolytes are polymers which have the ability to dissociate charges in polar solvents which results in charged polymer chains (macroion) and mobile counterions. They represent a broad and interesting class of soft matter\textsuperscript{1} that enjoys an increasing attention in the scientific community. For example, in technical applications polyelectrolytes are used as viscosity modifiers to reduce drag, for example, in oil pipelines, or to make low-fat dairy products creamy. They are used in sewage plants to clean water by precipitating heavy metal ions. They are also responsible for the ability of sanitary napkins or baby-diapers to absorb enormous amounts of water, and are hence called superabsorbers; see also Fig. 1 for an example of such an hydrogel.

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 biopolymer such as DNA, actin, or microtubules are polyelectrolytes. Also the cell membrane is charged, and electrostatic interactions between biomolecules are thought to play an important role in understanding the conformations and functions of biomolecules and their complex interactions\textsuperscript{2}. For example, in biotechnology, one is interested in the structure of synthetic liposomes, complexes of negatively charged DNA and positively charged lipids, which could serve as an efficient vehicle of transporting pieces of DNA into a cell. As another example, by adding multivalent counterions into a DNA solution one induces attractive interactions between like charges, leading to compact structures, such as tori, compare Fig. 1. The DNA in viruses is believed to be compactified due to such attractive interactions, and these play also a prominent role in the folding process of chromosomal DNA\textsuperscript{3}.

In contrast to the theory of neutral polymer systems, which is well developed, the theory of polyelectrolytes faces several difficulties. Simple scaling theories, which have been proven so successfully in neutral polymer theory, have to deal with additional length scales
Figure 1. a) shows DNA ejected from a virus into a liposome (a model bacterial cell) vesicle (adapted from Ref. 3) and b) depicts a swollen polyelectrolyte gel which can absorb huge amounts of water.

set by the long range Coulomb interaction. Furthermore there is a delicate interplay between the electrostatic interaction of the distribution of the counterions and the conformational degrees of freedom, which in turn are governed by a host of short range interactions, which renders the problem difficult. There are only two limiting cases that are easy to solve. These are the case of high salt excess, effectively screening out the electrostatic interaction (which in turn allows one to treat it as a perturbation), or the case of an overwhelming dominance of the Coulomb force, which results in a strongly elongated chain. Unfortunately it is often just the intermediate case, which proves to be the most interesting regime in terms of application, experiment and theory. The mean-field theory of charged systems is the Poisson-Boltzmann (PB) equation. Analytical or numerical solutions are easily obtainable for special geometries like a sphere, a plane or an infinitely long charged rod. These work nicely as long as excluded volume effects or charge correlations do not play a dominant role. However, it is not easy to implement a numerical solver of the PB equation for solution properties of many macroions. The Debye-Hückel approximation is the linearization of the PB equation, and makes even cruder assumptions, but is often a good starting point for scaling approaches.

Computational simulations provide some unique ways to elucidate the properties of charged systems. We first give a short introduction to the relevant simulation methods, and focus then on some recently obtained results. A more technical review of our latest results can be found, for example, in\textsuperscript{2,4-6}.

2 Simulational Model

The first step in developing a model which is suitable for computer simulations is to coarse-grain the chemical details. Our mapping from a chemical substance like sulfonated
polystyrene to a charged bead-spring model is depicted in Fig. 2. One relevant parameter is the charge density along the backbone, which is typically measured by the Manning parameter \( \xi = l_b/b \), where \( b \) is the distance between charges along the polymer backbone, 
\[
l_b = \frac{e^2}{4\pi\epsilon_0\epsilon_r k_B T} \]
the Bjerrum length, which measures the strength of the electrostatic interactions, \( e \) is the elementary charge, \( \epsilon_0 \) and \( \epsilon_r \) denote the vacuum dielectric, and relative dielectric constant, \( k_B \) the Boltzmann constant, and \( T \) the temperature. The size of the ions is denoted by \( \sigma \) and typically of the order of 4 r\( \AA \), which is the ion plus its first hydration shell. The simulation technique is a classical molecular simulation (MD), where one calculates the forces acting on each monomer, and then integrates Newton’s equations of motion for a discrete time step, and adds Brownian motion via a Langevin thermostat. Hydrodynamic interactions are thought to be negligible at our resolution and hence neglected. The model potentials basically consist of an excluded volume term modeled by a Lennard-Jones (LJ) interaction with variable cut-off. The bonding constraint within a polymer is modeled by a non-linear spring, a FENE potential, and all charges interact via a full Coulomb potential. The solvent quality can be modeled by the depth \( \epsilon \) of the short ranged LJ potential, when it is not cut-off in the minimum (for a pure repulsive interaction) but instead cut-off at \( 2.5\sigma \).

One of the biggest problems for the simulations of charged systems is the long range nature of the Coulomb interactions. In principle, each charge \( q_i \) interacts with all others, leading to a computational effort of \( O(N^2) \) already within the central simulation box. For many physical investigations one wants to simulate bulk properties and therefore introduces periodic boundary conditions to avoid surface effects. Then the charges also interact with all their images in the replicated simulation cells. The standard method to compute the energy \( E \) with the merely conditionally convergent Coulomb sum

\[
E = \frac{1}{2} \sum_{\vec{r} \in \mathbb{Z}^3} \sum_{ij} \frac{q_i q_j}{|\vec{r}_{ij} + \vec{n} L|},
\]
where the prime denotes that for \( \vec{n} = 0 \) the term \( i = j \) has to be omitted, is the traditional Ewald summation. The basic idea is to split the original sum via a simple transforma-
tion into two exponentially convergent parts, where the first one, $\phi_r$, is short ranged and evaluated in real space, the other one, $\phi_k$, is long ranged and can be analytically Fourier transformed and evaluated in Fourier space:

$$\frac{1}{r} = \frac{1 - f(r)}{r} + \frac{f(r)}{r} \simeq \phi_r(r_c, \alpha) + \phi_k(k_c, \alpha)$$  \hspace{2cm} (2)

Traditionally, one uses for $f$ the error function $\text{erf}(\alpha r) := \frac{2}{\sqrt{\pi}} - \frac{1}{2} \int_0^{\alpha r} e^{-t^2} dt$, though other choices are possible and sometimes more advantageous. For any choice of the Ewald parameter $\alpha$ and no truncation in the sums the formula yields the exact result. In practice one wants to cut-off the infinite sum at some finite values $r_c$ and $k_c$ to obtain $E$ to a user controlled accuracy, which is possible by using error estimates. The aforementioned procedure results in the well known Ewald formula for the energy of the box

$$E = E(r) + E(k) + E(s) + E(d)$$  \hspace{2cm} (3)

where the contributions from left to right are the real space, Fourier space, self, and dipole-correction energy terms.

The Ewald sum has complexity $O(N^{3/2})$ in its optimal implementation, and therefore is not suitable for the study of large systems ($N > O(1000)$). Implementing a fast Fourier transformation (FFT) for the Fourier part results in the so-called particle-mesh-Ewald formulations, which improve the efficiency to $O(N \log N)$, and which we also parallelized\(^6\) to work on the T3E architecture. We typically use between 500 and 5 000 charges, with chain length ranging from 100 - 400 monomers. The typical runtimes can range between one day and two weeks for a single data point. This is the main reason why one is forced to use supercomputers in order to be able to intensively study these systems.

To simulate the structure of water (or other dipolar solvents), one would also need to treat the dipolar interactions in a similar fashion. Also here the Ewald method can be used\(^10\), and error estimates exist\(^11\), however due to the enormously increase in computational effort we treated the aqueous solvent in the present study as a continuum, which enters only through its modified polarizability, giving us a relative dielectric constant of $\epsilon_r = 80$.

### 3 Poor Solvent Chains with Explicit Counterions

Many polyelectrolytes possess a carbon based backbone for which water is a poor solvent. Therefore, in aqueous solution, there is a competition between the solvent quality, the Coulombic repulsion, and the entropic degrees of freedom. The conformation in these systems can under certain conditions assume pearl-necklace like structures\(^12\). These also exist for strongly charged polyelectrolytes at finite densities in the presence of counterions\(^13,6\). Earlier simulations\(^13\) used 16 chains of length $N_m = 94$, with a charge fraction of $f = 1/3$, and monovalent counterions. The hydrophobic interaction strength was tuned by means of the Lennard-Jones parameter $\epsilon$. It was found that the polymer density $\rho$ can be used as a very simple parameter to separate different conformation regimes. At very high densities the electrostatic interaction is highly screened, so that the hydrophobic interaction wins, and the chains collapse to dense globules. If one slightly decreases the density, the chains can even contract further, because there are no more steric hinderences from the
other chains or counterions, and the screening is smaller. The collapsed globules, however, have still a net charge, and repel each other, so that this phase resembles a charged stabilized colloid or microgel phase. With decreasing density the electrostatic interaction will dominate over the hydrophobic one. The chains will tend to elongate, assuming pearl-necklace conformations, like in Fig. 3, as they have been predicted for weakly charged polyelectrolytes. The more the chain stretches, the smaller become the locally compact regions. Note that in contrast to the analytical theories, the pearls are stable, even though there are counterions localized near and/or inside the pearls.

Experimentally there are some hints for the existence of pearl-necklace chains seen in scattering experiments. One of the obstacles to observing them in scattering experiments could be related to the strong fluctuations of the pearl number. Even in equilibrium we have found coexistence of several pearl states. For example, looking at the time evolution of a five chain system composed of 382 monomers with a charge fraction $f = 1/3$, $l_B = 1.5$, and $\epsilon = 1.75$ at density $\rho = 1.48 \times 10^{-5}$, we observe many jumps between five and four pearl configurations even for a single chain. Also the position of the pearls move quite vividly.

To see what an experimentalist could observe in a scattering experiment we computed for this system the spherically averaged form factor $S_1(q)$ of the chain, shown in Fig. 4.

The maximum seen at $q = 6$ comes from the monomer extension. In the range $1 < q < 2$ we observe a sharp decrease in $S_1$, which comes from the scattering from the pearls, because it shows the typical Porod scattering of $S_1(q) \approx q^{-4}$. The kink at
Figure 4. Spherically averaged form-factor $S_1(q)$. Shown are the single chain form-factor (red line), together with the part of the form factor coming from the intra-pearl scattering (green line). The blue and pink lines show the elongated chain part, and the Porod scattering part (globular conformation).

$q \approx 1.66$ 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 $q \approx 0.5$ 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 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 only few pearl numbers, which could lead to stable and large signatures.

Next we performed simulations on stretching a single chain consisting of $N_m = 256$ charged monomers in the infinite dilution limit; there are no counterions present. This is nowadays a standard technique performed with Atomic-Force Microscopes (AFM) on single molecules. The LJ parameter was set to $\epsilon = 1.75$, and the Bjerrum length was set at $l_B = 0.08$, the chain is only weakly charged and close to the ones considered in the scaling approach\textsuperscript{16,17}. Technically we fix a certain end-to-end distance by using an equilibrated chain and fixing the end-monomer separation. Subsequently we enlarge the separation along the end-to-end vector. In the unstretched state the system is in a two-pearl configuration. The initial equilibrium end-to-end distance was $R_E = 21\sigma$, and the system was elongated up to $R_E = 88\sigma$. The first remarkable observation is, that the two-pearl state evolves into a three-pearl state under extension, which was predicted by scaling\textsuperscript{16}, but is counter intuitive from simple arguments. It is a result of the electrostatic inter-pearl interactions\textsuperscript{6}. At each separation the chain was equilibrated and sampled over a long time in equilibrium, and the resulting force on the end-monomers was measured, yielding a
Figure 5. Force-extension relation for a transition from two to three pearls. Shown are the average over all conformation (solid black), over only two-pearl configurations (red), and only three-pearl configurations (green).

Figure 6. Color coded distribution of the hydrodynamic radius \( R_H \) for different end-to-end distances \( R-E \). The color scheme runs from black for \( P = 0 \) over red, yellow, green, blue, and magenta for large values of \( P \). On the right hand side we show for three different values of \( R_e \) the probability distributions of \( R_H \). This corresponds to horizontal cuts through the histogram on the left.

Continuous curve, see Fig. 5. Only when averaged separately over the three-pearl and two-pearl states, we recognize a saw-tooth pattern in the force, as predicted by scaling theories. In equilibrium one would expect a plateau for a true first order transition, and the upper and lower part would correspond to metastable superheated and supercooled states. Again, due to the large chain fluctuation and coexistence region, all nontrivial signatures in the force-extension relation are basically washed out in equilibrium. To see hysteresis effects
in an AFM experiment one needs to pull at faster rates than the equilibration times of the chain.

In Fig. 6 we have plotted the distribution of the hydrodynamic radius $R_H$ for various end-to-end distances $R_E$. In the middle right picture we observe a clear bimodal distribution, showing that the transition between different pearl-states is of first order. This can be particularly well seen in $R_H$, because it is the thermodynamically conjugated variable to the pearl-number$^{18}$. Many more interesting results on poor solvent polyelectrolytes can be found in Ref.$^6,19,20$ and will be published soon.

## 4 Rodlike Polyelectrolytes

Stiff linear polyelectrolytes can be approximated by charged cylinders. This is a relevant special case, applying to quite a few biologically important polyelectrolytes with a large persistence length, like DNA, actin filaments or microtubules. Within PB theory and on the level of a cell model$^{21}$ the cylindrical geometry can be treated exactly in the salt-free case, providing for instance new insights into the phenomenon of the Manning condensation$^{22}$. For small Manning parameters, the agreement between PB theory and the simulations of the full interacting system is rather nice$^{4,5,22}$. Recently the osmotic coefficient of a synthetic stiff polyelectrolyte, a poly(para-phenylene), was measured in a salt-free environment$^{23}$. We have compared this data to predictions of PB theory, and a local density functional theory which includes a correlation correction of the basis of a recently proposed Debye-Hückel-Hole-cavity theory (DHHC)$^{24}$, and simulational results within the cell model. We find that correlation effects enhance condensation and lower the osmotic pressure, yet are not fully able to explain the discrepancy with the experimental data. Here the approach of working within the "primitive model" breaks down. In our opinion, specific interactions between the counterions, the macroion, and the solvent particles are needed to explain the discrepancy. Other theoretical approaches beyond the cell model which try to incorporate finite-size effects and interactions of the macroion itself will in general lead to a higher osmotic coefficient which is in contrast to the experimental data$^{25}$.

However, PB theory fails qualitatively (overcharging, charge oscillations and attractive interactions) for strongly charged systems$^{4,5,22}$. As an example we present in Fig. 7 measurements on the osmotic coefficient $\hat{p}$, which is the system pressure normalized by its ideal gas contribution. The measurements were performed on a forest of hexagonally aligned infinite rods at various concentrations, which is equivalent to varying the distance between neighboring rods. The rods have all a diameter of one $\sigma$, and have a charge parameter of $\xi = 0.9593$, hence are at the Manning limit for monovalent counterions. We see in Fig. 7 that the deviation to the PB predictions increase with increasing valence and with increasing concentration. For divalent counterions, one sees already a dip in the osmotic coefficient around densities $0.1 \sigma - 0.01 \sigma^{-3}$, which however develops into a negative valley for trivalent counterions. Because our rods are constraint to a lattice, a negative pressure simply means that the system wants to phase separate into a denser phase, hence we see attractions at these densities. There are nice rigorous results which prove that the attractions cannot be described within PB theory because of the neglect of important ionic correlations. The simulations have the clear advantage that all correlations are taken fully into account. Our findings are in agreement with experimental observations of attractions.
Figure 7. Osmotic coefficient \( \hat{p} \) for three equal rod systems as a function of concentration for three types of counterion valence. Heavy dots mark the measurements, while the solid lines are fits which merely serve to guide the eye. The dotted lines are the prediction of Poisson-Boltzmann theory. From top to bottom the valence varies like 1, 2, 3. The error in the measurement is roughly as big as the dots itself.

between like-charged macromolecules. Especially in the community of biological inspired physics\(^2\), these interactions are thought to be important for the clarification of the mechanism behind DNA compactification in viral heads, the chromatin structure, and novel methods for gene delivery, to name just the most prominent examples. There are numerous simulations which show similar attractions on a distance of few counterion diameters, which can be looked up in Ref.\(^4\).

The mechanism of the observed attractions for rod-like systems has been speculated to be correlations between the counterion layers around the macroion. However, until now, no unique theoretical picture has emerged that can clarify the detailed mechanism, i.e. which are the important correlations responsible for the attractions. There is the low temperature Wigner crystal theory, initiated by Refs.\(^{26}\), which postulates an ordered ground state of the counterions. Then there are theories which are based on Van-der-Waals type correlated fluctuations\(^{27}\), that are in principle high T theories. There are also theories which are fluctuation based, but are valid at low T\(^{28}\). Integral equation\(^{29}\) theories on various approximation levels have been demonstrating the existence of these attractions for a long time, but from these theories it is difficult to extract the detailed mechanism behind the observed correlations. Here also simulations can be helpful, because they have in principal access to all correlations, In Fig. 8 we see two rods separated by 2\(\sigma\) at two different interaction strength \((l_b = 0.5\sigma\) and \(l_b = 3\sigma\)) for trivalent counterions. The left pair of rods shows no attractive interactions, and we see only weak developed order in the counterion layer. The pair on the right side shows attraction, and we can clearly observe strong correlations in the layer. We have performed a detailed study on different ion correlations (entropic

393
Figure 8. Simulation snapshot of two DNA sized rods separated by a surface to surface distance of $2\sigma$. Picture a) shows a used Bjerrum length of $l_b = 0.5\sigma$ and in b) we used $l_b = 3\sigma$.

versus energetic contributions) trying to quantify when attraction is likely to start, and found out, that for most experimentally relevant cases the entropic part of the ion correlation is dominant. These will be presented elsewhere$^{30}$. Even a simple two macroion pair shows very complicated patterns in their counterion distribution which is theoretically by no means settled. Many more “simple” questions await their answers, and some surprises lie probably ahead of us. More details of our results in rod-like geometries can be found in Refs.$^{4,5,22,24,25,31,32}$.

Acknowledgments

We gratefully acknowledge partial funding by the DFG Schwerpunkt “Polyelektrolyte”, as well as a generous computer time donation by the NIC under grant No. hkf06.

References

27. F. Oosawa, Biopolymers 6, 1633 (1968); B. Jónsson, and H. Wennerström, in Ref. 2.
30. M. Deserno and C. Holm, to be published.
We use molecular dynamics simulations to study polyelectrolytes end-grafted to a surface. The simulation model includes counterions explicitly and treats the full Coulomb interaction. As a main result, we find a novel collapsed brush phase. Its nature can be understood by an extended scaling theory. In the collapsed phase, the height of the brush results from a competition between steric repulsion and attraction due to electrostatic correlations. As a result, the monomer density inside the brush becomes independent of grafting density and chain length.

1 Introduction

Polyelectrolytes, i.e., polymers being equipped with ionizable groups, have received a great deal of attention in recent years. In contrast to uncharged polymers, polyelectrolytes are soluble in water. This particular feature makes them environmentally friendly and in many cases the corresponding applications are cost effective. Anchoring polymer chains by one end sufficiently dense to a solid/liquid or air/liquid interface they form a so-called polymer brush, in which the chains are enforced to take a stretched conformation. A schematic representation of a uncharged polymer brush is shown in Fig. 1 (left-hand part). Polyelectrolyte brushes are interesting from the point of view of fundamental as well as of applied research. Already for some decades polymer brushes are known as an efficient means for preventing colloids from flocculation. Brushes made of charged chains have the particular advantage that the stabilization arises both from steric and electrostatic effects. In addition, a surface coated with polyelectrolytes is expected to be less sensitive

![Figure 1. Schematic picture of polymer brushes. Left-hand side: made of uncharged polymers, middle: polyelectrolyte brush at strong charging, left-hand side: polyelectrolyte brush at weak charging.](image-url)
to the salinity of the surrounding medium than a bare charged surface because a strongly charged brush is able to trap its own counterions and thus generates a layer of large effective ion strength (see Fig. 1, middle part). However, in the weak-charging limit (see Fig. 1, right-hand side) counterions become free to leave the brush and the properties of the polyelectrolyte brush can be drastically changed\(^2\).

Both in experiment and in theoretical work, polyelectrolyte brushes are an interesting subject with many unresolved problems. In this situation, simulations are a promising tool to validate theoretical models and to probe quantities and regimes which are not easily observable experimentally. However, simulations of polyelectrolytes itself remain still challenging despite a strong effort in recent years. The correct treatment of the long-range Coulomb interaction in a periodic system requires special methods which are computational expensive. This task is even more demanding in a slab geometry with spatial periodicity in two directions only which is the case for anchored chain. Under such boundary conditions, a straightforward application of most of the methods known for treating long-range interactions in three-dimensional periodic systems, as e.g. the so-called Ewald sum, is not possible.

In this contribution we report the results of extensive simulation studies on fully charged and on partially charged polyelectrolyte brushes. The paper is organized as follows. In Sec. 2 we shortly review the polyelectrolyte model and the simulation technique. The simulation results are discussed in Sec. 3. In Sec. 4 we compare simulation results with predictions of an extended scaling theory. Conclusions are given in Sec. 5.

### 2 Model and Method

We use a bead-spring model where the polyelectrolytes are composed of charged monomers (for partially charged polyelectrolytes, also uncharged species) which are connected by nonlinear springs and end-grafted to an uncharged surface. The counterions are

![Figure 2. Schematic picture of the simulation model.](image-url)
free charged particles and there is no additional salt in the system. The solvent (water) is
replaced by a dielectric background and a heat bath. A schematic picture of the simulation
model is shown in Fig. 2. The total potential of the system has four contributions:

\[ U = U_{\text{bond}} + U_{\text{LJ}} + U_{\text{wall}} + U_{\text{Coul}}. \]  

Besides the bond potential \( U_{\text{bond}} \) we have a purely repulsive Lennard-Jones term \( U_{\text{LJ}} \)
which models the situation of a polymer in a good solvent. All particles except the fixed
anchors interact repulsively with the surface. For this purpose, we use a short-ranged wall
potential \( U_{\text{wall}} \), which prevents the particles from crossing the surface. The last term in
Eq. (1) stands for the Coulomb interaction between all charged particles

\[ U_{\text{Coul}} = \lambda_B k_B T \sum_{n_x=0}^{\infty} \sum_{n_y=0}^{\infty} \sum_{i=1}^{N_{\text{tot}}-1} \sum_{j=i+1}^{N_{\text{tot}}} \frac{q_i q_j}{|\mathbf{r}_{ij} + n_x L \mathbf{e}_x + n_y L \mathbf{e}_y|}, \]  

where \( \lambda_B = e^2/(4\pi\varepsilon_0\varepsilon k_B T) \) is the Bjerrum length with \( \varepsilon_0 \) and \( \varepsilon \) being the vacuum
permittivity and the dielectric constant of the solvent, respectively. \( q_i \) is the charge of particle
\( i \) in units of the elementary charge \( e \). \( \mathbf{e}_x \) and \( \mathbf{e}_y \) are unit vectors in \( x \)- and \( y \)-direction,
respectively, and the indices \( n_x \) and \( n_y \) run over the periodic images of the simulation box.
\( N_{\text{tot}} \) is the total number of charges, and \( L \) is the planar box length. A major task in any
simulation of charged systems is to transform conditionally convergent sums like that in
Eq. (2) into rapidly convergent series. We use an approach proposed by Lekner and Sperb\(^3\),
which allows the transformation of Eq. (2) into an expansion in terms of Bessel functions
or Hurwitz zeta functions. Due to the asymptotic behavior of these functions one obtains
alternative expressions which are rapidly convergent either at large or small arguments.

We use stochastic molecular dynamics\(^4,5\) to study the system in equilibrium. The equation
of motion for particle \( i \) at position \( \mathbf{r}_i(t) \) is the Langevin equation,

\[ m \frac{d^2 \mathbf{r}_i}{dt^2} = -\nabla_i U - m \Gamma \frac{d\mathbf{r}_i}{dt} + \mathbf{W}_i(t), \]  

where all particles carry the same mass \( m \), and \( \Gamma \) is a friction constant which couples the
particles to a heat bath. \( U \) is the potential energy defined in Eq. (1). The system is held at
thermal equilibrium by a Gaussian random force \( \mathbf{W}_i(t) \),

\[ \langle \mathbf{W}_i(t) \rangle = 0, \quad \langle \mathbf{W}_i(t) \cdot \mathbf{W}_j(t') \rangle = 6mk_B T \Gamma \delta_{ij} \delta(t-t'), \]  

where the coupling to \( \Gamma \) is a consequence of the fluctuation-dissipation relation. Eq. 3 is
integrated using the velocity-Verlet algorithm\(^5\). The molecular dynamics code was parallelized
using a self-scheduling (master-slave) algorithm\(^6\) for the force loop. Upon testing
the parallel code for different processor numbers \( n \) on a Cray T3E, we found that for our
strongly charged systems the efficiency is optimal in the range of \( n = 32 \) to \( n = 64 \) pro-
cessors, where an efficiency of over \( 90\% \) is reached. Lower numbers of processors lead to
a loss in efficiency due to the idle master processor, while large processor numbers are less
efficient because communication time becomes a dominant contribution.

### 3 Simulation Results and Discussion

The molecular dynamics code used in this study allows the investigation of polyelectrolyte
brushes in a large parameter range. The contour lengths of the coarse-grained chains mod-
elled in the simulation correspond to about 20 nm - 50 nm which is comparable to experiments on polyelectrolyte brushes at the air/water interface. The grafting densities studied correspond to a surface area per chain between 600 $\text{rA}^2$ and 3000 $\text{rA}^2$, values being typical for experimental systems.

3.1 Completely Charged Chains

In Fig. 3 the dependence of the average brush height on anchoring density is shown for two different Bjerrum lengths ($\lambda_B$ equal to 0.7 $\sigma$ and 0.1 $\sigma$, with $\sigma$ being the Lennard-Jones diameter). Note that the average bond length $b$ is about 1.0 $\sigma$. Thus, for both Bjerrum lengths, the Manning parameter holds $\xi = \lambda_B / b < 1$. Thus, one would not expect strong counterion-condensation effects, at least for a fully stretched chain at infinite dilution. However, as we will see below, the situation can be quite different in the case of a strongly charged brush.

![Figure 3. Average brush height versus anchoring density for different interaction strength.](image)

For a moderate Bjerrum length $\lambda_B = 0.7\sigma$ (hereafter called strong coupling) we find a linear scaling law $<z_m> \sim \rho_a^{\alpha}$, which is consistent with a constant particle density inside the brush. Our observation of such a collapsed phase disagrees with the accepted scaling law for a charged brush in the so-called osmotic regime which predicts a brush height, independent of grafting density. We believe that this disagreement is caused by the assumption that in the osmotic regime the counterions inside the brush form an ideal gas. Indeed, analyzing the counterion-polyelectrolyte distribution function shown in Fig. 4, it is evident, that strong correlations are present. Note that the scaling law, we observe, is the same as that known for uncharged brushes in a poor solvent. On the other hand, recently it has been shown that counterion correlation (condensed counterions) gives rise to an effective interaction which can change the solvent quality from good to poor. Thus,
strongly charged polyelectrolyte brushes can be considered as a new example of systems where counterion-induced attractions can play an important role. This is a more general topic which received recently a lot of attention in theory and simulations. A more precise understanding of the nature of the collapsed phase we will discuss in Sec. 4.

Figure 4. Counterion-polyelectrolyte distribution function at $\lambda_B = 0.7\sigma$ and $\sigma_{ci} = \sigma_{m} = \sigma$. $r$ is the separation between polyelectrolyte bond and counterion.

For weak Coulomb coupling, i.e. at $\lambda_B = 0.1\sigma$, there occurs a quite different behavior. The collapsed regime disappears and we find a weak dependence of the brush height on anchoring density. In this case the counterion correlation is less pronounced as one can realize intuitively from a simulation snapshot shown in Fig. 5. The weak dependence of brush height on anchoring density is possibly a result of the repulsive Lennard-Jones interaction which is present in our finite simulation system. On the other hand, the osmotic regime is solely governed by the balance between the osmotic pressure of the counterions and the restoring force caused by the stretching energy of the chains.

Note that the third case shown in Fig. 3, i.e. small counterions, accidentally exhibits the same scaling law as the weak-coupling case. Nonetheless, the physical nature of these two regimes is quite different.

### 3.2 Partially Charged Chains

Simulations on partially charged polyelectrolyte brushes show quite different behavior for decreasing degree of charging $f$. In Fig. 6 brush heights versus anchoring densities are shown for the total range of charging $0 \leq f \leq 1$. On the one hand, at $f = 0$, we have the well known limit of a neutral brush $<z_m> \sim \rho_m^{1/3}$, on the other hand, at $f = 1$, we have the limiting scaling laws discussed above. At weak Coulomb coupling we find at any grafting density a monotonous stretching of the brush with increasing degree of charging. At strong interaction (left-hand side in Fig. 6) similar stretching is only found at dense
grafting. For loose grafting densities, we observe a pronounced shrinking of the brush with increasing $f$. In the following section we will discuss how this unusual behavior can be understood on the basis of an extended scaling theory.

4 Comparison to Extended Scaling Theory

Recently a number of theoretical and numerical results have stressed the possibility of polyelectrolyte chain collapse at high charge densities. Let us use the simplest description which is able to produce these effects in a qualitative manner. In the bulk and in the low-density limit, the Debye-Hückel free energy density per unit volume is given by

$$f_{DH}/k_B T = -\frac{\kappa^3_D}{12\pi}$$

with the Debye length $\kappa^{-1}_D$ for salt solutions of concentration $c_s$ defined by $\kappa^2_D = 8\pi\lambda_B c_s$. For low enough ion densities the Debye-Hückel approximation is a well-controlled approximation. More sophisticated approaches for treating the Coulomb correlations in simple electrolytes and polyelectrolytes (including the effect of the chain connectivity) lead to the same leading term given in Eq. (5).

Including the Debye-Hückel contribution $F_{DH}$ the total free energy for a polyelectrolyte
Figure 6. Average brush height (rescaled with the chains length $N$) versus anchoring density at different degree of charging, varying between $f = 1$ (red diamonds) and $f = 0$ (black stars). The limiting scaling laws are given as thin solid lines. Left-hand side: $\lambda_B = 0.7 \sigma$, right-hand side: $\lambda_B = 0.1 \sigma$. Because in the strong coupling case (left-hand side) the brush heights remain almost unchanged at small $f$, the smallest degree of charging plotted in this part of the figure is $f = 0.08$ (green circles).

brush in a good solvent becomes

$$F = F_{st} + F_{v2} + F_{os} + F_{electrost} + F_{DH},$$

where $F_{st}$ is the stretching energy of the chains and $F_{v2}$ the second virial (excluded volume) contribution. $F_{os}$ stands for the osmotic free energy associated with the entropy costs for confining the counterions inside the brush and $F_{electrost}$ is the direct electrostatic contribution occurring if counterions leave the brush which results in a violation of the local electro-neutrality.

Analyzing the total free energy Eq. (6), in addition to the previously known polyelectrolyte brush regimes we found that the Debye-Hückel correlation term induces a new collapsed brush regime (CB). Balancing the attractive interaction by the second virial the corresponding brush height becomes

$$h_{CB} \sim N \rho_a v_2^{-1 \lambda_B^{-3}}.$$  

Note first that the scaling $h_{CB} \sim N \rho_a$, which results in a constant particle density inside the brush, is in agreement with the simulation results discussed above. Secondly we emphasize that Eq. (7) contains an additional scaling $h_{CB} \sim f^{-3}$. Such a shrinking of the brush with increasing $f$ is in agreement with the behavior we have discussed above for strong coupling and loose grafting density. Thirdly, it should be noted that the second virial coefficient $v_2$ is an effective parameter. There is also an electrostatic contribution to $v_2$ which is repulsive for large ionic radii, but becomes negative for small ions.

The phase diagram in the $(f, \rho_a)$-space following from the total free energy Eq. (6) is shown in Fig. 7. Clearly, in the weak coupling limit, i.e. at $\lambda_B < v_2^{1/3}$, we get the diagram of states previously obtained by Borisov et al. On the other hand, at strong coupling and strong charging, i.e. at $(v_2/\lambda_B^{3})^{1/2} < f < 1$ the brush can undergo a first-order phase
transition from the osmotic into the collapsed state. Thus, the disappearance of the collapsed regime at small Bjerrum lengths, obtained in the simulations, is in agreement with the predictions of the theory. A qualitative understanding of the behavior at $\lambda_B = 0.7\sigma$ shown in the left-hand side of Fig. 6 can be obtained if one assumes a crossing of the CB - NB transition line in the upper part of the phase diagram Fig. 7 while reducing the degree of charging. Such an assumptions is reasonable, however, there is no way to succeed with a more quantitative picture. For the particular choice of simulation parameters, we have $\lambda_B^3/\nu_2 \approx 1$ which corresponds to the cross-over between strong and weak electrostatic coupling. In this case, the scaling theory allows no quantitative comparison with simulation results. However, it correctly describes the trends, namely that with increasing degree of charging $f$ one goes to the collapsed brush regime.

5 Conclusion

In our simulation studies on polyelectrolyte brushes we found a novel collapsed brush regime not understood by theories known from literature. This fact stimulated the generalization of the theoretical picture. Now, the nature of the collapsed phase can be understood on the basis of a simple scaling model: Including in the total free energy a Debye-Hückel correlation term, indeed a new collapsed regime is obtained where the attractive Debye-Hückel interaction is balanced by the second virial of monomers and counterions.
The accessible region of the diagram of states is very sensitively dependent on the choice of parameters, in particular on the relation between Bjerrum length $\lambda_B$ and second virial coefficient $v_2$. At this point we remind that in experiments on polyelectrolyte brushes no collapsed behavior has been reported so far, but the osmotic regime seems dominating. On the other hand, up to now no osmotic behavior has been seen in simulations, but brush collapse appears to be a major feature if Bjerrum length and second virial coefficient are straightforwardly chosen as ‘bare’ parameters. For obvious reasons this observation suggests that with such a choice of the simulation parameters we are studying the strong coupling case. Indeed, reducing the Bjerrum length at fixed second virial coefficient, considerable stretching of the polyelectrolyte brush together with a weak dependence on grafting density is obtained. Thus, to bridge the gap between simulation and experiment more insight is necessary in order to be able to chose appropriate and well-founded parameters.

**Acknowledgements**

The work was supported by the Deutsche Forschungsgemeinschaft within the Schwerpunktprogramm Polyelektrolyte. We gratefully acknowledge grants of computer time provided by NIC Jülich and ZIB Berlin.

**References**

Dissipative Particle Dynamics Simulations of Planar Amphiphilic Bilayers

Julian C. Shillcock and Reinhard Lipowsky

Max-Planck-Institut für Kolloid-und Grenzflächenforschung
14424 Potsdam, Germany
E-mail: {shillcock, lipowsky}@mpikg.mpg.de

The equilibrium structure and lateral stress profile of bilayer membrane patches are investigated using the Dissipative Particle Dynamics simulation technique. Although there are no attractive forces between the model amphiphiles, they spontaneously aggregate into planar bilayers under suitable conditions of concentration and amphiphile architecture. Pure bilayers of single-chain amphiphiles are simulated, and the amphiphile architecture and interaction parameters varied. We find that a strong chain stiffness potential is essential to create the degree of lamellar order typical in natural lipid membranes. The lamellar order of the bilayers is destabilised by reductions in the tail stiffness. The lateral stress profile across the bilayers contains detailed structure reflecting contributions from all the interaction potentials, as well as the amphiphile architecture. The stress profile is similar to that found in coarse-grained Molecular Dynamics simulations, but is here established at a fraction of the computational cost. Dissipative Particle Dynamics therefore allows the study of equilibrium behaviour of amphiphilic membranes hundreds of times larger than can be achieved using Molecular Dynamics simulations, and opens the way to the study of complex mesoscopic cellular phenomena.

1 Introduction

Lipid bilayer membranes surround living cells, protecting their interior from the outside world. They are much more than a static structural component, however, in that their composition and dynamics influence membrane-bound proteins, and contribute to the remarkable material properties of cells such as red blood cells\(^1\). Bilayer membranes also surround artificial vesicles, and have been constructed out of non-biological amphiphiles\(^2,3\), and diblock copolymers\(^4\). These membranes continually undulate owing to the thermal motion of their constituent lipids. Thermal forces combine with specific molecular forces to create complex, dynamic, multi-component systems\(^5\). Dynamic processes taking place within a membrane can involve co-operative changes over distances large compared to the molecular size, and occur on time-scales much longer than molecular vibrational periods.

The complexity of natural membranes has led experimentalists to focus on simpler model systems: lipid bilayer vesicles\(^5,6\). These are often composed of a single type of amphiphile, and usually lack embedded inclusions. Much progress has been made in the last decade in experiments designed to probe the vesicle membrane’s material properties. The area stretch modulus of pure stearoyl oleoylphosphatidylcholine (SOPC) vesicles and SOPC/cholesterol mixtures has been measured using micropipette aspiration\(^7\), and is found to increase on adding cholesterol as a cosurfactant. The lateral diffusion of lipids in a vesicle membrane\(^8\) and the membrane viscosity\(^9\) have both been recently determined. Detailed studies\(^10\) have been published of the dependence of the bilayer elastic bending and area stretch moduli on lipid tail length and the degree of unsaturation. Amphiphilic diblock copolymers have been shown to form vesicles that are an order of magnitude stronger, and
less permeable to water, than natural phospholipid bilayers. Such toughened vesicles, or polymersomes, offer greater control over the membrane material properties than lipid vesicles owing to the possibility of cross-linking the copolymers and changing the block size or molecular weight.

We have used the Dissipative Particle Dynamics (DPD) simulation method to investigate the structure and lateral stress profile of bilayer membranes containing approximately 3200 amphiphiles as a function of the amphiphile architecture and interaction parameters. This represents a membrane patch at least one order of magnitude larger than previously published results, and allows us to study the membrane’s mesoscopic properties while the (presumably) irrelevant short length-scale motions of the individual amphiphiles are averaged out. The DPD technique was first introduced almost a decade ago, and recent articles provide a comprehensive survey of the method.

The paper is organised as follows. We first provide the motivation for our work and then briefly review the Dissipative Particle Dynamics simulation technique, referring the interested reader to Ref. 17 for a detailed description. Then we present results on the equilibrium bilayer structure and stress profile of amphiphilic bilayers as a function of the amphiphile architecture and interaction parameters. Finally, we discuss the implications of this work for simulating complex processes in lipid bilayers.

2 Motivation

Experiments on complex biomembranes have created a demand for a theoretical understanding of the dependence of membrane material properties on the constituent amphiphile’s molecular structure and membrane composition. Because mean-field and lattice-based models ignore membrane undulations, restrict amphiphile headgroups to a planar interface, and cannot easily incorporate the effects of arbitrary molecular architecture, attention has turned to predicting material properties using coarse-grained Molecular Dynamics (MD) simulations. This approach has been used to extract the area compression modulus and bending modulus of single-component lipid bilayers, and their lateral stress distribution. The latter quantity is believed to be important in modulating membrane-bound protein behaviour. Recently, coarse-grained MD simulations have been used to compare the equilibrium structure of a dimyristoylphosphatidycholine bilayer to that obtained from atomistic simulations. This provides the opportunity to move up in length-scale towards the mesoscopic regime. However, a major drawback of these MD simulations is that they are restricted by current computing technology to membrane patches containing only a few hundred amphiphiles plus the requisite solvent molecules. Mesoscopic simulation techniques, such as DPD, offer the possibility of extracting information about the material properties of biomembranes, and other complex fluid systems, well beyond the length and time scales achievable by MD simulations. Our aim in this work is to determine whether DPD simulations can take the investigation of membrane material properties to length and time scales beyond those achievable in coarse-grained MD simulations, whilst still exhibiting the structure found in lateral stress profiles. Given that a micron-size vesicle can contain from one million to a billion amphiphiles, and that a few percent of its surface area is involved in processes such as pore formation or fusion events, the ability to model large systems is essential if these processes are to be studied using computer simulations.
3 Dissipative Particle Dynamics Simulation Method

The elementary units in a DPD simulation are fluid elements or soft beads. A bead represents a volume of fluid that is large on a molecular scale, and hence contains at least several molecules of the fluid, but still macroscopically small. Beads interact via effective forces chosen so as to reproduce the hydrodynamic behaviour of the fluid without reference to its molecular structure. DPD differs in this respect from MD simulations, in which the forces are chosen to model the inter-molecular interactions of a system as accurately as possible. Forces in DPD are pairwise additive, conserve momentum, have no hard core and are short-ranged, the range of the force defining the size of the soft beads. All beads have the same mass, \( m_0 \), and radius, \( r_0 \), unless otherwise stated, and these set the mass and length scales in the simulation. A time-scale must be extracted from the dynamics of relevant processes in the simulated fluid, such as the diffusion of a micelle’s centre of mass, or the in-plane viscosity of a bilayer membrane. Because we study equilibrium properties of the bilayers, we estimate the time-scale of the simulations from the generic time,

\[
t_0 = \sqrt{\frac{m_0 r_0^2}{k_B T}},
\]

set by the bead mass and radius and the system temperature, where \( k_B \) is Boltzmann’s constant and the temperature \( T \) is defined in Ref. 17. We take the diameter of one DPD bead as 1 nm and assume that it has the density of water at room temperature, \( T = 300 K \). The simulation time-step then corresponds to 5 ps, and a typical run of \( 10^5 \) steps is equivalent to approximately a microsecond of real time.

Simulations take place within a cuboidal box of constant volume \( V = L_X L_Y L_Z \), where \( L_X, L_Y, L_Z \) are the simulation box side lengths in units of the bead diameter, \( r_0 \). Periodic boundary conditions are used in all three dimensions to minimise edge effects. The simulation box is filled with beads to the chosen density which represents \( \rho \) beads/unit volume. We are interested in studying the properties of a single bilayer in water. The number of amphiphiles in the bilayer, \( N \), is determined by the box area, \( L_X L_Y \), and the desired projected area per amphiphile, \( A_{pr}/N \) according to

\[
N = \frac{2L_X L_Y}{A_{pr}/N}.
\]

Because the bead radius, \( r_0 \), defines the length scale for the simulations, we quote dimensional quantities in their dimensionless form, e.g., the area per amphiphile is \( A/Nr_0^2 \). In a similar manner, the mass and time scales are obtained from the bead mass and radius, and the system temperature.

Various polymer architectures are used to represent bilayer-forming amphiphiles. They are composed of hydrophilic head beads, designated \( H \), and hydrophobic tail beads, designated \( C \). The simplest architecture has a single \( H \) bead attached to a linear chain of \( C \) beads. The number of tail beads is varied to investigate the dependence of bilayer properties on the degree of amphiphile hydrophobicity. An amphiphile containing one head and \( n \) tail beads is represented, using an obvious symbolism, as \( HC_n \). The amphiphiles are contained within bulk solvent composed of \( W \) beads. Each solvent bead represents a small volume of bulk water consisting of several molecules. Because each \( W \) bead represents several molecules of solvent, there is no explicit modelling of hydrogen bonds or entropic forces. The interactions between beads in DPD simulations are to be interpreted as a coarse-graining of a fluid rather than a simulation of the molecules of a fluid. In this way, only structure and behaviour that occurs on a length-scale larger than the elementary beads in the simulation have physical relevance.

We refer to the polymers as surfactants or amphiphiles rather than lipids both to emphasise the generality of the simulation technique, and to avoid suggesting that the model...
amphiphiles should be viewed as atomically-detailed representations of the complex structure of lipid molecules. In the same vein, we refer to DPD beads connected by springs as polymers rather than molecules, but this is not meant to imply that the polymers are necessarily long nor that they are composed of identical repeating units. The interpretation of a polymer composed of DPD beads requires some care. We take the view that the head bead in a model lipid represents the hydrophilic glycerol-phosphate-head region while each tail bead represents several methyl groups in a hydrocarbon chain. In this view, each hydrophobic bead represents, say, 2 to 5 methyl groups. The same interpretation applies to non-biological amphiphiles, such as alkyl phosphate surfactants that consist of an single carbon chain attached to a phosphate head group.

4 Results on Equilibrium Structure of Bilayer Membranes

The simplest amphiphile architecture that is found to self-assemble into a bilayer consists of a single, hydrophilic head bead attached to a linear chain of hydrophobic tail beads designated $HC_n$. A snapshot of a bilayer composed of $HC_0$ amphiphiles is shown in Figure 1. The simulation box size is $V/r_0^3 = 32^3$, and the overall bead density is $ρr_0^3 = 3$ giving approximately 100,000 beads of all types. A bilayer readily self-assembles when the amphiphiles are initially randomly distributed in water for surfactant number fractions in the fairly restricted range of $3 - 6\%$. Below this range micelles occur, and above it complicated three-dimensional structures are formed (data not shown). The amphiphiles aggregate into a bilayer because of the strong repulsion between their tail beads and the solvent beads, mimicking the hydrophobic force of, for example, lipid molecules in water. Because of the large parameter space of the simulations, we investigate the effects of just a few parameter combinations. Given three bead types (H, C and W) there are 6 independent bead-bead interaction parameters. Their values are: $a_{HH} = a_{CC} = a_{WW} = 25$, $a_{HW} = 35$, $a_{HC} = 50$, $a_{CW} = 75$. The self-interactions are determined by the requirement that a pure fluid of each bead type has the compressibility of water, while the unlike bead-bead interactions are varied to represent the degree of solubility of one species in another, or the hydrophobicity of the hydrocarbon chain of the amphiphiles. The beads are connected into semi-flexible polymers using a bond potential and a chain bending potential that contribute two more parameters each. We reduce these by fixing the spring constant and unstretched bond length to $k_B r_0^2/kT = 128$ and $l_0 = 0.5 r_0$, respectively, and the preferred bond angle to $0^\circ$. This choice of values is explained later. This leaves 4 parameters to be investigated: $a_{HC}, a_{HW}, a_{CW}$ and $k_3/k_B T$ in addition to the number density of amphiphiles which is set by the projected membrane area, $A_{pr}/N r_0^2$ and the simulation box size. We have not investigated the effects of varying the dissipation coefficients because they should be irrelevant for the equilibrium structures we are interested in. Instead, they are chosen according to the procedure in Ref. 17 to yield a well-ordered bilayer. Because of the isotropy of the simulation box, bilayers do not always assemble with the same orientation, but to simplify the discussion we refer to the bilayer normal as the Z axis in all our results. A small bilayer containing 830 amphiphiles in an area $16^2$ readily self-assembles from an initially random configuration in 20000 time-steps (data not shown). Because self-assembly of bilayers has been studied previously and we are interested in measuring equilibrium properties of large membranes, we pre-assemble the amphiphiles into a planar bilayer, and allow it to relax to an equilibrium state before constructing ensemble averages of observables.
Figure 1. Snapshots of a bilayer containing 3321 $HT_6$ amphiphiles. The simulation box has size $(32r_0)^3$, and the bead density is $\rho r_0^3 = 3$ giving almost 100,000 beads of all types. The area-per-amphiphile is $A_{pr}/N r_0^2 = 0.62 \pm 0.05$ leading to a surface tension of $\sigma = -0.03 \pm 0.1$. The amphiphile head beads are dark grey, and the tail beads are light grey. The average membrane width, measured from the centre of the head beads of one monolayer to those of the other, is $<\ell_{me}/r_0> = 6.61 \pm 0.01$. Water beads are invisible for clarity. In the lower snapshot the terminal tail bead is coloured darker to demonstrate that the amphiphiles terminate near the bilayer midplane. The $L_\alpha$ order of the bilayer is evident, and the two monolayers are not very inter-digitated. Thermally-excited undulations are small here because of the chain bending stiffness of the amphiphiles.

The bilayer shown in Figure 1 contains 3321 $HC_6$ amphiphiles and has a relaxed area per amphiphile of $<A/N r_0^2> = 0.63 \pm 0.01$, and an approximately zero surface tension, $\sigma r_0^2/k_B T = -0.03 \pm 0.1$ (see next section for the measurement of surface tension). The snapshots in Fig. 1a and 1b come from separate simulations with identical parameters, but the terminal tail bead is coloured differently from the intermediate tail beads in 1b to show that the amphiphiles are approximately straight and terminate near the bilayer midplane. We measure the bilayer width and area from a triangulation of the two monolayer surfaces. A rectangular grid is placed over the simulation XY plane and the amphiphiles are assigned to grid cells according to the X,Y coordinates of their head beads. The upper monolayer contains those amphiphiles whose head Z coordinates are greater than their tail Z coordi-
nates. The lower monolayer contains all other amphiphiles. This defines the number of amphiphiles in each monolayer. The average Z coordinate of the head beads in each grid cell for each monolayer is used to define the height of the monolayers at each point \((X,Y)\). This procedure yields a two-dimensional height field, \(h(X,Y)\), for each monolayer. The bilayer width, \(\ell_{me}\), is the difference between these heights averaged over all grid points, and the surface area of each monolayer, \(A\), is the sum of the areas of the triangles composing it. The grid cell width is typically twice the bead radius, so that there are at least several amphiphiles per cell. This smears out the excursions of single amphiphiles from the monolayers ensuring that the monolayer area is not overly sensitive to small-scale fluctuations.

Figure 2. Bead density profiles for head and tail beads in the bilayer of Fig. 1 and the bulk water. The simulation box is divided into 128 slices of thickness \(r_0/4\) parallel to the bilayer surface, and the time average of the number of beads of each type in the slices is used to generate the density profiles. Water is excluded from the bilayer interior by the strong hydrophobic repulsion of the tail beads, and the head beads are localised at the water-hydrophobic interface. The tail bead profile includes all the tail beads, and shows that the density in the hydrophobic region is uniform except for a small dip at the bilayer midplane. All error bars are similar to the one shown and indicate the statistical accuracy of the ensemble averages extracted from the simulations.

The lateral density profiles of the hydrophobic and hydrophilic beads in the \(HC_6\) bilayer, together with the water beads, are shown in Figure 2. It is clear from Figures 1 and 2 that the amphiphiles terminate close to the bilayer midplane, with the final tail bead confined to that region, and that they exhibit strong \(L_\alpha\) order. The profiles are calculated by averaging the bead densities over thin slices (of width \(r_0/4\) parallel to the bilayer surface. The water is excluded from the hydrophobic region of the bilayer by the strong tail-water repulsion, whereas the head beads can sometimes penetrate to the centre, although this is hardly visible in this figure. Two effects contribute to this penetration: amphiphiles can burrow their way through the bilayer emerging into the apposed monolayer, a process known as "flip-flop"; and the thermally-excited undulations of the bilayer lead to density profiles that include contributions from non-planar bilayer configurations. The tail bead
density shows a small dip at the midplane of the bilayer, indicating that the monolayers are not significantly interdigitated. Bilayers composed of linear amphiphiles \( HC_n \) with \( n = 4 – 10 \) exhibit these general properties, although the shape fluctuations are stronger for shorter-chain amphiphiles. Amphiphiles containing four or more tail beads are not observed to leave the bilayer and exchange with the solvent on the time-scale of the simulations, which is of the order of microseconds.

Although a bilayer forms from \( HC_n \) surfactants for a wide range of tail-water repulsions, \( \sigma_{TW} \), two constraints should be satisfied before the simulated bilayer has properties that match the typical structure of a lipid bilayer. The average amphiphile end-to-end length should be approximately one half of the bilayer width, so that the two monolayers do not interdigitate; and the amphiphiles should be straight and oriented along the bilayer normal. We present results for a region of the simulation parameter space satisfying these constraints by adjusting the bond stretching and stiffness potential parameters. Under these constraints, the amphiphile length can be meaningfully compared to that of lipid molecules, and the bilayer width scales linearly with the amphiphile tail length. The end-to-end length, \( L_{ee} \), of an \( HT_n \) amphiphile is measured from the centre-to-centre distance from the H bead to that of the last T bead. Two bond types are present in these amphiphiles: \( HT \) and \( TT \) although, for simplicity, both types have the same bond strength parameters of \( k_2r_0^2/k_BT = 128, l_0/r_0 = 0.5 \). These values yield average bond lengths of \( < l_{HT}/r_0 > = 0.62 \pm 0.01 \) and \( < l_{TT}/r_0 > = 0.55 \pm 0.01 \). The larger \( < l_{HT}/r_0 > \) length reflects the repulsion of the head from the tail beads, and its greater freedom to fluctuate into the adjacent water than is possessed by the tail beads deeper in the hydrophobic bulk. The standard deviation of the bond lengths is less than 2% for all tail lengths and bilayer tensions. The amphiphile end-to-end length for the tensionless \( HT_6 \) bilayer is \( < L_{ee}/r_0 > = 3.23 \pm 0.01 \), which can be compared with the bilayer thickness \( < \ell_{mc}/r_0 > = 6.61 \pm 0.01 \), showing that the two monolayers are not interdigitated.

The end-to-end length grows linearly with the number of tail beads for all \( HT_n \) amphiphiles for a fixed bending stiffness \( k_3/k_BT \) (data not shown) and is only slightly affected by the tension on the bilayer in the regime we study. It decreases slowly as the projected area per amphiphile, \( A_{pr}/Nr_0^2 \), increases, with the decrease being larger for longer amphiphiles. For \( HT_8 \) amphiphiles, it decreases from \( < L_{ee}/r_0 > = 4.28 \pm 0.01 \) for \( A_{pr}/Nr_0^2 = 0.60 \) to \( < L_{ee}/r_0 > = 4.24 \pm 0.005 \) for \( A_{pr}/Nr_0^2 = 0.64 \). The absence of a corresponding variation in the bond lengths indicates that this shortening occurs by the bonds rotating relative to each other. Such rotations are resisted by the amphiphile chain stiffness potential discussed next. The surface tension increases from \( \sigma_0^2/k_BT = -3.0 \) to \( \sigma_0^2/k_BT = 7.3 \) over this span of areas indicating that \( HT_8 \) bilayers are quite rigid. All of our results for tensionless bilayers are taken from the regime \( < \ell_{mc}/L_{ee} > \approx 2.0 \), in which the monolayers are not significantly interdigitated. Bilayers composed of \( HT_6 \) or longer amphiphiles are quite rigid and show only small shape fluctuations even close to zero surface tension.

At small values of \( A_{pr}/Nr_0^2 \), the \( HT_4 \) membranes show substantial undulations, the fluctuations being gradually suppressed for amphiphiles with longer tails. In addition, a tensionless state could not be found for \( HT_4 \) bilayers simply by reducing the projected area because some amphiphiles inverted and buried their heads inside the hydrophobic region generating a positive surface tension. A well-ordered bilayer forms with a projected area of \( A_{pr}/Nr_0^2 = 0.70 \), and surface tension \( \sigma_0^2/k_BT = 0.013 \pm 0.016 \), when the
head-water repulsion parameter is decreased from $a_{HW} = 35$ to $a_{HW} = 27$, keeping all other potential parameters unchanged. Increasing $A_{pr}/N r_0^2$ for all bilayers composed of $HT_n$ amphiphiles reduces their shape fluctuations and width and, beyond a certain point, ruptures the membranes by the appearance of a pore. The amphiphiles around the pore rim reorient so as to shield their hydrophobic tails from the surrounding solvent (data not shown). Although interesting as a possible model of pore formation, we do not investigate bilayer rupture further here.

In the absence of chain stiffness, $k_3/k_B T = 0$, stable bilayers are formed in which the amphiphile tails represent (almost) freely-jointed chains, but the hydrophobic region lacks the characteristic $L_{\alpha}$ order of lamella phase lipid bilayers. The hydrophilic head beads also show a tendency to penetrate into the hydrophobic interior, although this effect decreases as the chain stiffness increases. A chain bending stiffness of $k_3/k_B T = 20$, and a preferred angle of $\phi_0 = 0$, causes the amphiphiles to align approximately parallel to the bilayer normal as seen in Figure 1. The linearity of the amphiphile chains may be estimated from the angle between adjacent bonds averaged over the chain length. More specifically, an orientational order parameter may be defined as the second Legendre polynomial of the scalar product of adjacent bond vectors. In the absence of a bending stiffness potential this order parameter is close to zero. With the above-mentioned values of $k_{2r_0^2}/k_B T = 128$, $l_0/r_0 = 0.5$ and $k_3/k_B T = 20$, $\phi_0 = 0$, the order parameter for amphiphiles containing from 4 to 10 tail beads is always in the range $0.8 - 0.9$.

5 Lateral Stress Profile for Bilayers of Linear Amphiphiles

Lipid bilayers around cells and vesicles are often close to a tension-free state. The distribution of stresses within the bilayer is not, however, uniform. Recent calculations indicate that the presence of cosurfactants, and hydrocarbon chain stiffness due to unsaturated C-C bonds, modulate the membrane stress distribution. We have measured the stress distribution within the model bilayers as a function of amphiphile tail length and stiffness.

The calculation of the stress tensor for a system composed of point particles interacting via continuous potentials is described in Ref. 12, which extends earlier work of Schofield and Henderson. In this method, the contributions to the stress profile of the bead-bead interactions, bond stretching and bond bending stiffness potentials are averaged over thin slices parallel to the bilayer surface. We define the stress profile, $\sigma(z)$, as the difference of the normal and lateral components of the stress tensor summed over all potentials. For further details of the calculation, the reader is referred to previous work. The stress profile for the $HT_6$ bilayer described in Section A is shown in Fig. 3, and exhibits very similar structure to that seen in the MD simulations of Ref. 12. However, the absence of Lennard-Jones potentials in our simulations does not visibly modify the stress profile, indicating that the infinite barrier of Lennard-Jones potentials is not essential for capturing the stress profile structure in mesoscopic simulations. The integral of $\sigma(z)$ across the bilayer yields the surface tension. We approximate this integral by a summation across the whole simulation box in the $Z$ direction because the contribution from regions containing only solvent vanishes. Figure 4 shows that the surface tension of bilayers composed of $HT_5$ to $HT_{10}$ amphiphiles increases as $A_{pr}/N r_0^2$ increases. The increase is steepest for the longer amphiphiles, and becomes sub-linear at large values of the projected area.

Reducing the tail bending stiffness for $HT_6$ amphiphiles to $k_3/k_B T = 10$, while
Figure 3. Stress profile for the $HT_6$ bilayer shown in Fig. 1. The total stress is the sum of three contributions: the repulsion potentials between head, tail and water beads; the Hookean spring potentials connecting adjacent beads in the amphiphiles, and the chain stiffness potential between adjacent triples of hydrophobic beads. The outer positive peaks at the monolayer-water interfaces are due to the repulsion between the hydrophobic tail beads and the hydrophilic head beads and water. The adjacent negative troughs arise from the Hookean bond potential compressing the amphiphiles, while the inner positive peaks near the monolayer midplanes are due to the chain stiffness potential.

keeping all other parameters constant, causes the bilayer shown in Fig. 1 to swell as some amphiphiles invert so that their heads are buried in the hydrophobic region. This reduces the lamella order of the tails to 0.7, and creates a large positive surface tension, $\sigma r_0^2/k_BT = 1.62$, across the bilayer. The surface tension remains positive as the projected area is varied, although it reaches a minimum, and the bilayer regains some of its ordered nature, for $A_{pr}/N r_0^2 = 0.7$, at which point the mean bilayer width is $<\ell_{me}/r_0>= 5.71 \pm 0.03$. The bilayer may be restored to a tensionless state by reducing the head-water repulsion parameter, $a_{HW}$ to 30, and increasing the projected area to $A_{pr}/N r_0^2 = 0.71$, but this leaves the chain order in the hydrophobic region at the reduced value of 0.7, and the bilayer width at $<\ell_{me}/r_0>= 5.72 \pm 0.02$. This result shows that the chain bending stiffness and amphiphile head-water repulsion parameters cannot be independently varied in a simulation to produce a tensionless bilayer, but that they play effectively opposite roles in controlling the tendency of amphiphiles to invert and bury their heads in the hydrophobic region. Reducing the chain stiffness of amphiphiles in a tensionless bilayer requires a simultaneous reduction in their head-water repulsion and an increase in the projected area to restore the bilayer to a tensionless state. The surface tension for bilayers near their tensionless state is very sensitive to changes in the area per amphiphile: adding or removing just three amphiphiles from the tensionless $HT_6$ bilayer of Fig. 1, which contains 3321 amphiphiles, changes its surface tension from $\sigma r_0^2/k_BT = -0.03 \pm 0.1$ to $\sigma r_0^2/k_BT = -0.15$ and $\sigma r_0^2/k_BT = 0.10$ respectively. The
surface tensions for an $HT_n$ bilayer containing 3368, 3365, 3362 amphiphiles in a fixed simulation box of size $(32/r_0)^3$ are $\sigma r_0^2/k_BT = -0.24, 0.024, 0.21$ respectively. Away from the tensionless state, the surface tension changes more rapidly with projected area for longer tail amphiphiles. Similarly, although the equilibrium bilayer structure is insensitive to the exact value of the tail-water repulsion parameter in the range $a_{TW} = 65 - 85$, the width of the bilayer changing by less than 0.2% to $6.60 \pm 0.01$ and $6.62 \pm 0.01$ for the two extreme values, the peak heights in the lateral stress profile, and therefore the surface tension, are very sensitive to this parameter. Changing $a_{TW}$ for the bilayer of Fig. 1 to 65 and 85 changes the surface tension from approximately zero to $\sigma r_0^2/k_BT = -0.58, 0.55$ respectively, all other parameters being constant. This has the effect of moving the tensionless bilayer state to higher and lower projected areas respectively.

Figure 4. Variation of the surface tension with the projected area per amphiphile, $A_{pr}/N r_0^2$, for bilayers composed of $HT_n$ amphiphiles for several tail lengths. The surface tension is obtained by integrating the stress profile across the bilayer as described in the text. The lines connecting the points are only to guide the eye but show that the surface tension varies linearly around its zero point, and shows a sub-linear dependence on area at large projected areas. The preferred area per amphiphile, at which the surface tension vanishes, decreases as the amphiphile tail length increases.

6 Concluding Remarks

A major goal of computer simulations is to predict the material properties of mesoscopic, or possibly macroscopic, aggregates given only the elementary molecular constituents composing them. Although, perhaps, less useful for simulating interactions between hard colloidal particles in solution, DPD shows great promise when applied to soft complex fluids. We have shown that model amphiphiles consisting of a hydrophilic headgroup attached to
a hydrophobic tail form planar bilayers whose density profile and lateral stress distribution, and the dependence of these properties on the tail length, agree at least qualitatively with experiments and previous coarse-grained MD simulations. The task of capturing just those microscopic properties of amphiphiles that give rise to the material properties and dynamical behaviour of lipid bilayers or polymersomes is a challenge whose solution will have direct applications to chemistry, materials science and medicine. We believe that DPD is highly suited to complex fluid simulations because it allows the mesoscopic behaviour of large systems to be followed for long times at relatively small computational cost.

Acknowledgments

Membrane simulations were performed with the aid of a grant of computer time provided by the NIC of the Research Centre Jülich.

References

Massively Parallel Direct Numerical Simulation of Turbulent Combustion

Marc Lange$^{1,2}$ and Jürgen Warnatz$^1$

$^1$ Interdisciplinary Center for Scientific Computing (IWR)
Heidelberg University, INF 368, 69120 Heidelberg, Germany

$^2$ High-Performance Computing-Center Stuttgart (HLRS)
Stuttgart University, Allmandring 30, 70550 Stuttgart, Germany
E-mail: lange@hlrs.de

A parallel code for the direct numerical simulation of turbulent reactive flows is presented. Detailed models are used for chemical kinetics and molecular transport. The performance of this code on the Cray T3E is briefly discussed. Several application examples regarding non-premixed and premixed turbulent combustion are given.

1 Introduction

Combustion processes are important for a wide range of applications like automotive engines, electrical power generation, and heating. This leads to an enormous interest in improved predictive methods to aid the design of practical combustion systems with improved economy and reduced pollutant emissions. In most of these applications the reactive system is turbulent and the reaction progress is influenced by turbulent fluctuations and mixing in the flow. Due to the broad spectrum of length and time scales apparent in turbulent reactive flows, a direct numerical simulation (DNS), i.e. the direct solution of the equations describing the dynamics of reacting gas mixtures as given in Sect. 2, of most practical combustion devices will remain computationally prohibitive for the near future. The simulation of such systems must therefore continue to be done by the formulation and solution of model equations involving some form of averaging$^1$. The nonlinear equations of fluid mechanics when averaged contain additional unknown quantities, the Reynolds stresses, so the averaged equations must be supplemented by appropriate model expressions before a closed system of equations can be obtained. The coupling between the chemical kinetics and fluid dynamics constitutes one central problem in turbulent combustion modeling$^2$.

During the last couple of years, DNS have become an important tool to study turbulent combustion at a fundamental level. The detailed information about turbulence-chemistry interactions provided by DNS has been extremely valuable in the development and validation of turbulent combustion models$^3$. However, many of the DNS carried out so far have used simple one-step reaction mechanisms and/or oversimplified models for molecular transport. It has been shown$^4,5$ and will also become clear from the applications presented later in this paper that several important effects cannot be captured by simulations with such oversimplified models. By making efficient use of the computational power provided by parallel computers, it is possible to perform DNS of reactive flows using detailed chemical reaction mechanisms at least in two spatial dimensions$^6,7$. Nevertheless, computation time is still the main limiting factor for the DNS of reacting flows, especially in the case of using realistic models for chemical kinetics and molecular transport.
Chemically reacting flows can be described by a set of coupled partial differential equations expressing the conservation of total mass, the masses of the $N_S$ chemical species, momentum, and energy. Employing tensor notation, where no summation over the index $\alpha$ which denotes the chemical species is performed, these equations can be written as

$$
\frac{\partial \vec{u}}{\partial t} + \frac{\partial (\vec{u} \vec{u})}{\partial x} = 0,
$$

$$
\frac{\partial (\rho Y_\alpha)}{\partial t} + \frac{\partial (\rho Y_\alpha u_i)}{\partial x_i} = -\frac{\partial j_{i\alpha}}{\partial x_i} + M_\alpha \dot{\omega}_\alpha \quad (\alpha = 1, \ldots, N_S),
$$

$$
\frac{\partial (\rho u_i)}{\partial t} + \frac{\partial (\rho u_i u_j)}{\partial x_j} = \frac{\partial \tau_{ij}}{\partial x_i} - \frac{\partial p}{\partial x_i},
$$

$$
\frac{\partial e_t}{\partial t} + \frac{\partial ((e_t + p)u_i)}{\partial x_i} = \frac{\partial (u_i \tau_{ij})}{\partial x_i} - \frac{\partial q_i}{\partial x_i}.
$$

Herein, $\rho$ denotes the density and $\vec{u}$ the velocity. $Y_\alpha$, $j_{i\alpha}$ and $M_\alpha$ are the mass fraction, diffusion flux, and molar mass of the species $\alpha$. $\tau_{ij}$ denotes the viscous stress tensor and $p$ the pressure. $\vec{q}$ is the heat flux and $e_t$ is the total energy given by

$$
e_t = \rho \left( \frac{u_i u_i}{2} + \sum_{\alpha=1}^{N_S} h_\alpha Y_\alpha \right) - p,
$$

where $h_\alpha$ is the specific enthalpy of the species $\alpha$. The term $\dot{\omega}_\alpha$ on the right-hand sides of Eq. (2) is the chemical production rate of the species $\alpha$, which is given as the sum of the formation rates in all $N_R$ elementary reactions,

$$
\dot{\omega}_\alpha = \sum_{\lambda=1}^{N_R} k_\lambda (\nu_{i\alpha\lambda}^{(p)} - \nu_{i\alpha\lambda}^{(r)}) \prod_{\alpha=1}^{N_S} c_{\alpha\lambda}^{(r)}.
$$

Herein $\nu_{i\alpha\lambda}^{(p)}$ and $\nu_{i\alpha\lambda}^{(r)}$ denote the stoichiometric coefficients of reactants and products respectively and $c_{\alpha\lambda}$ is the concentration of the species $\alpha$. The rate coefficients $k_\lambda$ of the elementary reactions are given by a modified Arrhenius law

$$
k_\lambda = A_\lambda T^{\beta_\lambda} \exp \left( -\frac{E_{a\lambda}}{RT} \right).
$$

The parameters $A_\lambda$, $\beta_\lambda$ of the pre-exponential factor and the activation energy $E_{a\lambda}$ are determined by a comparison with experimental data. A detailed mechanism with 37 elementary reactions among 9 species by Maas and Warnatz has been used to describe the chemical kinetics of the $H_2/O_2$/$N_2$ system. The methane-air mechanism employed for the DNS presented in Sect. 6 comprising 84 elementary reactions among 15 species is based on that of Smooke. Besides the computation of the reaction kinetics, detailed models are also utilized for the computation of the thermodynamical properties, the viscosity and the molecular and thermal diffusion velocities. These equations are complemented by a state equation, generally the ideal gas law

$$p = \frac{\rho}{\bar{M}} RT
$$

with $R$ being the gas constant and $\bar{M}$ the mean molar mass of the mixture.
3 Structure and Performance of the Parallel Code

A code has been developed for the DNS of reactive flows using detailed models for chemical kinetics and molecular transport as described in the last section. This code has been designed for parallel computers with distributed memory like the Cray T3E as the target platform\textsuperscript{12,7}. As at least parameter studies using DNS with detailed chemistry and transport in three spatial dimensions are computationally prohibitive, we restrict ourselves to two spatial dimensions. As turbulence is an inherently three-dimensional phenomenon this is a serious drawback. Nevertheless, we consider the inclusion of a realistic description of chemical reactions and molecular transport to be at least as important as an extension to three dimensions.

The spatial discretization in this code is performed using a finite-difference scheme with sixth-order central-derivatives, avoiding numerical dissipation and leading to very high accuracy. Depending on the boundary conditions, lower order schemes are used at the outermost grid points. The integration in time is carried out using a fourth-order fully explicit Runge-Kutta method. The time step of the integration is controlled by three different limiting conditions. A Courant-Friedrichs-Lewy (CFL) criterion and a Fourier criterion for the diffusion terms are checked to ensure the stability of the integration. An additional accuracy control of the solution is obtained through time-step doubling.

The parallelization strategy is based on a regular two-dimensional domain decomposition with a three points wide region of halo elements at the domain boundaries. Using these locally stored values, an integration step on each subdomain is carried out independently from the other nodes. After each integration step, the new inner boundary values of the subdomain are sent to and the new values of the surrounding region are received from the neighbouring nodes via message-passing communication.

Below, we present performance results for the Cray T3E optimized implementation of our code using MPI as message-passing library. All computations have been performed on Cray T3E-900 systems, i.e. 450 MHz clock speed and stream buffers enabled. Having access to a node with 512 MB RAM allows us to perform a one-processor reference computation for the \( \text{H}_2/\text{O}_2/\text{N}_2 \) system with \( 544 \times 544 \) grid points, a problem size which corresponds to some real production runs. The achieved speedups and efficiencies for this benchmark are given in Table 1. An average rate of 86.3 MFlop/s per PE is achieved in this computation using 64 processors\textsuperscript{13}.

The classical parallel efficiency for a problem of fixed size as given above obviously includes the loss due to the overhead for communication and synchronization as well as differences caused by the decreasing load per processor. (Of course the amount of data which has to be sent to and received from other processors is also decreasing in this case.) An alternative measure is the scaled parallel efficiency, i.e. the ratio of the computing time for solving an problem of the size \( p \cdot N \) using \( p \) processors and the computing time for

<table>
<thead>
<tr>
<th>processors</th>
<th>1</th>
<th>4</th>
<th>8</th>
<th>16</th>
<th>32</th>
<th>64</th>
<th>128</th>
<th>256</th>
<th>512</th>
</tr>
</thead>
<tbody>
<tr>
<td>speedup</td>
<td>1</td>
<td>4.3</td>
<td>8.1</td>
<td>15.9</td>
<td>30.5</td>
<td>57.9</td>
<td>108.7</td>
<td>189.0</td>
<td>293.6</td>
</tr>
<tr>
<td>efficiency</td>
<td>100.0</td>
<td>106.6</td>
<td>100.7</td>
<td>99.2</td>
<td>95.3</td>
<td>90.4</td>
<td>84.9</td>
<td>73.8</td>
<td>57.4</td>
</tr>
</tbody>
</table>

Table 1. Scaling on a Cray T3E for a DNS with 9 species and 37 reactions on a \( 544^2 \) points grid
solving the corresponding problem of size $N$ on one processor. Choosing a constant load per processor of $32^2$ grid points and the chemical mechanism for the hydrogen-air system, a scaled parallel efficiency of 76.8% has been achieved using 256 PEs of a Cray T3E-900.

One of the most time-consuming parts in DNS like those presented in the following sections is the computation of the chemical source terms on the right-hand sides of the species-mass conservation equations by evaluating Eq. (6). In several cases of interest, a vast reduction of this time can be achieved without a significant loss in accuracy by using an adaptive evaluation of the chemical source terms. A dynamic load-balancing procedure is used in the parallel adaptive computations to maintain a high efficiency. This method is described in detail in Ref. 14 for non-premixed turbulent combustion and has been extended to premixed turbulent flames as described in Ref. 15.

4 Autoignition in a Turbulent Mixing Layer

Autoignition takes place in combustion systems like Diesel engines, in which fuel ignites after being released into a turbulent oxidant of elevated temperature. The influence of the turbulent flow field on the ignition-delay time and the spatial distribution of ignition spots is studied in a model configuration shown in Fig. 1. A cold fuel stream in the left of the computational domain and an air stream with an elevated initial temperature in the right half of the computational domain are superimposed with a turbulent flow field computed by inverse FFT from a von-Kármán-Pao-spectrum with randomly chosen phases. Non-reflecting outflow conditions are used at the boundaries in $x$-direction and periodic boundary conditions are used in $y$-direction. In the simulation presented here, the fuel stream consists of 10% hydrogen and 90% nitrogen (mole fractions) at a temperature of $T_1 = 298$ K and the initial temperature of the air stream is $T_2 = 1100$ K. The Reynolds number of the turbulent flow field based on the integral length scale is $Re_A(t_0) = 252$.

![Figure 1. Initial conditions for the DNS of autoignition in a turbulent mixing layer](image1)

![Figure 2. Ignition spots at the time of strongest increase of maximum heat release rate](image2)
After a specific temporal delay, which is characteristic for chemical reactions which are
governed by a chain-branching mechanism, a very fast increase of the heat release rate
$\dot{\Omega}$ in distinct spatial regions is observed. Fig. 2 shows these ignition spots for one such
simulation. The flow field visualized by line integral convolution\cite{17} after one millisecond
is shown in Fig. 3 while the colors mark the temperature, which has risen up to
$T_{\text{max}}(\vec{x}, t_0 + 1 \text{ ms}) = 1549 \text{ K}$ due to the heat released in the flame front.

Main results of these DNS regard the influence of turbulence on the ignition delay
time and the spatial occurrence of the ignition spots\cite{5}. For the latter problem, it has been
shown analytically using simplified models that in the corresponding laminar situation aut-
ignition takes place at the coordinate where the mixture fraction has a specific value\cite{18}.
The ignition spots in the DNS of turbulent mixing layers also appear along isolines of the
mixture fraction. As in the laminar case, the specific value of the most reactive mixture
fraction depends on the temperatures and chemical compositions of the fuel and oxidizer
streams. The curvature of the mixture-fraction isolines has been identified as an additional
parameter determining the spatial location of the ignition spots. In regions with curvature
being convex towards the fuel side, the focussing of the faster diffusing species like atomic
hydrogen increases their concentration on the hot fuel-lean side of the mixing layer. Obvi-
ously, it would not have been possible to identify this effect using an oversimplified model
for chemical kinetics or a description of the transport omitting differential diffusion.

5 Induced Ignition of a Turbulent Hydrogen-Air Mixture

Induced ignition is another phenomenon of practical importance, e. g. in Otto engine com-
busion and safety considerations. DNS studies of this process have been performed using
simple one-step chemistry in a model configuration of an initially uniform premixed gas under turbulent conditions which is ignited by an energy source\textsuperscript{19}. The same configuration has been used for the detailed chemistry DNS presented here. A cold ($T = 298$ K) hydrogen-air mixture is superimposed with a turbulent flow field. During the first 15\,\mu{s} of the simulation an energy-source in a small circular region at the center of the computational domain is active. During this time the mixture at the center of the domain heats up, radicals are formed, and the mixture ignite. A shockwave is observed which propagates outwards towards the boundaries of the domain. Non-reflecting outflow conditions based on characteristic wave relations\textsuperscript{16} are imposed on all boundaries to allow the shock wave to leave the domain without disturbing the solution. Above a minimum ignition energy an expanding flame kernel is observed.

Figure 4 shows the temporal evolution of such a flame kernel in the turbulent flow field after the ignition. The first row of images shows the mass fraction of $\text{H}_2\text{O}_2$, a radical which is confined to a very thin layer in the flame, at $t = 0.2$\,ms, $t = 0.4$\,ms, and $t = 0.6$\,ms, respectively. Below, the temporal evolution of vorticity is shown. The lack of a vortex-stretching mechanism in two-dimensional simulations of decaying turbulence leads to an inverse cascade with growing structures. There is a very strong damping of the turbulence in the hot region of the burnt gas due to the high viscosity. A comparison shows a good qualitative agreement of the DNS presented with the results of an experimental investigation of turbulent flames performed under similar conditions\textsuperscript{20}.

A quantitative analysis requires specialized tools for the postprocessing of the large and complex datasets generated by DNS. An extensible tool has been developed which allows to extract several features from these datasets. This is illustrated in Figs. 5 and 6 which exemplify results gained from the simulation of the turbulent flame kernel shown.

![Figure 4: H$_2$O$_2$ mass fraction (top) and vorticity (bottom) at $t = 0.2$\,ms, $t = 0.4$\,ms, and $t = 0.6$\,ms (left to right) in a turbulent flame kernel evolving after induced ignition.](image)
in Fig. 4. In Fig. 5 the strain rate and the flame thickness (based on the local temperature gradient) along the flame front at $t = 0.6 \text{ ms}$ are shown. The correlation of the $\text{H}_2\text{O}_2$ mass fraction and the curvature of the flame is shown in Fig. 6 for subsets of the DNS data with $T = (T_{\text{max}} + T_{\text{min}})/2 = 1140 \text{ K}$. High concentrations of the $\text{H}_2\text{O}_2$ radical in the reaction zone evidently occur in regions with high negative curvature, i.e. convex towards the burnt gas. Such changes of the chemical composition in curved flame fronts are caused by preferential diffusion\textsuperscript{21,22}.

## 6 Turbulent Premixed Methane Flame

The structure and propagation of turbulent flames has also been studied by DNS of the temporal evolution of an initially planar flame in a turbulent flow field\textsuperscript{23,5}. A snapshot of such an synthetic turbulent premixed methane-air flame is shown in Fig. 7. The third coordinate direction is used here to visualize the fuel mass fraction $Y_{\text{CH}_4}$ where the contour colors denote the mass fraction of OH which is an important intermediate radical. The initially flat flame has been interacting with the turbulent flow field for $1.35 \text{ ms}$ which is about 1.8 times the large-eddy turnover time $t_\Lambda$. An interesting phenomenon to be seen in Fig. 7 is the isolated pocket of cold unburnt fuel which has propagated into the hot burnt gas side of the flame. The occurrence of such pockets in turbulent flames is well known from experiments\textsuperscript{24}. The transient process of pocket formation involves mutual annihilation of parts of the flame front and is discussed and shown in more detail in Refs. 23, 25 for the case of hydrogen flames.

Figure 8 shows the spatial distribution of the local heat release rate in the same turbulent flame at the same time as shown in Fig. 7. The main heat release occurs inside a thin layer with strong variations along the flame front which are mainly determined by the curvature of this reaction layer. The local heat release rate is one of the most important properties of turbulent flames with respect to turbulent combustion modeling. The direct experimental measurement of the heat release rate is extremely difficult, if not impossible. The only practical approach is to measure some other quantity that has some correlation...
with this rate over the relevant range of flame and flow parameters. It is e.g. possible to perform spatially resolved measurements of concentrations of trace species like OH, CH, and HCO. An extremely useful result is therefore the fact that the local heat release rate is very well correlated with the HCO concentration as can be seen in Fig. 9. This holds for

Figure 7. Snapshot of fuel (height) and OH (colors) mass fractions in a turbulent premixed methane-air flame

Figure 8. Heat release rate in a turbulent premixed methane-air flame at $t = 1.8 \tau_A$  

Figure 9. Correlation of local heat release rate with HCO mass fraction

426
the complete temporal evolution and also for methane-air flames with different equivalence ratios and for varied turbulence intensity.

7 Conclusion

A code for the DNS of turbulent reactive flows using detailed models for chemical kinetics and molecular transport has been presented. This code exhibits good performance and an excellent scaling behaviour on Cray T3E systems. It has been used to study non-premixed and premixed turbulent combustion processes in several configurations. Exemplary application results presented in this paper regard the influence of turbulence on autoignition in a mixing layer, the structure of a flame kernel evolving after induced ignition of a turbulent hydrogen-air mixture, and pocket formation in a premixed methane-air flame propagating in a turbulent flow field. In this last case, a correlation analysis reveals the suitability of HCO as a trace species for the heat release rate. Making use of adaptive techniques like the adaptive evaluation of chemical source terms mentioned in Sect. 3 and harnessing the power of modern supercomputers will remain to be of key importance for future DNS. The ongoing increase of computational power available will further strengthen the usage of DNS as a tool in combustion research.

Computational Resources

All presented DNS have been performed using the Cray T3E systems of the John von Neumann Institute for Computing at Jülich (NIC) and of the High-Performance Computing Center at Stuttgart (HLRS). The CPU-time for individual DNS depends on several parameters including the size of the computational grid, the number of chemical species and elementary reactions considered, the physical time for which the temporal evolution is computed, the achievable timesteps for the integration, etc. A good example for a typical production run is the DNS of a turbulent reactive mixing layer as presented in Sect. 4 up to a physical time of one millisecond on an $800 \times 800$ points grid. Using 256 PEs of the Cray T3E-1200 this computation took about 23 hours. Our study of this configuration includes several such DNS in which parameters like turbulence intensity and the initial temperature of the air stream have been varied.

References


Earth system modelling has always put highest demands on both memory and speed aspects of supercomputing. It can be foreseen, that this will not change in the future due to at least four reasons:

1. Any subsystem we want to model, be it the atmosphere, the ocean or subsections of it, like boundary layers or limited areas, has quite strong interactions with many other subsystems. To take account of these interactions increasing the model area to include the interacting systems in one model is mostly inevitable. A good example are weather forecast models, which nowadays have to be global and soon have to include parts of the world oceans in order to further improve our ability to predict reliably e.g. the temperature in Jülich only several days ahead.

2. The interactions within Earth systems and at their boundaries cover always a very broad continuum of scales in both space and time. Any resolution cut-off demanded by restrictions in memory and computing speed must be payed for very dearly by the necessity to parameterize the cut-off processes as functions of the larger scales. It is obvious, that this will limit the quality of any simulation. Again weather forecast models may serve as an example: Their resolution has been constantly increasing from hundreds of kilometers in the beginning of numerical weather forecasting when the first electronic computers became available to now 7 km (e.g. the Lokalmodell (LM) of the German weather Service (DWD)). Still, the processes which affect our daily life most like clouds and precipitation, but also the air turbulence, are not yet modelled on the basis of first principles in the models but parameterized within the so-called physics-modules.

3. The complexity and number of independent variables, which need to be taken into account in Earth system modelling, increases steadily and is virtually unlimited. Until recently, e.g. weather forecast models were based on the prediction of only five parameters (windspeed and horizontal direction, temperature, pressure and humidity) by solving simultaneously five interdependent partial differential equations in space and time at any point of the globe. Explicit treatment of the vertical velocity and cloud water have only recently entered the models. It can be foreseen, that first several condensate classes of hydrometeors (snow, hail, rain) will need be added, followed by subdividing these in several size classes, and followed by the addition of aerosol classes and trace gases (chemical weather).

4. Finally, by the extensions of the models as described above, both the amount of data to initialize the models, but even more the output produced during a simulation will multiply and has already achieved such a status, that output storage and output analysis has become a subject of major concern in Earth system modelling.

The three papers in this sections provide excellent examples especially for the first three topics described above. The papers are ordered according to scale beginning with mesoscale ocean modelling around 100 km resolution (Beismann and Redler), where sub-
scale parametrizing of exchange processes is crucial. It follows mesoscale atmospheric chemistry modelling at a few kilometer resolution (Wolke et al.) where effort is put on intelligently organizing computer efficiency. The last contribution concerns large eddy simulation of the structure of the marine atmospheric boundary layer at 100 m resolution where both the inclusion of wet processes and a sufficiently large modelling area is stressed. (Raasch and Schröter).

The importance of sub-scale mixing processes for both the uptake and redistribution of tracers, and the formation and spreading of water masses in the world ocean is shown by Beismann and Redler. They use a medium resolution model (4/3 x 4/3 degrees) of the Atlantic Ocean basin to simulate the behaviour of the North Atlantic Deep Water (NADW) water, which is believed to be a highly influential water mass for decadal climate variations. They show, that the modelled behaviour of this water mass is strongly dependent on the mixing parametrisation applied. It becomes obvious, that small-scale phenomena like the filament-like Deep Western Boundary Current (DWDC), which cannot be adequately resolved by a model of this resolution, is a crucial component of the whole system. Higher resolution models are necessary to correctly account for its influence.

Wolke et al. address in detail the problems encountered when coupling a weather forecast model (LM) with a chemistry transport model (MUSCAT). In the present version coupling is only one-way: The meteorological fields simulated by the LM are used for solving the equations of 73 species on the basis of 237 reactions. Both models are state-of-the-art models in their respective categories and extremely demanding concerning computer memory and speed. Multi-grid techniques are used to concentrate computing power in model areas, where highest resolution is necessary by dynamic allocation of processor loading. The coupling itself poses additional demands on efficiently organizing memory and data exchange. The authors can show, that especially the way MUSCAT is parallelized has achieved a very high efficiency. The expense of CPU power on coupling between the individual processors has been minimized.

Finally, Raasch and Schröter show a very successful application of massive by parallel computing for so-called Large Eddy Simulation (LES). LES extend atmospheric modelling from the scale applied for weather forecast models (usually around 10 km) down to below 100 m. A novel feature of this particular model is the inclusion of cloud physics. The authors were able to prove, probably for the first time, that the differences in aspect ratio (height to horizontal extend) between the classical dry Benard convection and the cellular convection in cold air outbreaks over warm ocean areas is caused by the moist processes. Liberation of latent heat during condensation of cloud water and its use-up during evaporation lead to much flatter observed aspect ratios. An important prerequisite for the successful simulations was the possibility to largely increase the model area while keeping the high spatial resolution. It was necessary to model simultaneously many cells, which gives another hint, that this observed atmospheric process, most probably also many others, cannot be reproduced even in its basics by idealized settings which ignore both the effects of small scale processes and the large scale influences.
1 Introduction

The frequent occurrence of so-called mesoscale cellular convection (MCC) patterns over vast regions of the oceans has been revealed by satellite imaging and its properties have been reviewed in detail by various authors\(^1\). Areas of convective activity are usually associated with the flow of cold air over warm water such as in cold air outbreaks (CAOB). In such a case an initially cloud-free and presumably stable stratified air mass is advected from land or from an ice surface over a warmer sea surface. The heating and moistening from below causes the development of a convective boundary layer, in which a typical sequence of convection patterns can be observed: As soon as the cold air hits the warm sea, roll-like secondary flow patterns develop, appearing in satellite images as so-called cloud-streets. Further downstream the roll pattern gradually changes to hexagonal cellular convection patterns, which can occur with either open or closed cells. Open cells consist of broad central areas of descending air, enclosed by relatively narrow rings of updrafts, whereas closed cells are characterized by a reversed circulation. As an example Figure 1 shows a satellite image of a CAOB situation which was observed on 10 March 1982. During this typical case of a CAOB, the whole Northern part of the Atlantic Ocean was covered by the characteristic organized convective cloud pattern.

Studying organized convection during CAOBs is of interest for at least two reasons: Firstly, the organization of convection may affect the vertical transport of heat and moisture from the water surface into the atmosphere. In CAOB regions, the sensible and latent heat flux near the surface can reach very large values (both may sum up to more than 1000 W m\(^{-2}\)). Thus CAOBs significantly contribute to the atmospheric energy balance. In weather forecast and climate models these fluxes are parameterized and therefore it is crucial to understand, how they are modified by organized convection. Secondly, there are still some open questions concerning basic features of organized convection during CAOBs. While in classical laboratory experiments of Rayleigh-Bénard convection the aspect ratio of hexagonal cells (the ratio of the cell diameter to its height) is about 3, values between 10 and 30 are typically observed during CAOBs. There are indications given by numerical simulations of CAOBs, that the release of heat due to condensation within the clouds is responsible for the cell broadening\(^7\). The study of the evolution of convective
structures of CAOBs with numerical models is only possible, when these models are able to resolve the main energy containing turbulence elements of the flow by their numerical grids. Due to insufficient computer resources, these so-called large-eddy simulations (LES) of CAOBs could not be satisfactorily carried out so far, because the energy containing eddies have sizes ranging from a few hundred meters (single bubbles of warm air) up to several ten kilometers (hexagonal cells). The computational domains have to be large enough to contain the hexagonal cells and must have a spatial resolution fine enough in order to resolve also the smaller scales because possible interactions between these scales may affect the cell structure. The results of previous CAOB studies are afflicted with different uncertainties: Either the model domain was too small so that one single convection cell filled the whole model domain at the end of the simulation or a coarse resolution had to
be used in order to be able to cover a larger area.

In meteorology, LES has been used since the early 1970s as a powerful tool for studying atmospheric turbulence. First investigations relating to LES in meteorology trace back to Lilly and Deardorff. LES models are known for their large demands on computer resources. An LES of a typical flow during a cold-air outbreak with well developed turbulence will need a computational domain of about 100 km$^2$ and a grid resolution of 50 to 100 m, resulting in about $10^8$ grid points per variable. In case of the LES-model employed here the demand on main memory will be in the order of 30 Gbytes (provided that one word of memory is represented by 8 bytes). While in the past such simulations were impossible due to limited computer resources, the availability of parallel and massively parallel systems now offer new opportunities.

However, to get an access to the enormous computer power of parallel systems, existing models need to be extensively modified, i.e. parallelized, before they can be used efficiently on these machines. Therefore, we had to totally rewrite an existing LES code for the use on massively parallel systems. Using this new model combined with the capacity of today’s massive parallel computers like the CRAY-T3E, we are able to use model domains which are large enough to guarantee that the evolving structures are not significantly restricted by the boundaries of the model domain. Additionally we can use grid spacings which are fine enough to resolve small scale convective structures like single up- and downdrafts with horizontal extensions down to 100 m, which may interact with the larger cells.

2 The Parallelized Large-Eddy Simulation Model PALM

The LES model applied here is specially designed for the use on massively parallel computers and carries the name PALM (parallelisiertes LES-Modell). It is based on the non-hydrostatic Boussinesq-approximated Navier-Stokes equations and contains a water cycle with cloud formation and precipitation processes. It takes into account infrared radiative cooling in cloudy conditions. The spatial derivatives are approximated by second-order
centered finite differences. Time integration is performed by using the leap-frog scheme. Lateral boundary conditions of the model are cyclic and Monin-Obukhov similarity is assumed between the surface and the first computational grid points above.

The parallelization is achieved by horizontal two-dimensional domain decomposition as it is shown in Figure 2. The total domain is divided into subdomains, which are assigned to the processor elements (PEs) – one subdomain per each PE – and each PE solves the whole set of equations on its subdomain. Communication between the PEs is realized by the message passing interface MPI. The cyclic horizontal boundary conditions are realized implicitly by creating a virtual two-dimensional topology of processors, which is shown together with the interconnecting links between the different PEs for a $3 \times 3$ grid of PEs in Figure 2.

**PALM** parallelizes very well and shows a good performance on distributed memory machines (CRAY-T3E) as well as on shared memory systems (SGI-Origin with 128 PEs). Figure 3 shows some results of our scalability tests on the CRAY-T3E of NIC. An almost linear speed up is achieved. Up to very large numbers of PEs (here 256 PEs) computational time is halved by doubling the number of PEs used. During all these test runs total communication time between the PEs does not exceed 8% of the total execution time (not shown). The largest run performed during our tests needs about 45 GBytes of main memory. Using 512 PEs of the CRAY-T3E at NIC and a computational grid of $1216 \times 1216 \times 160$ one timestep needs about 12 s of CPU-time on each PE for this large test run (this run seems to be a world record in resolution for LES models).

For a detailed model description and a detailed discussion of our test runs the reader is referred to Raasch and Schröter and to the world wide web.

---

Figure 3. Speed up as a function of the number of PEs. Lines show the results for a set of test runs using $160 \times 160 \times 64$ grid points as well as the ideal case (red line)

**http://www.muk.uni-hannover.de/~raasch/PALM-1/intro.html**

436
3 Broadening of Convective Cells During Cold Air Outbreaks

To study the broadening of mesoscale convection cells we refer to a CAOB situation observed during the ARKTIS 1991 experiment. On 8 March 1991 a transition of convection patterns was observed, as shown exemplarily in Figure 1. Details on the experiment are given by Brümmer. Two principal runs were performed, where we used the initialization according to the observed situation. For both runs the model domain covers an area of 70 km $\times$ 70 km in the horizontal and 5 km in the vertical direction with a grid spacing of 100 m resulting in $700 \times 700 \times 80$ grid points. The simulations were performed on 256 PEs on the NIC CRAY-T3E covering a period of 12.5 h. Each PE required 115 CPU-h (4.8 days) in total. The first run (RUN1) includes the whole water cycle, whereas the second run (RUN2) takes the same initial parameters but the water cycle was switched off in order to study the influence of adiabatic heat sources on cell broadening.
Figure 5 shows the power spectra of vertical velocity as a function of aspect ratio for different time levels of RUN1 and RUN2 calculated at mid boundary layer. The broadening of the dominating scales, corresponding to the mesoscale convection cells, can be seen clearly by increasing aspect ratios with increasing time. Until the end of the simulation the dominating aspect ratio increases from approximately 3.5 at 2 h and 4 h to 10 at 12.5 h. These results are in good agreement to earlier findings of other authors. In contrast the power spectra of RUN2, which was performed without considering condensation and longwave radiation processes, do not show a growing of the dominating aspect ratios. Within the whole period of the simulation the aspect ratios remain between 3 and 3.5.

Signals of the mesoscale convection cells can also be identified by visual analysis. In Figure 6 cross sections in the $x$-$y$-plane through the field of liquid water $q_l$ and the vertical velocity field $w$ at the end of RUN1 ($t = 12.5$ h) in the middle of the cloud layer are shown. The centers of the cells are represented by a homogeneous distribution of relatively large values of $q_l$. The centers are enclosed by narrow rings represented by small values of $q_l$. This kind of cellular pattern is typical for so called closed convection cells. The typical diameters of the cells are in the order of 20-30 km, which is in good agreement to the observation during ARKTIS. On the first view the $w$-field could be characterized by spatially randomly-distributed downdrafts (light areas) and updrafts (dark areas). But on the second view cellular patterns could be identified in the $w$-field, due to positive correlation with the field of liquid water. In areas of small liquid water contents predominantly downdrafts can be identified. In the center of the cells the updrafts seem to be more frequent than downdrafts. Each mesoscale convection cell is an organized conglomeration of many up-
and downdrafts. A phenomenon which could not be observed in earlier studies by Müller and Chlond\(^7\), who used much coarser grid resolutions (up to 1600 m).

During RUN2 no formation of mesoscale convection cells could be identified. The structure of the turbulent flow and the aspect ratios observed during RUN2 agrees with the ones detected in cloudless convective boundary layers. Since in RUN2 the formation of mesoscale convection cells fails to appear, we can draw the conclusion that diabatic heat sources are responsible for the existence and for the broadening of mesoscale convection cells.

In order to point out the correlation between dynamics of the flow and its thermodynamics (temperature and liquid water content) figures 7 and 8 present cross sections in the \(x\)-\(z\)-plane through a single mesoscale convection cell. Shown are isopleths of the liquid water content and of the potential temperature for \(y = 20\) km (see fig. 6). Additionally the flow dynamics are indicated by mean velocity vectors \((u, w)\) averaged over five grid points. As it should be expected from discussion above, large amounts of liquid water are in positive correlation to upwardly vertical velocities. Reaching the condensation level, water vapor condenses resulting in increasing liquid water contents (dark blue areas) and a thicker cloud layer. Due to latent heat release during the condensation process the temperature increases (red areas for \(2000 \, m < z < 3200 \, m\)). Correspondingly in areas with descending air, marked by downward pointing arrows, liquid water content and temperature decrease due to evaporation \((x \approx 15\text{-}20\, km\)). Together the areas of ascending air and areas of descending air form the circulation of the mesoscale convection cells.

First attempts to answer the question about the physical processes which lead to the broadening of convective cells were made by Dörnbrack\(^4\) as well as by Müller and Chlond\(^7\). Currently, we try to verify their hypothesis using the results of our simulations. Figure 8 shows that within the cloud layer broader coherent warm and cold areas with less heterogeneity appear. These homogenous areas are invoked by stronger exchange processes which are forced by the diabatic heat sources. As a consequence the dynamic of the flow is directly affected by these broader structures.

## 4 Conclusions

For examinations of turbulent flows in the atmospheric boundary layer the parallelized large-eddy simulation model PALM was developed for the use on massively parallel computers. Problems like the simulation of interactions between small scale and (larger scale) organized convection can be tackled now. Within these studies of organized cellular convection a simulation was performed for the first time using a turbulence resolving numerical model with a fine grid spacing combined with large horizontal extensions of the model domain. In this way we generated a dataset to study the broadening of mesoscale cellular convection resolving turbulent eddies covering a bandwidth of wavenumbers which is unmatched so far. The uncertainties of earlier investigations, using undersized model domains\(^4\) or using a coarse grid spacing\(^7\), were removed by the results of our simulations. Now it is definitely shown that diabatic heat sources like latent heat release due to condensation processes and longwave radiative cooling at the top of the cloud layer are responsible for the existence and for the broadening of mesoscale convections cells.
Figure 7. Contour plots of the liquid water content and a vector plot of the averaged velocity field for a vertical $x$-$z$ cross section at $y = 20$ km at 12.5 h.

Figure 8. Contour plots of the potential temperature fluctuations and a vector plot of the averaged velocity field for a vertical $x$-$z$ cross section at $y = 20$ km at 12.5 h. Temperature fluctuations are obtained by subtracting the horizontal average value calculated for each vertical grid-plane.
References


Sensitivity Studies with Numerical Models of Medium Resolution of the Atlantic Ocean

Jens-Olaf Beismann¹ and René Redler²

¹ Institut für Meereskunde, Universität Kiel
Düsternbrooker Weg 20, 24105 Kiel, Germany
E-mail: jobeismann@ifm.uni-kiel.de

² C&C Research Laboratories, NEC Europe Ltd.
Rathausallee 10, 53757 Sankt Augustin, Germany
(previously at GMD SCAI, Schloss Birlinghoven Sankt Augustin)
E-mail: rene.redler@ccrl-nece.de

The numerical simulation of the uptake and redistribution of chlorofluorocarbons (CFC) provides a powerful tool for studying the mechanisms of water mass formation and spreading in ocean circulation models. Uptake of atmospheric trace gases of which CO₂ is another prominent example is strongly influenced by oceanic physical variability at spatial scales below 100 km. A realistic representation of this uptake in numerical models is essential for future climate studies.

In the present study we use dissolved CFCs as a tracer to analyze the renewal of North Atlantic Deep Water and its pathways in different prognostic 3-d models of the Atlantic. Results of several long-term model integrations are presented to discuss the physical processes determining the temporal evolution of the 3-d CFC distribution in the deep ocean. Effects of different subgrid scale parameterizations are discussed. In particular we address the southward transport of CFC-enriched water with the Deep Western Boundary Current.

1 Introduction

The North Atlantic Deep Water (NADW) with its source water masses is a major component of the global oceanic thermohaline circulation. Understanding the formation processes of the different water mass components of NADW and the uptake of anthropogenic trace gases going along with this is still a key issue when investigating the global circulation, the coupled ocean-atmosphere climate system and the ocean in particular as one of the major sinks for anthropogenic CO₂. As often pointed out in the literature such processes take place on a relatively small scale compared to the basin wide circulation. Formation rates of NADW turn out to be highly variable in time from interdecadal timescales down to days. Prominent examples of these are the interdecadal variability of Labrador Sea convection, the Denmark Strait Overflow or the Iceland Faroe Overflow.

Due to the sparse coverage in both space and time of field measurements the question of how these processes are related to the climate system and how they interact with each other cannot be answered from observations alone. Numerical models are valuable and widely accepted tools to gain deeper understanding and insight into the natural phenomena. To simulate the global circulation with a certain degree of realism, sufficiently high resolution in both space and time is required.

Tracing the uptake and redistribution of chlorofluorocarbons (CFC) provides a powerful tool to study the ocean circulation and the mechanisms of water mass formation, both in
the real ocean and in numerical simulations\textsuperscript{1,2}. With respect to model results the intercomparison of field measurements and modeled CFC distributions serves as a useful means for assessing the models capability of detecting and describing processes in key regions of water mass formation and their propagation in current regimes like the Deep Western Boundary Current (DWBC). In combination with observed CFC distributions these tracers thus help to determine whether a model correctly simulates NADW formation rates, pathways and mean current speeds\textsuperscript{3}.

CFCs have frequently been used for this purpose (cf. the recent review by England and Maier-Reimer\textsuperscript{4}), mostly with models of rather coarse horizontal resolution of about 2° to 4° in latitude and longitude. In comparing results of these models with observations the question arises which aspects of the model solutions can be assumed to be fairly independent of the model resolution and/or model configuration, and which features are likely to change with future generations of models of higher resolution.

Another issue of very coarse resolution models addresses the high diffusivity and its impact on the water mass propagation and distribution. Ideally, the sub-grid scale parameterization employed by any numerical model should not affect the propagation rate of a signal when compared with higher resolution models and observations.

A series of systematic studies on the impact of various mixing schemes has already been published by England and Hirst\textsuperscript{5} and England and Holloway\textsuperscript{3}. While all the numerical experiments presented in these studies show reasonable production rates and southward transports of NADW, this must go along with low velocities at individual grid points located in the DWBC region: The DWBC as a narrow current along the continental shelf is hardly resolved with coarse resolution models, and individual velocities represent an average over the whole volume of the respective grid box which has to represent both a core of the very fast DWBC and weak currents further offshore. One of the logical consequences is an underestimation of ventilation rates of water masses directly influenced by NADW.

Uptake of CO\textsubscript{2} (another prominent anthropogenic trace gas) by the ocean is a key issue for understanding and predicting the behavior of the climate system. In contrast to CO\textsubscript{2}, CFCs are easier to include in numerical models because of their chemical inertness and their well known timeseries of nearly homogeneous atmospheric concentrations over the last decades. Detailed knowledge about how the representation of gas uptake affects calculated inventories are crucial for a correct interpretation of scenarios addressing the uptake e.g. of anthropogenic CO\textsubscript{2}. Previous authors\textsuperscript{6,4} have explored the sensitivity of the simulated air-sea gas fluxes to parameterizations of the gas transfer velocity, but since they restricted themselves mainly to the comparison of tracer concentrations on certain ocean sections, very little can be found in the literature concerning the influence of gas flux parameterizations on oceanic inventories of dissolved trace gases. Furthermore the numerical models used in the above mentioned studies were of rather coarse resolution (typically of the order of 4° ), leading to large uncertainties in the representation of key processes governing the trace gas uptake like e.g. convection in the Labrador Sea. It is thus possible that these earlier findings might change when models with finer grid spacing are employed.

In section 2 we introduce the technical details of the numerical model we employed, and we present some results of our experiments in section 3.
2 Experimental Design

The FLAME (Family of Linked Atlantic Model Experiments) code we use for our model studies is based on the Modular Ocean Model (MOM) from the Geophysical Fluid Dynamics Laboratory in Princeton.

Technical details addressing the model configuration and parallelization strategy have already been published at the beginning of this project. For further details the reader is kindly referred to this document accessible via


2.1 Physical Setup

Since the physical processes under consideration are mainly restricted to the Atlantic and not much influenced by the rest of the world ocean we limit ourselves to model the Atlantic only. Our model (Figure 1) area covers the region of the Atlantic Ocean and is situated between 70°N and 70°S with open boundaries across the Antarctic Circumpolar Current in the Drake Passage (70°W) and south of Africa at 30°E following the approach of Stevens.

The northern and southern boundaries are closed. To include the effect of water mass formation in the Greenland Sea a restoring area taken from the DYNAMO experiments is used for all our case studies.

The horizontal resolution of the locally rectangular grid is 4/3° in longitude and 4/3° × cos(φ) in latitude (φ) resulting in a mesh size of 148 km at the Equator decreasing to 50 km at the subpolar boundaries. The vertical is discretized in 45 levels, with a spacing of 10 m in the uppermost level and a smooth increase to 250 m at 2500 m depth. Below 2500 m the vertical grid box thickness is a constant 250 m up to a maximum depth of 5500 m.

MOM is a Fortran code based on the so called “primitive equations” which are derived from the Reynolds-averaged Navier-Stokes equations under various assumptions, like the Boussinesq, spherical and hydrostatic approximation. From these equations MOM calculates prognostic values of the horizontal velocity, potential temperature (Θ) and salinity (S) distribution. Density and pressure are constructed from the Θ and S fields. In addition to that, we calculate the distribution of two species of chlorofluorocarbons (CFC-11 and CFC-12) as supplementary tracers. The equations are discretized on a regular three-dimensional grid with depth as vertical coordinate.

For solving the momentum equations, the horizontal velocities are split up into a vertical average and its deviations. The vertically averaged transport can be expressed by a 2-dimensional stream function which is described by a Poisson equation. The latter is solved using an iterative conjugate gradient method. Additional constraints are needed to keep the values around the islands constant during every iteration step.

Due to the hydrostatic approximation the vertical component of the momentum equation degenerates, and the vertical transport due to hydrostatic instabilities (convection) has to be parameterized. Vertical velocities are diagnosed from the divergent horizontal velocity field.

The remaining 2nd order non-linear partial differential equations for Θ, S and the vertical shear of horizontal momentum are solved on regular 3-D grids using the finite vol-

\footnote{http://www.ifm.uni-kiel.de/fb/fbl/tm/research/FLAME/index.html}
Figure 1. Sea surface temperature as simulated by the FLAME 4/3° Atlantic model. Horizontal lines indicate the data partitioning. By applying 10 PEs each processor has to work on 15 rows. Unequal spacing between horizontal lines are due to the map projection used.

Volume/difference approach. Explicit time stepping is used. Central differences are applied in space. Time discretization is done with central differences (so called “leap-frog” scheme) as well but replaced at regular intervals by a backward Euler time step in order to damp the computational mode.

For horizontal diffusion of tracers, e.g. $\Theta$, $S$ and CFCs we use mixing along isopycnal surfaces\textsuperscript{12,13}. This implicit approach requires the inversion of a tridiagonal matrix accompanied by a rotation of the mixing tensor into the direction of surfaces of constant density. Some experiments employ the eddy induced tracer advection parameterization according
to Gent and McWilliams\textsuperscript{14}, in some cases with horizontal background diffusion. Both coefficients, isopycnal diffusivity and thickness diffusivity for the Gent&McWilliams parameterization (GM90, hereafter), are set to $2 \cdot 10^7$ cm$^2$/s decaying with depth to $0.5 \cdot 10^7$ cm$^2$/s below 4000 m. Along boundaries the coefficients are kept zero. Constant coefficients are chosen for horizontal viscosity of $10^8 \cdot \cos \phi$ cm$^2$/s.

Simulating the oceanic uptake and spreading of anthropogenic tracers requires an appropriate representation of the flow of dense water masses across shallow sills which in our model is achieved by a bottom boundary layer parameterization according to Beckmann and Dösser\textsuperscript{15}.

The atmospheric climatological forcing applied to our experiments is based on a 3-year monthly mean climatology of ECMWF (European Center for Medium Range Weather Forecast) analysis (1986 to 1988)\textsuperscript{16}, and relaxation to surface salinity given by Boyer and Levitus\textsuperscript{17}. The thermal boundary condition is described by Barnier et al.\textsuperscript{16} and uses linearized bulk formulas to define a model dependent air-sea net heat flux as a relaxation to an equivalent sea surface temperature. The wind stress is also derived from the same ECMWF analysis\textsuperscript{18}.

The ECMWF wind stresses, atmospheric equilibrium temperature, friction velocity and sea surface salinity fields were all converted to pseudo fields as proposed by Killworth\textsuperscript{19}, in order to achieve correct monthly means. For several experiments we use monthly mean heat flux and wind stress anomalies (with respect to a long term mean) taken from the NCEP/NCAR reanalysis for years 1958 to 1996\textsuperscript{20} which had been added to our climatology\textsuperscript{21}.

All individual experiments are initialized with January $\Theta$ and S fields and are spun up from a state of rest for an integration period of 20 model years followed by an analysis phase of 47 years to hindcast the time span from 1950 to 1996.

### 2.2 Computational Requirements

Our model domain encompasses 533250 grid points. With 7 prognostic variables to calculate and a time step of 1 hour (8760 time steps per model year) we need about 37.5 hours CPU time to simulate 1 model year on a Cray T3E-1200. Running the executable on 15 PEs, this accounts for roughly 2.5 hours elapsed time and a total of 165 hours elapsed time is required to complete a single experiment for a full 66 year period. These requirements allow us to perform the sensitivity studies described above by exploring various parameterizations of subgrid scale physics.

The disk requirements for this model are moderate: For each individual job (1 year simulation) the diagnostic output sums up to 1.5 Gbyte which is reduced to 300 Mbyte in a separate postprocessing phase. For each individual experiment the total amount of data for a full 67 year integration to be stored permanently for further analysis sums up to approximately 20 Gbyte.

It is tempting to compare results of the configurations described so far with a model of somewhat higher resolution in space and time which is able to resolve the sub-grid scale physics explicitly to a wider extent.

The numerical model we designed for this purpose ($1/3^{\circ} \times 1/3^{\circ} \cos(\phi)$ horizontal resolution) encompasses 8.3 million grid points. Using a time step of 20 minutes (25920 time steps per model year) the required CPU time to perform 1 year of simulation sums...
up to 1080 hours which (running the executable on 120 PEs on a T3E-1200) results in an elapsed time of 9 hours. Due to limitations of the queue system the jobs had to be split up into 3 spanning the time frame of 4 months each. The temporary disk space required for the output of each individual job then adds up to 4.2 Gbyte. For each model year 3.9 Gbyte of pre-analyzed data have to be stored permanently for later analysis. Furthermore a permanent disk space of 1 Gbyte is required to provide the necessary input data to the model.

Due to limitations in CPU time and disk space, performing a high resolution run comparable to the experimental setup described above would have taken 1 year real time on the given system and environment without allowing us to perform any further sensitivity studies beside this one to compare with during this time.  

3 Simulation Results

Basin scale numerical ocean models do not resolve mesoscale eddies due to their coarse resolution. Instead they have to rely on some parameterization to represent the mixing effect of eddies on tracer distributions.

Here we will discuss briefly the CFC-12 distribution in the upper part of NADW (uNADW) from a series of experiments (B1.7, B1.9 and B1.N4) that have been conducted on the CRAY-T3E. In our 4/3° configuration different mixing parameterizations have been employed. For the qualitative comparison of tracer distributions we use CFC-12 concentrations on an isopycnal surface (surface of constant density) corresponding to the tracer maximum in the uNADW, usually \( \rho_0 = 1027.78 \text{ kg/m}^3 \). \( \rho_0 \) is the density of sea water for a given temperature and salt content referenced to the sea surface.

Numerical models at this resolution suffer from unrealistic uNADW propagation pathways marked by the CFC signal when simple mixing parameterizations are used. In general the simulated CFC distribution is much too diffusive compared to field observations, leading e.g. to a broad DWBC instead of a narrow and concentrated boundary current. In addition, the CFC signal in the eastern parts of the subtropical and subpolar Atlantic reaches too far to the south (Figure 2, upper right panel).

The introduction of the GM90 parameterization of eddy-induced tracer advection improves our simulation in some aspects (Figure 2, upper left panel): The eastward outflow from the Labrador Sea is reduced, leading to a less intense CFC propagation in the eastern basin of the North Atlantic. The southernmost extension of a given isoline can now clearly be found in the western basin, whereas isolines have a much more zonal direction in the experiment without GM90. At the same time, the already sluggish DWBC in experiment B1.9 has slowed down in the GM90 run (B1.7), and the associated tracer signal cannot make its way around the Grand Banks into the subtropical Atlantic. The core of the CFC-12 maximum is not connected to a boundary current structure, but propagates southward in the interior of the basin. Also, the tracer tongue does not reach as far across the equator and into the South Atlantic as in experiment B1.9. The effect of the GM90 parameterization to mimic the conversion of available potential energy to eddy kinetic energy by flattening

An integration of the 1/3° (2 \times 50 years) has been started 2 months ago on the SX5 at HLR Stuttgart. At the time of writing this paper the first phase of the integration was completed. Numerical results are currently under investigation.
isopycnal surfaces and modifying the tracer advection speed has important consequences for our simulated DWBC: underneath the Gulf Stream, the slope of the isopycnal surfaces is so weak that it is not possible to reverse the flow direction with increasing depth (via thermal wind). Therefore, NADW outflowing from the subpolar North Atlantic to the south has to pass on the off-shore side of the Gulf Stream which leads to the tracer distribution described above. Tracer observations on meridional sections southwest of the Grand Banks\textsuperscript{22,23} show that in addition to the uNADW CFC maximum on the continental shelf, local maxima of lower concentration exist in the interior. These have been attributed to recirculation cells in the DWBC, but they do not support the single interior maximum simulated in our GM90 experiment.

Holloway\textsuperscript{24} proposed a parameterization that tries to represent the interaction of (un-
resolved) mesoscale eddies with bottom topography. Based on statistical mechanics, this parameterization induces a cyclonic (counter-clockwise on the northern hemisphere) circulation along topographic slopes even if the ocean model is not forced by atmospheric fluxes of momentum or buoyancy. Our experiment B1.N4 (Figure 2, lower panel) includes this parameterization, and we can identify important differences to the simulations with the more “classic” eddy parameterizations: A Deep Western Boundary Current has been established all along the American continental shelf where the fastest tracer propagation can be found. The southward extension of the tracer tongue reaches further across the equator than in experiments B1.7 and B1.9, and the unrealistic eastward spreading in the subpolar gyre has been further reduced. While these findings underline the role of bottom topography for the deep circulation, a more quantitative assessment of experiment B1.N4 will be necessary to explore implications for the integrated tracer uptake.

4 Conclusions

A model hierarchy of the Atlantic has been used to study formation processes of NADW as one of the most prominent water masses for the global ocean circulation which can easily be detected due to its high content of CFC. To complement previous studies of this kind performed with models of rather coarse resolution, we employed a model with a horizontal resolution of $4/3\times 4/3\cos(\phi)$ encompassing the Atlantic in the latitudinal range $70^\circ$N to $70^\circ$S. Sensitivity studies with respect to various mixing parameterizations are investigated.

Although not addressed in the very details in this paper, a first comparison with observed CFC data showed that our model is well able to capture features like the surface concentration, ventilation of intermediate layers in the subpolar North Atlantic and the formation of the DWBC. Nevertheless when looking at important details of the simulation it turns out that e.g. the velocities in the DWBC are too weak in models of this resolution and the CFC signal in general is too diffusive. From this point of view higher resolution is necessary in order to achieve realistic simulations.

This may have important implications for modeling studies of the ocean’s role in the global carbon cycle and for climate change scenarios which, because of the long integration periods have to rely on very coarse grid ocean models. The modeled global circulation pattern as well as renewal rates of water masses in the deep ocean basins may be subject to changes once the hardware is available to apply models of higher resolution in long time integrations for climate scenarios.

Acknowledgments

The authors gratefully acknowledge the work of Klaus Ketelsen, SGI Munich for his work on optimizing the code for the T3E. The work of the FLAME group is acknowledged for its contribution at various stages of the project. Model results have been analyzed using the Ferret plotting package, kindly provided by Steve Hankin and his team from NOAA/PMEL. The sensitivity studies have been carried on the Cray T3E-1200 at the John-von-Neumann Computing Center Jülich, while one of the authors (RR) was working at the GMD (now Fraunhofer) Institute for Algorithms and Scientific Computing, Sankt Augustin, Germany.
References

18. L. Siefridt. *Validation des données ERS-1 et des flux de surface du CEPMMT dans le*


Air quality models base on mass balances described by systems of time-dependent, three-dimensional advection-diffusion-reaction equations. To date, one limitation of schemes for the numerical solution of such systems has been their inability to solve equations both quickly and with a high accuracy in multiple grid cell models. This requires the use of fast parallel computers. Multiblock grid techniques and implicit-explicit (IMEX) time integration schemes are suited to take benefit from the parallel architecture. A parallel version of the multiscale chemistry-transport code MUSCAT is presented which is based on these techniques. The meteorological fields are generated simultaneously by the non-hydrostatic meteorological model LM. Both codes run in parallel mode on a predefined number of processors and exchange informations by an implemented coupler interface. The ability and performance of the model system are discussed.

1 Introduction

Atmospheric chemistry-transport models are useful tools for the understanding of pollutant dynamics in the atmosphere. Such air quality models are numerically expensive in terms of computing time. This is due to the fact that their resulting systems of ordinary differential equations (ODE) are nonlinear, highly coupled and extremely stiff. In chemical terms, a stiff system occurs when the lifetime of some species are many orders of magnitude smaller than the lifetime of other species. To illustrate the magnitude of the problem, we note that the stiffness ratio for a typical tropospheric photo-chemistry system is usually greater than $10^{10}$. Because explicit ODE solvers require numerous short time steps in order to maintain stability, most current techniques solve stiff ODEs implicitly or by implicit-explicit schemes.

For our air quality studies, the chemistry-transport-model MUSCAT (MUltiScale Chemistry Aerosol Transport) has been utilized. The code is parallelized and is tested on several computer systems. Presently, MUSCAT has an online-coupling to the parallel, non-hydrostatic meteorological code LM which is the operational regional forecast model of the German Weather Service. Both parallel codes work on their own predefined fraction of the available processors and have their own separate time step size control, see Fig. 1. The coupling scheme simultaneously provides time-averaged wind fields and time-interpolated values of other meteorological fields (vertical exchange coefficient, temperature, humidity, density). Coupling between meteorology and chemistry-transport takes place at each horizontal advection time step only. This scheme can easily be extended to the coupling of MUSCAT with other mesoscale meteorological codes. In the previous version of our model system, it was used for the coupling of the parallel MUSCAT to the mesoscale model METRAS. This meteorological model is also a non-hydrostatic model developed for regional applications. In this version, METRAS runs only on one processor.
In MUSCAT, the horizontal grid is subdivided into so-called “blocks”. The code is parallelized by distributing these blocks on the available processors. This may lead to load imbalances, since each block has its own time step size control defined by the implicit time integrator. Therefore, well-suited dynamic load balancing is proposed and investigated. In Sec. 5 the benefit of a developed strategy is discussed for a "Berlioz" ozone scenario. Finally, the ability of MUSCAT for environmental studies is demonstrated for the "Black Triangle" area.

2 The Chemistry-Transport Code MUSCAT

Air quality models base on mass balances described by systems of time-dependent, three-dimensional advection-diffusion-reaction equations

\[
\frac{\partial y}{\partial t} + \frac{\partial}{\partial x_1}(u_1 y) + \frac{\partial}{\partial x_2}(u_2 y) + \frac{\partial}{\partial x_3}(u_3 y) = \frac{\partial}{\partial x_3}(\rho K z \frac{\partial y}{\partial x_3}) + Q + R(y). \tag{1}
\]

\(y\) denotes a vector of species concentrations or aerosol characteristics which will be predicted and \(\rho\) is the density of the air. The wind field \((u_1, u_2, u_3)\) and the vertical diffusion coefficient \(K_z\) are computed simultaneously by a meteorological model. The conversion term \(R\) represents the atmospheric chemical reactions and/or the aerosol-dynamical processes. \(Q\) denotes prescribed time-dependent emissions.

**Multiblock Grid.** In MUSCAT a static grid nesting technique\(^{12,6}\) is implemented. The horizontal grid is subdivided into so-called “blocks”. Different resolutions can be used for individual subdomains in the multiblock approach, see Fig. 2. This allows fine resolution for the description of the dispersion in urban regions and around large point sources. The multiblock structure originates from dividing an equidistant horizontal grid (usually the meteorological grid) into rectangular blocks of different size. By means of doubling or halving the refinement level, each block can be coarsened or refined separately. This is done on condition that the refinements of neighbouring blocks differ by one level at the
The maximum size of the already refined or coarsened blocks is limited by a given maximum number of columns. The vertical grid is the same as in the meteorological model.

The spatial discretization is performed by a finite-volume scheme on a staggered grid. We implemented a first order upwind and a biased upwind third order procedure with additional limiting[^1]. This scheme has to be applied to non-equidistant stencils which occur at the interface of blocks with different resolutions[^6]. Fig. 3 illustrates the choice of the grid values involved in the interpolation formula for the two different upstream directions. In this case the grid cell has five cell wall interfaces.

**Time Integration.** For the integration in time of the spatially discretized equation (1) we apply an IMEX scheme[^12]. This scheme uses explicit second order Runge-Kutta methods for the integration of the horizontal advection and an implicit method for the rest. The fluxes resulting from the horizontal advection are defined as a linear combination of the fluxes from the current and previous stages of the Runge-Kutta method. These horizontal fluxes are treated as "artificial" sources within the implicit integration. Within the implicit integration, the stiff chemistry and all vertical transport processes (turbulent diffusion, advection, deposition) are integrated in a coupled manner by the second order BDF method. We apply a modification of the code LSODE[^2] with a special linear system solver. The error control can lead to several implicit time steps per one explicit step. Furthermore, different implicit step sizes may be generated in different blocks. The "large" explicit time step is chosen as a fraction of the CFL number. Higher order accuracy and stability conditions for this class of IMEX schemes are investigated in Knoth and Wolke[^5].

**Gas Phase Chemistry.** The chemical reaction systems are given in ASCII data files in a notation that is easily understandable. For the task of reading and interpreting these chemical data we have developed a preprocessor. Contained in its output file are all data structures required for the computation of the chemical term $R(y)$ and the corresponding Jacobian. Changes within the chemical mechanism or the replacement of the whole chemistry can be performed in a simple and comprehensive way. Several gas phase mechanisms (e.g. RACM, RADM2, CBM IV) are used successfully in 3D case studies. Time resolved anthropogenic emissions are included in the model via point, area and line sources. It is distinguished between several emitting groups. Biogenic emissions are parameterized in terms of land use type, temperature and radiation.
Aerosol Dynamics. For simulation of aerosol-dynamical processes the model MADMAcS 11 was included in MUSCAT. The particle size distribution and the aerosol-dynamical processes (condensation, coagulation, sedimentation and deposition) are described using the modal technique. The mass fractions of all particles within one mode are assumed to be identical. Particle size distribution changes owing to various mechanisms, which are divided into external processes like particle transport by convection and diffusion, deposition and sedimentation as well as internal processes like condensation, evaporation, homogeneous nucleation and coagulation.

Parallelization. Our parallelization approach is based on the distribution of blocks among the processors. Inter-processor communication is realized by means of MPI. The exchange of boundary data is organized as follows. Since the implicit integration does not treat horizontal processes, it can be processed in each column separately, using its own time step size control. An exchange of data over block boundaries is necessary only once during each Runge-Kutta substep. Each block needs the concentration values in one or two cell rows of its neighbours, according to the order of the advection scheme. The implementation of the boundary exchange is not straightforward because of the different resolutions of the blocks. The possibilities of one cell being assigned to two neighbouring cells or of two cells receiving the same value must be taken into account. We apply the technique of "extended arrays": the blocks use additional boundary stripes on which incoming data of neighbouring blocks can be stored. Hence, each processor only needs memory for the data of blocks that are assigned to it.
3 Online-Coupling to the Parallel Meteorological Model LM

In Wolke and Knoth\textsuperscript{12} an online-coupling technique between the chemistry-transport code MUSCAT and the mesoscale non-hydrostatic meteorological model METRAS (Schlünzen, 1990) is proposed. The meteorological and the chemistry-transport algorithms have their own separate time step size control. The coupling procedure is adapted to the applied IMEX schemes in the chemistry-transport code. All meteorological fields are given with respect to the equidistant horizontal meteorological grid. They have to be averaged or interpolated from the base grid into the block-structured chemistry-transport grid with different resolutions. The velocity field is supplied by its normal components on the faces of each grid cell, and their corresponding contravariant mass flux components fulfill a discrete version of the continuity equation in each grid cell. This property has to be preserved for refined and coarsened cells. Because only the mass flux components are needed for the advective transport of a scalar, the necessary interpolation is carried out for these values. The interpolation is done recursively starting from the meteorological level.

The same approach is applied to the coupling with the meteorological driver LM. Since LM solves a compressible version of the model equations an additional adjustment of the meteorological data is necessary. The velocity components are projected such that a discrete version of the continuity equation is satisfied. The main task of this projection is the solution of an elliptic equation by a preconditioned conjugate gradient method. This is also done in parallel on the LM processors. The projected wind fields and the other meteorological data are gathered by one of the LM processors. This processor communicates directly with each of the MUSCAT processors.

4 Parallel Performance and Load Balancing

In this section, the parallel performance and the run time behaviour of the model system is discussed for a typical meteorological "summer smog" situation ("Berlioz ozone scenario"). In the following we focus on the computational aspects. A discussion of the simulation from an environmental point of view would go beyond the scope of this paper.

The model area covers approximately 640 km x 640 km with a multiscale grid with resolutions of about 4 km–8 km–16 km, see Fig. 4. In the vertical direction the model domain is divided into 19 non-equidistant layers between the surface and a height of approximately 3000 m. The used chemical mechanism RACM\textsuperscript{10} considers 73 species and 237 reactions. Time-dependent emission data from both point and area sources are taken into account.

All tests are run on a Cray T3E with various numbers of processors. In all cases, 20\% of the available processors are used for the meteorological code LM, the others for MUSCAT. The equidistant meteorological grid is decomposed into several subdomains whose number corresponds to the number of LM processors. The MUSCAT runs are performed with two different prescribed error tolerances $Tol$ of the BDF integrator.

**Dynamical Load Balancing.** Consider a static partition where the blocks are distributed between the processors only once at the beginning of the program’s run time. Here, we use the number of horizontal cells (i.e., of columns) as measure of the work load of the respective block. Therefore, the total number of horizontal cells of each processor is to be balanced. This is achieved by the grid-partitioning tool ParMETIS\textsuperscript{4}. It optimizes
both the balance of columns and the “edge cut”, i.e., it takes care for short inter-processor border lines.

In order to improve the load balance, techniques allowing for redistribution of blocks have been implemented. A block’s work load is estimated using the numbers of Jacobian and function evaluations applied during a past time period. According to the work loads of the blocks, ParMETIS searches for a better distribution, besides minimizing the movements of blocks. The communication required for the exchange of block data can be done by means of similar strategies as for the boundary exchange.

Performance and Scalability. In the three cases, only the work load of the MUSCAT processors is changed by the step size control and by the dynamic load balancing procedure. The run time behaviour of the LM processors differs only in their idle times (see also Fig. 6). The parallel performance of the model runs is presented in Fig. 5. In the figure with CPU times, a logarithmic scale is used for the “number of processor” axis. Higher accuracy requirements increase the work load needed for the implicit integration in MUSCAT significantly. By using dynamic load balancing the CPU times can be reduced by a factor of about 0.8–0.9 in comparison to the static grid decomposition. The speed-up is defined with respect to the runs with 5 processors. For Tol = 10^{-3} we take the dynamic case as reference case. An optimal speed up is reached if the CPU curves show a linear

Figure 5. CPU times and speed-up on the Cray T3E.

Figure 6. Analysis of the required CPU time for runs with 40 processors.
behaviour with slope one. The measured speed-up values show that the code is scalable up to 80 processors.

The computational behaviour for the 40 processor run is presented in more detail in Fig. 6. In all cases the amount for coupling is negligible. As expected, the idle times of the LM processors growth if the work load of the MUSCAT processors caused by tighter tolerances is increased. Therefore, the number of LM processors can be reduced especially in the runs with higher accuracy. However, the basic strategy consists in the avoidance of idle times on the MUSCAT processors side. The costs for the projection of the wind fields are comparable with that of the LM computations. But it seems that this work load can be reduced by more suitable termination criteria of the cg-method.

5 Air Quality Applications for the "Black Triangle Area"

The "Black Triangle" area has been one of the most polluted areas in middle Europe. Changes in the political situation have altered the emission situation as well and, therefore, also air quality. These changes in air pollution have been investigated and monitored using the previous version METRAS–MUSCAT of the model system. Relevant limit values for protection of human health defined or suggested in directives by the European Commission were checked in worst case scenarios for the year 2005. The species sulfur dioxide, ozone, and aerosol particles have been studied. In the following only a summary of the sulfur dioxide and aerosol particle results is presented.

The model area covers Southern Saxony, Northern Bohemia and parts of Western Poland. As horizontal resolution, 2 km has been chosen. In the vertical, a non-equidistant grid with 27 layers has been used (model top at about 8 km height).

The model results have been evaluated for the general pollutant situation for past (1996) and future (2005) emission scenarios as well as in terms of the contributions of several emission source types to the total pollution. For the future scenarios it has been checked if the relevant EU guidelines will possibly be met or under which conditions they can be fulfilled. The anthropogenic emission data base refers to 1996, later scenarios have been estimated using trend factors and information about developments in industry. The data are sorted after emission source types such as households, industry, traffic, and power plants.

Sulfur Dioxide. The simulations have been carried out for meteorological conditions where high values of SO2 concentrations typically occur. In the example below winter high
pressure conditions with moderate winds from the east, ground temperatures of about -6°C and a strong inversion have been prescribed. Sulfur dioxide is mainly emitted by large power plants (point sources), smaller industrial complexes and coal fires from households. Emissions have been reduced considerably in the last years due to desulfurization and changes in heating systems. The most striking improvement has taken place between 1996 and 1998.

In 1996 high SO2 concentrations mainly originate from emissions of large power plants in Northern Bohemia. Large amounts of SO2 can be transported over wide distances and also into Saxony. Due to the complex topography of the mountain ridge "Erzgebirge" and the position of the major sources, some regions e.g. around Annaberg are more affected than others. Households and small industrial complexes contribute to SO2 pollution to a lesser extent and only very locally around the source areas. Extensive emission reduction particularly at the power plants lead to a major improvement in air quality especially between 1996 and 1998 (not shown).

The relevant EU guideline prescribes for SO2 a 24-hour limit value of 125 $\mu g m^{-3}$ for the protection of human health, that should not be exceeded more than three days times a calendar year. The simulations for the year 2005 show, that even for assumed worst case meteorological conditions this limit value will be met. So, the pollution due to SO2 should not be of importance for human health in this region for the future.

**Aerosols.** Particles smaller than PM10 are the relevant part of particulate matter in terms of human health. Recently an EU-guideline for a limit value exists for the mass concentration of the PM 10 fraction only. This part of the total particle emissions has been taken into account in these studies. The resulting amounts have then been distributed into two fractions with two different mean radii as listed in Fig. 9. Particle concentrations in the lowest model layer show similar distributions as sulfur dioxide (winter conditions, winds from the southeast), compare Figs. 11 and 8. From predictions for 2005 it can be concluded that the relevant EU-limit value for the protection of human health ($50 \mu g m^{-3}$ 8-hour mean value, not more than 35 times a year) can possibly be met even for the assumed worst case meteorological conditions. Some small problems may exist further on near the large power plants in Northern Bohemia.
<table>
<thead>
<tr>
<th>Emittance group</th>
<th>PM10 mass fraction of emitted particles [%]</th>
<th>% of mass fraction of particles with mean radius of 0.3 µm</th>
<th>% of mass fraction of particles with mean radius of 7.5 µm</th>
</tr>
</thead>
<tbody>
<tr>
<td>power plants</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a) brown coal</td>
<td>95</td>
<td>80</td>
<td>15</td>
</tr>
<tr>
<td>b) oil</td>
<td>75</td>
<td>60</td>
<td>15</td>
</tr>
<tr>
<td>c) gas</td>
<td>100</td>
<td>100</td>
<td>-</td>
</tr>
<tr>
<td>other point sources</td>
<td>75</td>
<td>45</td>
<td>30</td>
</tr>
<tr>
<td>households</td>
<td>90</td>
<td>80</td>
<td>10</td>
</tr>
<tr>
<td>small consumers</td>
<td>75</td>
<td>45</td>
<td>30</td>
</tr>
<tr>
<td>traffic</td>
<td>100</td>
<td>100</td>
<td>-</td>
</tr>
</tbody>
</table>

Figure 9. Splitting of the PM10 emission inventory.

Figure 10. PM10 emissions 1996.

Conclusions and Outlook

The model system LM–MUSCAT is a powerful instrument for environmental studies. The parallel efficiency is encouraging. In near future, the treatment of cloud chemistry processes will lead to more complex codes. It is well-known that such multiphase processes increases the computational requirements significantly. Furthermore, clouds are dynamic objects with high spatial and temporal variability. Therefore, efficient dynamic load balancing strategies seems to be more important for these simulations.

Acknowledgements

The work was supported by the NIC Jülich, the DFG and the Ministry for Environment of Saxony. Furthermore, we thank the ZHR Dresden and the DWD Offenbach for good cooperation.
Figure 11. Maximal 8-hour-mean PM 10 concentrations near the ground, 25 µg m$^{-3}$ background concentration of PM 10.

References


Both nuclear and atomic physics are of course vast subjects, and the research done at NIC supercomputers can cover only some narrow aspects. Indeed there were four major projects in this field, all dealing with fundamental questions. All four dealt with few-body problems and had some connections to particle physics.

Two projects were concerned with few-nucleon systems, i.e. with the structures and dynamics of the lightest nuclei, with $A = 3$ to 12. The challenge here is that neither the precise basic interaction nor any exact scheme to solve the equations of motion are known. Thus any theory must check carefully whether discrepancies with experiment are due to numerical difficulties or due to imperfections in the interaction ansatz. One main problem here is that 3-nucleon forces are known to be present, mainly because of relativistic and other degrees of freedom (e.g. baryon resonances) which cannot be treated explicitly in a non-relativistic Schrödinger equation. Moreover, while the Schrödinger equation for three nucleons can still be solved numerically with high precision, this is no longer true for more than three nucleons.

In the only paper on these problems included in this volume, W. Glöckle and colleagues discuss mainly three-baryon systems from this point of view. In addition to the “classic” problem of scattering in three-nucleon systems, where a wealth of precise experimental results is available, they also discuss hypernuclei and scattering processes involving photons. This paper shows impressively how high-performance computing has led to important recent progresses, but points also to difficulties which will keep researchers busy for the next few years.

The other big project on the nuclear few-body problem at NIC, by H.M. Hofmann et al. from Erlangen University, deals with somewhat larger nuclei. There, a straightforward solution of the Schrödinger equation is impossible, and one has to resort to variational methods. In particular, these authors use the Refined Resonating Group Model where the wave function is decomposed into a set of Gaussians. They use a genetic algorithm for optimizing the parameters of this set. If, in addition, realistic forces are used (including three-nucleon forces), one arrives again at a problem where high-performance supercomputing is essential. One reward is then that one finds good agreement with experimental scattering data in four-nucleon systems (e.g., n-t and p-3He scattering).

The two projects on atomic physics (by J. Eichler et al. from the Hahn-Meitner-Institute in Berlin, and by W. Scheid et al. from Gießen University) both are concerned with electron-positron pair production in relativistic heavy-ion collisions. In spite of its deceptive simplicity, this has been a fascinating problem ever since it was demonstrated experimentally nearly ten years ago. It obviously probes the structure of the vacuum, it cannot be treated perturbatively, and it needs a fully relativistic treatment. The latter is highly non-trivial. Indeed, as shown by Eichler et al., all known numerical schemes violate Lorentz symmetry, and the induced errors are not small. Scheid and colleagues, on the
other hand, show that previous perturbative treatments were substantially wrong.

The numerical treatment of pair production in heavy-ion collisions is made feasible since the two nuclei can be assumed to move on straight lines. In spite of this simplification, the problem remains a formidable challenge for high-performance supercomputing. Indeed, the various numerical methods used by both groups are quite different, although all of them finally burn down to the integration of large systems of differential equations.

In summary we see again that high performance supercomputing is essential if one wants to make detailed comparison with experiments, given the inherent complexities of even simple-looking systems. This is not to say that no progress can be made with much less numerical efforts, in particular if the theorist can select the problem he wants to work on. But there will always be problems – and these often are the great challenges – which require computations which can only be done on state of the art supercomputers.
Systems of few nucleons can be solved in a numerically precise manner on supercomputers, which allows to probe nuclear forces in an unprecedented manner. We present achievements and challenges for three- and four- nucleon bound states, three-nucleon scattering reactions and electron induced processes in the three-nucleon system. The need for three-nucleon forces is documented. We also demonstrate that the presently available most modern hyperon-nucleon forces fail to describe the lightest few-body hypernuclei.

1 Introduction

The question, what is the nature and what are the properties of nuclear forces, is as old as nuclear physics. Despite an intensive research of more than half a century this question is still not satisfactorily answered. As we know now, the reason is the composite nature of the nucleons, where the subnucleonic degrees of freedom are governed by a mathematically very complex nonlinear field theory, the Quantum Chromodynamics. A rigorous treatment is not yet possible. Instead one still relies on a more effective picture, going back to Yukawa, that nuclear forces are generated by the exchange of various types of mesons between the nucleons. Qualitatively this mechanism describes quite well the rich set of measured nucleon-nucleon (NN) scattering data at not too high energies. When fine tuned and augmented by additional phenomenological structures it leads to several so-called high precision NN interaction models, which perfectly describe the NN data set.

Now, the NN system alone is apparently not all of nuclear physics. There are hundreds of nuclei with many nucleons bound together by nuclear forces. One would like to understand quantitatively the amount of binding energy, which is defined as the mass of the nucleus minus the nucleon number times the mass of a nucleon. That difference, the binding energy, is smaller than zero and it requires an input of energy to separate the nucleus into free nucleons. The overall attractive nuclear forces work against that separation. One can also hit a nucleus by a nucleon leading to a so called nuclear reaction process with many facets, which one would like to understand theoretically. There are of course many more questions to be posed but already these most elementary ones serve to illustrate the basic challenge: One has to solve the dynamical equation for the motion of the nucleons under the action of nuclear forces in order to answer these questions. For low energy nuclear physics this is the Schrödinger equation of quantum mechanics.

At this point enters the spirit of few-nucleon physics. That Schrödinger equation can be solved on supercomputers in a numerically rigorous manner for a small number of nucleons, not yet for a large number. Therefore systems of say 3,4 nucleons are an ideal test laboratory to probe nuclear forces and to find out whether the present day nuclear force
models are good enough to lead to correct binding energies and to a correct description of nuclear reaction processes. When we talked up to now of nuclear forces we meant forces acting between two nucleons. The composite nature of the nucleons, however, and thus the possibility of internal excitations, lead in the very natural manner, as we shall sketch below, to additional new types of forces, three-nucleon forces. Trivially they can act the first time, when three nucleons are together. Will that go on and one has to expect proper four-nucleon forces if four nucleons are together? Then the task of understanding of nuclear physics would be hopeless. Thus we face a very first question: will NN forces as determined from NN scattering data be sufficient to a large extent to also describe systems of 3,4... nucleons correctly or will many-nucleon forces play a never ending role? We shall describe the present inside in our first result section.

Normal nuclei are composed of neutrons and protons, to which we referred generically as nucleons. There exist more particles in nature with similar masses as the nucleons. Though they are not stable, they live long enough to be combined with nucleons to form new types of nuclei, hypernuclei. It is a great theoretical challenge to study the strong forces acting between these new types of particles, called hyperons (Λ’s and Σ’s), and the nucleons, as well as the forces among the hyperons. Again light systems are an ideal laboratory since they are accessible by supercomputers. Our results will be displayed in the second result section.

Finally we would like to select one further approach to study nuclear forces, namely scattering electrons on nuclei. The electron emits a photon, which is absorbed by the nuclear system and triggers nuclear reactions. This is an important tool to probe nuclear forces and is intensively pursued experimentally at electron facilities all over the world. We present our results in the third result section and end with an outlook.

2 Results

The Schrödinger equation for n nucleons reads

$$\left( \sum_{i=1}^{n} T_i + \sum_{i<j}^{n} V_{ij} + \sum_{i<j<k}^{n} V_{ijk} + \ldots \right) \Psi = E \Psi$$

The Hamiltonian acting on the wavefunction \( \Psi \) is composed of the kinetic energy \( T_i \), the NN interactions \( V_{ij} \), three-nucleon interactions \( V_{ijk} \) and possible higher-order nuclear forces. In configuration space this is a 3n-fold partial differential equation. In case of bound states E is directly related to the binding energy of a nucleus and in case of nuclear reactions E is related to the energy by which a nucleon hits a nucleus. In the form of Eq. (1) the numerical task is not yet tractable even on the most powerful present day supercomputers. The wave function \( \Psi \) describes the very complex correlations in the motion of the n nucleons: two nucleons cannot approach each other too closely due to the strong short range repulsion of NN forces and 3N forces impose even more complex restrictions. A break-through was established in the form of the Faddeev-Yakubovsky equations (FY) which break \( \Psi \) into pieces according to the various sequences of subclusters of the whole n-body problem. These pieces of \( \Psi \) are responsible only for certain subcluster correlations and not for all at the same time as \( \Psi \). These pieces obey a set of mathematically well defined coupled equations, the FY equations.
The motion of the interacting nucleons can be classified according to the many ways orbital angular momenta and spins are coupled to a total conserved angular momentum. This introduces a great amount of analytical insight and reduces the number of variables dramatically. Nevertheless the resulting number of unknowns for the discretized form of the FY equations is quite large. It amounts to typically $10^5$ in a three-body system and up to $10^8$ in a four-body system. In other words one has to handle (essentially) full matrices up to the order $10^8 \times 10^8$. Apparently this requires supercomputers and our results are based on both, highly vectorized as well as massively parallelized codes.

Let us now regard the problems and our results.

2.1 The Lightest Nuclei and Nuclear Reactions

The starting point are the most modern high precision NN forces as mentioned in the introduction. What will be the outcome if these pair forces act between two neutrons and one proton, like in case of $^3$H, or between two protons and two neutrons in case of $^4$He? Will the Schrödinger equation or the mathematically equivalent FY equations yield the correct masses of these nuclei and thus the correct binding energies? How are the nucleons spatially arranged in these light nuclei?

We show in Fig. 1 their most probable symmetric geometries as they are predicted by $\Psi$ from Eq. (1). The most probable pair distances are close to 1 fm=$10^{-13}$ cm (1 fm is a typical unit for the range of nuclear forces). The blue background indicates that all orientations in space of these geometrical patterns have to be superimposed. The amount of extension of the nucleons shown as little balls is just an artistic view. The binding energies (in units of MeV) are presented in Table 1 for various high precision NN forces and compared to the experimental values. We see, theory and experimental values disagree. The theoretical numbers are smaller in magnitude, which tells that the NN forces alone do not bind the nuclei as strongly together as in nature.

This result should not come as a surprise in view of the remarks made in the introduction. 3N forces might act in nature in addition. To understand their properties is a hot topic right now. One very natural mechanism for a 3N force is depicted in Fig. 2, where processes following one another in time are shown from left to right. First three nucleons move upwards along vertical lines. Then nucleon number 1 emits a pion, which is absorbed by nucleon 2 and converts thereby nucleon 2 into an excited state, the famous $\Delta$-excitation of the nucleon. In that intermediate state there are no longer three nucleons in their ground state present but one of them is excited and appears as a new particle, the $\Delta$ with different quantum numbers. Further on in time the $\Delta$ gets deexcited by emitting a pion which is
then absorbed by nucleon 3. We end up again with 3 nucleons. This process induces a force between the three nucleons. In mathematical terms one leaves the Hilbert space of 3 nucleons at the intermediate state and consequently that process can not be reduced to a sequence of pair forces between nucleons. It is a proper 3N force mechanism.

![Diagram](image-url)

Figure 2. Two-pion-exchange 3NF.

<table>
<thead>
<tr>
<th>NN force</th>
<th>$^3$H</th>
<th>$^4$He</th>
</tr>
</thead>
<tbody>
<tr>
<td>CD Bonn</td>
<td>-8.013</td>
<td>-26.26</td>
</tr>
<tr>
<td>AV18</td>
<td>-7.628</td>
<td>-24.28</td>
</tr>
<tr>
<td>Nijm I</td>
<td>-7.741</td>
<td>-24.98</td>
</tr>
<tr>
<td>Nijm II</td>
<td>-7.659</td>
<td>-24.56</td>
</tr>
<tr>
<td>Nijm'93</td>
<td>-7.668</td>
<td>-24.53</td>
</tr>
<tr>
<td>Experimental value</td>
<td>-8.482</td>
<td>-28.30</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Nuclear forces</th>
<th>$^3$H</th>
<th>$^4$He</th>
</tr>
</thead>
<tbody>
<tr>
<td>CD Bonn + TM</td>
<td>-8.478</td>
<td>-29.15</td>
</tr>
<tr>
<td>AV18 + TM</td>
<td>-8.478</td>
<td>-28.84</td>
</tr>
<tr>
<td>AV18 + Urbana IX</td>
<td>-8.484</td>
<td>-28.50</td>
</tr>
<tr>
<td>Experimental value</td>
<td>-8.482</td>
<td>-28.30</td>
</tr>
</tbody>
</table>

Table 1. $^3$H and $^4$He binding energies obtained with different NN force models compared to the experimental values.

Table 2. $^3$H and $^4$He binding energies obtained with different nuclear force models compared to the experimental values.

Related to the mechanism of Fig. 2 is the Tucson - Melbourne (TM) 3NF. It contains
one free parameter, which can be used to fix the 3N binding energy. Thus like NN data are used to adjust model parameters in the NN force the 3N binding energy is used to adjust a model parameter in the 3NF. This is not as bad as it might look in the very first moment. Though one has lost the predictive power for the 3N binding energy one can now predict the 4N binding energy. Furthermore there is a huge amount of other 3N observables in the realm of nuclear reactions, which can now be predicted.

Table 2 shows the adjusted $^3$H binding energies (the partially not perfect adjustment is of no relevance) and the predictions for $^4$He. It also includes predictions based on a second very popular 3NF, the Urbana IX. The additional 3NF’s improve significantly the $^4$He results, but there is now a small overbinding. At present time that small overestimation can be considered as an indication for the possible action of a very small 4N force of repulsive character.

2.2 Nuclear Reactions for Three Nucleons

The famous classical 3-body problem is still unsolved analytically; the quantum mechanical one is not easier and only thanks to supercomputers it can now be solved. One can distinguish elastic scattering of a nucleon on a deuteron or inelastic processes, where the deuteron breaks apart and in the final state three nucleons emerge from the interaction region. While the three nucleons are close together within the range of nuclear forces classically spoken they follow very complex intricate trajectories. Quantum mechanically one has to solve the Schrödinger equation in the form of FY equations nonperturbatively, because the nuclear forces are rather strong. The nucleon as a spin 1/2 particle can be experimentally prepared in two spin states, say up and down; the deuteron with overall spin 1 in 8 spin states (vector and tensor polarized). This refers for instance to the states of the two reaction partners before the reaction takes place, but also the final states, can be separated according to their different spin orientations. Thus there is quite a rich number of possible scattering processes starting from a certain spin orientation and ending up in the same or another one in the final state. The probabilities for all these detailed processes are governed by the spin-dependent nuclear forces. In case of the deuteron break up process the three outgoing nucleons share the total available energy in a continuously distributed manner and they can leave the reaction zone into arbitrary directions. This together with spin orientations provides an even richer amount of information when measured to challenge the Schrödinger equation together with its dynamical input, the nuclear forces. Lack of space forces us to select just four examples which demonstrate successes and failures of our present day understanding.

We show in Fig. 3 a schematic set up for elastic nucleon-deuteron scattering. We see the initial nucleon at the left approaching the deuteron (in the middle). Then after the elastic scattering has taken place in a certain reaction volume in space nucleon and deuteron separate from each other and are registered under certain angles $\theta$ and $\theta'$ in detectors. The arrows on the initial nucleon (one of them in bracket) indicate two possible spin orientations and the three arrows on the deuteron indicate a situation, where one averages over the various spin orientations of the deuteron. In other words this set up corresponds to the case that only the initial nucleon is polarized (spin oriented). Fig. 4 shows the angular distribution of the scattered nucleon for a totally unpolarized situation, where also an average over the two spin orientations of the initial nucleon has been performed. That distribution
is a measure for the probability that a nucleon is elastically scattered into a certain angle $\theta$, whereas the deuteron is scattered into a corresponding angle $\theta'$. The left and right parts refer to two different energies, 65 and 135 MeV of the incoming nucleon, respectively. The blue bands comprise the predictions based on the present day most modern high precision NN forces and fail to reproduce the experimental results, given as dots. If we add the 3NF adjusted in $^3$H as described above theory, as given by the red band, moves without any further modification right away into the data. We consider this to be a beautiful signature of a 3NF effect.

Let's move on to observables which depend on spin orientations in Fig. 5. The left figure shows the situation that a difference of angular distributions is formed for nucleon spin up and down related to their sum. One talks of an analyzing power $A_y$. The right figure displays a corresponding analyzing power now with a spin flip of the deuteron leading to $A_y(d)$. Again we see the two bands. For $A_y(d)$ the 3NF moves theory beautifully into the data, but for $A_y(N)$ the 3NF effect which we have taken into account up to now is negligible and one is left with a serious problem, called the low-energy $A_y$ puzzle.

In case of the deuteron break up process three nucleons emerge from the reaction region. Due to kinematical constraints (energy and momentum conservation) it is sufficient to detect two nucleons, which fixes then also the momentum of the third nucleon. We select...
Figure 5. Nucleon and deuteron vector analyzing powers in elastic nucleon-deuteron scattering.

Figure 6. Examples of the breakup cross sections.

In Fig. 6 two examples out of an unlimited number of different breakup configurations. In both cases the frequency of the breakup events (cross section) is displayed as a function of the way the energy is distributed over the three final particles (parameter S). The left figure is for a situation that two of the final nucleons leave the interaction region in the same direction; in the right figure the final three nucleons leave under \( 120^\circ \) interparticle angles in a specially oriented plane. We see a rich structure in the number of events in the first case, which is beautifully described in theory when only NN forces act (3NF effects are negligible) and we see a clear discrepancy in the second case. Though the data fluctuate a bit there are clearly more events in nature than in theory and 3NF effects considered up to now turned out to be negligible.

Apparently there are successes and failures. These numerically rigorous calculations stimulated world wide experimental activities at several accelerator facilities to enrich and improve the experimental data and to challenge theory even stronger. One can say that these few-nucleon studies are a perfect tool to probe and finally fix the Hamiltonian underlying low-energy nuclear physics.
2.3 The Lightest Hypernuclei

Hyperon-nucleon (YN) forces are not strong enough to form two-body bound states. Thus the lightest hypernucleus $^3\Lambda H$, the so called hypertriton, plays the role of the deuteron in hypernuclear physics. Next come the two hypernuclei $^4\Lambda H$ and $^4\Lambda He$. We shall investigate these three nuclei using the FY equations and the most modern YN forces from the Nijmegen group together with the high precision nuclear forces used before. The numerical challenge is much higher than for nuclei since there are more coupled equations due to the distinguishability of the hyperon from the nucleon and due to an important dynamical effect. There are strong transitions which change the $\Lambda$-hyperon into a $\Sigma$-hyperon during the interaction of the hyperon with a nucleon. This $\Lambda - \Sigma$ conversion brings in an interesting facet: the hypertriton is at one moment a system of $\Lambda$, neutron and proton and in another moment a system of $\Sigma$, neutron and proton. Our results\(^6\) for the separation energies (there is a fixed link to binding energies) based on two versions of the Nijmegen YN forces are displayed in Table 3. The lightest hypernucleus $^3\Lambda H$ is well described by only one of the two YN forces. Because of the very low separation energy its dominant structure is a deuteron surrounded far out by a $\Lambda$ particle. The other appearance of the same hypernucleus is a deformed deuteron with a $\Sigma$-hyperon sitting very close by. This fascinating fluctuating picture is depicted schematically in Fig. 7. Quantitatively the most probable distance of the $\Lambda$-hyperon from the center of mass of the two nucleons (deuteron) is 11 fm, which means that the $\Lambda$ particle in $^3\Lambda H$ remains most of the time outside the range of the strong interaction. Nuclei with this astonishing property are called “halo” nuclei and $^3\Lambda H$ is a very spectacular one.

<table>
<thead>
<tr>
<th>NN and YN forces</th>
<th>$^3\Lambda H$</th>
<th>$^4\Lambda H$</th>
<th>$^4\Lambda H^*$</th>
<th>$^4\Lambda He$</th>
<th>$^4\Lambda H^*$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nijm’93 + SC89</td>
<td>0.143</td>
<td>1.80</td>
<td>–</td>
<td>2.14</td>
<td>0.02</td>
</tr>
<tr>
<td>Nijm’93 + SC97e</td>
<td>0.023</td>
<td>1.47</td>
<td>0.73</td>
<td>1.54</td>
<td>0.72</td>
</tr>
<tr>
<td>Experimental value</td>
<td>0.130 ± 0.050</td>
<td>2.04</td>
<td>1.00</td>
<td>2.39</td>
<td>1.24</td>
</tr>
</tbody>
</table>

Table 3. Theoretical and experimental separation energies of the lightest hypernuclei.

Regarding now the separation energies for the 4-body hypernuclei we face serious discrepancies as displayed in Table 3. Neither the separation energies in relation to the ground state nor to the excited state (marked with a star) are acceptably close to the data and this for both YN force models. It is only due to our rigorous calculations that this has been discovered. Previous approximate evaluations have hidden these facts. Because of the lack of hyperon beams the YN data set is extremely scarce and therefore the separation energies of the light hypernuclei have to serve as the important test cases for probing theoretical models of YN forces. Without supercomputers this could not have been carried out.

2.4 Probing Nuclear Dynamics by Photons

The interaction of a photon with a charged particle is weak and can be treated perturbatively. This opens the way to probe the strong dynamics, which can not be treated perturbatively, in a well controlled manner. Worldwide electron and photon facilities, like MAMI
Figure 7. Schematic picture of the Λ and Σ appearances in the $^3\Lambda H$ wave function.

and ELSA in Germany, take advantage of that to investigate nuclei and nucleons themselves. With respect to nuclei the very lightest ones are more and more in the focus, since few-nucleon systems can be treated rigorously. Regarding for instance electron scattering on $^3$He various processes can occur: elastic scattering where in the final state $^3$He emerges intact and inelastic scattering, where in the final state fragments of $^3$He occur, either a proton and a deuteron or two protons and a neutron. Moreover one can fix the spin orientations of the electron and of the nucleus $^3$He in the initial state before the reaction takes place, which increases the amount of information one can extract from the measurement. Here we have to restrict ourselves to just one process, inclusive electron scattering of polarized electrons on polarized $^3$He: $^3$He($e^-, e'^- - e^+$). This is illustrated in Fig. 8. On the left we see an electron emerging from the accelerator and its spin (red arrow) is oriented parallel to its velocity. Then the electron emits a photon, which travels to the right and the electron changes its direction of flight by an angle $\theta$. This final electron (denoted by $e'$) is then registered in a detector. The photon hits the nucleus $^3$He and breaks it apart. The outcoming nucleons are not measured in that so called inclusive reaction. But on the way out of the reaction zone they interact strongly with each other by the nuclear forces, which is indicated by the wiggly lines. It is a property of the nuclear forces that the total spin of $^3$He is carried to more than 90% by the neutron (in green), while the spins of the two protons are oriented opposite to each other, which cancels their effect. This opens the fascinating possibility to extract neutron properties if one selects judiciously those pieces of the photon interaction with the nucleon which are sensitive to its spin. This can be achieved by performing the measurement as depicted in Fig. 8 twice, one with the spin of the initial electron pointing into the direction of flight and one pointing into the opposite direction. Then one forms the ratio of the difference of the number of measured events to their sum. That ratio called an asymmetry, $A$, is plotted in Fig. 9 as a function of the energy $\omega$ which the photon carries and injects into the nucleus. First of all we see the experimental results presented by open circles. Then we see the green curve, which describes the theoretical results by neglecting all interactions among the final nucleons. This is not at all the case in nature and it has to be taken into account properly and then leads to the blue curve, which now already comes closer to the data. The last very important dynamical ingredient comes from a closer look into the mechanism of the photon absorption by $^3$He. There are basically two processes illustrated in Fig. 10. Either the photon is absorbed just by one nucleon and the others are spectators or it is absorbed by the charged mesons during their flight from one nucleon.
to another. It is the inclusion of the second type of the photon absorption process which finally brings theory right into the data (red curve). This is of course an important result and insight by itself but one should add that this experiment and theoretical analysis has been performed also with the aim to extract electromagnetic properties of the neutron, namely its electromagnetic form factors. This has been achieved beautifully and can, because of lack of space, not described here further. The knowledge of the electromagnetic form factors provides important information how
charges and magnetic properties are distributed inside a neutron. Since there are no free neutron targets experiments on the deuteron or $^3\text{He}$, each with one neutron inside, are mandatory to extract that information on the neutron. The theoretical analysis of that requires as we have seen a profound knowledge of the nuclear forces and a very reliable understanding of these light nuclei and of the interaction processes going on among the nucleons after the photon has been absorbed. We were able to perform the investigations of these electromagnetically induced processes based on the basic research described before.

Another important remark should be added. The preparation of spin oriented $^3\text{He}$ nuclei necessary for this sort of experiments required world wide efforts and was carried through in view of gaining access to the properties of the neutron. Thus it was pure basic research. It could not been foreseen that polarized $^3\text{He}$ can be used to “illuminate” the human lungs and to localize defects of various types in the lungs in an unprecedented clear manner. Polarized $^3\text{He}$, developed on the basis of pure nuclear research is a beautiful example for a spin off having important application in medicine.

3 Outlook

Our results clearly demonstrate that the aim to establish the basic Hamiltonian for low-energy nuclear physics expressed in nucleon degrees of freedom is reachable. Three-nucleon forces will thereby play an important role. Chiral perturbation theory, used as a new approach to nuclear forces, will provide an input linked to the underlying QCD. Intensive research along this new line is being performed. Without supercomputers nuclear dynamics could not be probed and the growing amount of experimental data could not be analyzed.

Acknowledgments

We are indebted to the John von Neumann Institute for Computing for the possibility to perform these investigations.

References

The dramatic performance increase of computers in general, but especially of parallel supercomputers in recent years, opened up a new and exciting way to research in natural and engineering sciences: numerical simulation. Besides the traditional experimental and theoretical research approaches, a scientist may now use computers to gain insight into complex real world processes. To this end, a mathematical model based on experimental and observational data is developed. This model is then discretized with an appropriate numerical method and the resulting discrete problem must then be solved as efficiently as possible. Obviously the quality of the results is significantly dependent on the numerical method used and the quality of the discretization. In principle, this approach is applicable to a wide range of problems. But the resulting demands for storage and compute time on available computers is a limiting factor in practical applications. Hence, a major goal is the development of fast numerical methods with minimal storage requirements and their efficient implementation on large parallel computers. Another goal is to apply these techniques to practical problems from the natural and engineering sciences.

Most interesting problems can be modelled mathematically by systems of partial differential equations. Therefore, appropriate numerical discretization techniques such as finite elements, finite differences or finite volumes are of central interest. Here, the use of error estimators for the control of adaptive discretization techniques is a must to reduce the storage demands and to make three-dimensional problems with non-smooth solutions approachable. Furthermore, the fast and efficient numerical solution of the resulting large linear systems by multilevel and multigrid methods is important for efficiency reasons. Finally, all these components need to be combined and parallelized in a load-balanced fashion to be able to tackle real world problems.

These ingredients are the key to any successful and meaningful numerical simulation of three-dimensional problems. Consequently they are subject of intensive research. This is reflected in the following three articles which report on the results of numerical projects conducted on the Cray T3E in Jülich and other parallel supercomputers.

In the contribution Hash based adaptive parallel multilevel methods with space filling curves by M. Griebel and G. Zumbusch, the authors address the problem of load balancing. Especially for adaptively refined meshes, the optimal partitioning of data is in general a $NP$-hard problem. Therefore heuristics are needed. Here, a parallel and cheap method based on space-filling curves is used. The method is applied to elliptic boundary value problems. Its performance is demonstrated on two large parallel computers, namely the ASCI Blue Pacific and Jülich’s Cray T3E.

The article A parallel software-platform for solving problems of partial differential equations using unstructured grids and adaptive multigrid methods by P. Bastian, K. Johannesen, S. Lang, S. Nägele, C. Wieners, V. Reichenberger, G. Wittum and C. Wrobel deals with the software aspects of complex simulation programs. The package UG is a
system to discretize and solve PDEs on unstructured grids. It uses finite element and finite volume techniques, features adaptive refinement and various parallel multigrid solvers, especially algebraic multigrid. Load balancing for parallelization is achieved by recursive coordinate bisection and similar partitioning techniques. Applications range from classical PDEs, over two-phase flows, elasticity and plasticity to the Navier-Stokes equations and turbulence modelling.

In the third contribution *High performance FEM simulation via FEAST and application to parallel CFD via FEATFLOW*, C. Becker, S. Buijssen, S. Kilian and S. Turek report on their efforts to simulate non-stationary laminar and turbulent flow problems in complex situations. Based on their FEAST and FEATFLOW program systems, the authors aim at a scalable recursive clustering approach. They discuss computational bottlenecks for parallel multigrid and domain decomposition, motivate the approach and show the results of their simulations for an industrial problem from steel casting.

Alltogether, these three numerical projects show how to efficiently bridge the gap between adaptive discretization techniques, multigrid solvers and parallel computing when it comes to practical three-dimensional simulations of real life problems on large supercomputers like the Cray T3E. It is on the basis of such techniques that numerical simulation already is and in the future even more will be a competitive and cost effective third way to gain further insight in the natural and engineering sciences.
Hash Based Adaptive Parallel Multilevel Methods
with Space-Filling Curves

Michael Griebel\(^1\) and Gerhard Zumbusch\(^1\)

Institut für Angewandte Mathematik, Universität Bonn
Wegelerstr. 6, 53115 Bonn, Germany
E-mail: \{griebel, zumbusch\}@iam.uni-bonn.de

The solution of partial differential equations on a parallel computer usually follows the data
parallel paradigm. The grid is partitioned and mapped onto the processors. In this paper a parallelisable and cheap method based on space-filling curves is proposed. The partitioning is embedded into the parallel solution algorithm using multilevel iterative solvers and adaptive
grid refinement. Numerical experiments on two massively parallel computers prove the
efficiency of this approach.

1 Introduction

In this paper, we consider parallel versions of adaptive multigrid solvers and adaptive sparse grid discretisations for elliptic partial differential equations. The multigrid method operates on a finite difference discretisation on quad-tree and oct-tree meshes, which are obtained by adaptive mesh refinement. The adaptive sparse grid discretisation is also based on a finite difference scheme and is well suited for higher dimensional problems. A fast parallel load balancing strategy for both approaches is proposed which is defined by a space-filling Hilbert curve. It is furthermore applicable to arbitrary shaped domains. Some numerical experiments demonstrate the parallel efficiency and scalability of the approach.

2 Multigrid Solver

Our goal is to solve an elliptic partial differential equation efficiently. The PDE is discretised by finite differences on a 1-irregular quad-tree or oct-tree grid. Here, we set up the operator as a set of difference stencils from one node to its neighboring nodes in the grid, which can be easily determined: Given a node, its neighbors can be only on a limited number of levels. The distance to the neighbor is determined by the refinement level the nodes share, see Figure 1.

So pure geometric information is sufficient to apply the finite difference operator to some vector. Hence we avoid the storage of the stiffness matrix or any related information. For the iterative solution of the equation system, we have to implement matrix multiplication, which is to apply the operator to a given vector.

We use an additive version of the multigrid method for the solution of the equation system, i.e. the so called BPX preconditioner\(^4\). This requires an outer Krylov iterative solver. The BPX preconditioner has the advantage of an optimal \(O(1)\) condition number and an implementation of order \(O(n)\), which is optimal, even in the presence of degenerate grids. Furthermore, this additive version of multigrid is easier to parallelise than multiplicative multigrid versions.
The straightforward implementation is analogous to the implementation of a multigrid V-cycle. However, the implementation with optimal order is similar to the hierarchical basis transformation and requires one auxiliary vector. Two loops over all nodes are necessary, one for the restriction operation and one for the prolongation operation. They can be both implemented as a tree traversal or alternatively by iterating over the nodes in the right order.

A parallel version of multiplicative multigrid usually is based on a partition of all nested grids. The domain $\Omega$ is decomposed into several sub-domains $\Omega_j$, which induces partitions of all grids. Each processor holds a fraction of each grid in such a way that these fractions of each grid form a nested sequence. Hence each operation on a specific level is partitioned and mapped to all processors. Furthermore the communication during grid transfer operations is small because of nested sequences on a processor. This means that one has to treat global problems on each level, which are partitioned to all processors. The intra-grid communication has to be small, i.e. the number of nodes on the boundary of the partition should be small. Furthermore the amount of work on coarse grid levels usually is small and each processor does not compute much. There are several strategies to deal with the coarse grid problem in general, such as to centralize the computation on a master processor, to perform identical computations on all processors or to modify the coarse grid correction step.

A static partition of the domain into strips or squares can be used for uniform grids and has been used for the first parallel multigrid implementations, see also the survey Ref. 5. In contrast to the geometry oriented parallelisation of multiplicative multigrid methods, the additive multigrid version or additive multilevel preconditioners can be parallelised in a more flexible way. The overall workload has to be partitioned, but we do not have to consider individual levels. Here, the communication takes place in a single step for all nodes, which are located on the boundary of at least one grid of the nested sequence. The multilevel BPX preconditioner for a uniform grid has been parallelised. These approaches can easily be generalized to block-structured grids.

Figure 1. Adaptively refined quad-tree grid (left) and a finite difference stencil (right)
3 Sparse Grids

Before we consider the parallelisation in more detail, we introduce our second discretisation strategy, which is especially well suited for higher dimensional problems. Sparse grids are a multi-dimensional approximation scheme, which is known under several names such as ‘hyperbolic crosspoints’, ‘splitting extrapolation’ or as a boolean sum of grids. Probably Smolyak\textsuperscript{8} was the historically first reference. Directly related to the boolean construction of the grids was the construction of a multi-dimensional quadrature formula. Both, quadrature formulae and the approximation properties of such tensor product spaces were subject to further research, see Temlyakov\textsuperscript{9} and others\textsuperscript{10–13}. The curse of dimension was also subject to general research on the theoretical complexity of higher-dimensional problems. For such reasons, sparse grids play an important role for higher-dimensional problems. Besides the application to quadrature problems, sparse grids are now also used for the solution of PDEs. They were introduced for the solution of elliptic partial differential equations by Zenger\textsuperscript{14}, where a Galerkin method, adaptive sparse grids and tree data structures were discussed. At the same time a different discretisation scheme based on the extrapolation of solutions on several related, regular grids was proposed, the ‘combination’ technique\textsuperscript{15}.

In this paper we consider a finite difference (FD) scheme\textsuperscript{16, 17}. It is simpler to implement and to apply to different types of equations, but there is not that much known analytically yet. We focus on the parallelisation of such a FD scheme on sparse grids. So far only the ‘combination’ technique can be run on a distributed-memory parallel computer\textsuperscript{18}, which is essential for many large scale simulations. The advantage of the FD scheme is the flexibility of the grid. The sparse grid may be refined adaptively, while the ‘combination’ technique relies on regular sparse grids. However, the adaptive grid refinement poses a severe load-balancing problem, which has to be resolved during runtime of the simulation.

The multi-dimensional approximation scheme of sparse grids can be constructed as a subspace of the tensor-products of one-dimensional spaces represented by a hierarchical multi-resolution scheme, such as the hierarchical basis, see the historical reference Faber\textsuperscript{19}, or generally any basis system of pre-wavelets or wavelets\textsuperscript{20}. Each one-dimensional basis function can be derived from a model function $\phi$ by a scaling of $2^{-l}$ (also called level $l$) and a translation by a multiple of $2^{-l}$. In the case of the piecewise linear hierarchical basis, the model function $\phi$ is the well known hat function. We denote the one-dimensional space of functions up to level $l$ by $T_l$. On level $l$, the standard grid space can be written as

$$\hat{T}_l = \langle T_i \otimes T_j \otimes \ldots \rangle_{i,j,\ldots \leq l}.$$  \hfill (1)

In contrast to the regular grid, the corresponding sparse grid space consists of fewer functions. On level $l$, it can be written as

$$\tilde{T}_l = \langle T_i \otimes T_j \otimes \ldots \rangle_{i+j+\ldots \leq l}.$$  \hfill (2)

This is a subset of the regular grid space. A regular grid has about $2^{d+l}$ nodes, which is substantially more than the $2^{l \cdot l^{d-1}}$ nodes of the sparse grid.

The major advantage of sparse grids compared to regular grids is their smaller number of nodes (or grid points) for the same level $l$ and resolution $2^{-l}$. This is especially true in higher dimensions $d \gg 1$.

Of course, the question whether sparse grids have an advantage compared to regular grids does also depend on the discretisation accuracy of a solution obtained on a grid.
Furthermore the number-of-operations complexity is of interest, because it is an estimate for the computing time a specific algorithm needs. For details see Ref. 16, 17.

We define the hierarchical transformation $H$ as the hierarchical basis transformation on the regular grid from nodal values to hierarchical values, which are restricted to the sparse grid nodes. All wavelet-type basis functions provide such fast $O(n)$ transformation to and from the nodal basis representation. The transformation is especially simple for the one-dimensional hierarchical basis: Given the nodal values $u_j$ with $j = 0, 1, \ldots, 2^j - 1$, the hierarchical representation for interior points can be obtained by

$$
\hat{u}_j = u_j - \frac{1}{2} (u_{\text{left father}} + u_{\text{right father}})
$$

and the boundary nodes $u_0$ and $u_{2^j-1}$ remain unchanged. The nodal values are replaced by their hierarchical excess or deterioration, compared to the value obtained by interpolation between on the next coarser level grid. The inverse transformation can be implemented similarly. However, the coarse nodes have to be computed before the finer grid nodes. Furthermore, the transformation can be implemented in place, without an auxiliary vector. The hierarchical basis transformation $H$ is also abbreviated by the stencil $[1/2 1 1/2]$.

Based on the hierarchical basis transformation $H$, we define the action of a one-dimensional finite difference operator for the discretisation of a differential operator: We apply the associated standard difference stencil $D_i$ along the $x_i$-axis to values located on the sparse grid nodes in a specific basis representation. To this end the values are given in nodal basis in direction $i$ and in hierarchical basis representation in all other directions $I \{ i \}$. The associated transformation is denoted by $H_{I \{ i \}}^{-1}$. The stencil $D_i$ for each node itself is chosen as the narrowest finite difference stencil available on the sparse grid. It is equivalent to the corresponding stencil on a regular, anisotropic refined grid. The finite difference stencil can be a 3-point Laplacian $1/2 \left[ \begin{array}{rr} 1 & -2 \\ -2 & 1 \end{array} \right]$, an upwind-stabilized discretisation of the convective term $\frac{\partial}{\partial x_i}$, some variable coefficient operators and so on. In nodal values the finite difference operator reads

$$
\frac{\partial^2}{\partial x_i^2} u \approx H_{I \{ i \}}^{-1} \circ D_i \circ H_{I \{ i \}} u.
$$

A general difference operator is then obtained by dimensional splitting. A Poisson equation, as a simple example, can be discretised in nodal basis representation as usual as a sum of operators (4). Here the one-dimensional difference operators $D_i$ may be chosen as a three point centered Laplacian $\frac{1}{6} \left[ \begin{array}{rr} 1 & -2 \\ -2 & 1 \end{array} \right]$. On adaptively refined grids, the nearest neighbor nodes are chosen, which may lead to asymmetric stencils, i.e. non-uniform one-dimensional stencils. Further higher order modifications of the stencils have been tested, too. In the presence of a transport term in the equation, the unsymmetry is believed to be no problem. There are ways to create discretisations for all kinds of equations, e.g. for general diffusion problems, convection-diffusion problems, reaction-diffusion problems, for the Navier-Stokes equations or for hyperbolic conservation laws. Known facts on consistency and stability of this scheme are summarised in Ref. 17.
4 The Load-Balancing Problem

Finite-Element, Finite-Volume and Finite-Difference methods for the solution of partial differential equations are based on meshes. The solution is represented by degrees of freedom attached to certain locations on the mesh. Numerical algorithms operate on these degrees of freedom during steps like the assembly of a linear equation system or the solution of an equation system. A natural way of porting algorithms to a parallel computer is the data distribution approach. The mesh with attached degrees of freedom is decomposed into several partitions and mapped to the processors of the parallel computer. Accordingly also the operations on the data are partitioned. Goals of a partitioning scheme are load-balancing and little communication between the processors. Sometimes also singly-connected partitions are required. If the partitions are determined during run-time, furthermore a fast partitioning scheme itself is sought. This is e.g. the case within adaptive mesh refinement of a PDE solver.

The partitioning problem in general is \(NP\)-hard\(^{22}\). There are many heuristics based on graph connectivity or geometric properties to address this problem\(^{23–27}\). In practice fast heuristics are known. However, there is not much known about the general quality of these methods\(^{28}\). In contrary there exist examples, where single heuristics give really bad results.

In this paper we propose a specific geometry based heuristic with space-filling curves. It is cheap and helps to simplify the implementation of parallel algorithms\(^{29–34}\). We are interested in bounds for the quality of the partitions. This will lead us to general estimates on the parallel performance of advanced numerical algorithms on these partitions.

5 Space-Filling Curves

First we have to define curves. The term curve shall denote the image of a continuous mapping of the unit interval to the \(\mathbb{R}^d\). Mathematically, a curve is space-filling if and only if the image of the mapping does have a classical positive \(d\)-dimensional measure. The curve fills up the whole domain. For reasons of simplicity we restrict our attention to simple domains. We are interested in a mapping

\[
f : [0, 1] := I \mapsto \Omega \subset \mathbb{R}^d, \quad f \text{ continuous and surjective.} \tag{5}
\]

There are classical space-filling curves like the Hilbert-, the Peano- and the Lebesgue-curve\(^{35}\). However, we will also construct special space-filling curves on an unstructured mesh.

A space-filling curve can also be used for the inverse mapping \(f\) from a domain \(\Omega \subset \mathbb{R}^d\) to the unit interval \(I\). This means that we can map geometric entities in \(\mathbb{R}^d\), such as elements or nodes, to the one dimensional interval. Entities, which are neighbors on the interval, are also neighbors in the volume \(\mathbb{R}^d\). Unfortunately the reverse cannot be true and neighbors in the volume may be separated through the mapping.

However, we can solve the resulting one-dimensional partition problem: We cut the interval \(I\) into disjoint sub-intervals \(I_j\) of equal workload with \(\bigcup_j I_j = I\). This gives perfect load-balance and small separators between the partitions. The partition \(f(I_j)\) of the domain \(\Omega\) induced by the space-filling curve with \(\bigcup_j f(I_j) \supset \Omega\) also gives perfect load-balance. However, the separators \(\partial f(I_j) \setminus \partial \Omega\) (boundary of the geometrical sets) are larger than the optimal separators in general as we will see in the following.
6 Quality of a Partition

We use a basic performance model for a distributed memory computer. The execution time of a program consists of computing time, which is proportional to the number of operations on a processor, and of communication time. Communication between the processors is implemented with message passing through some network and requires time linear in the size of data $t = t_{\text{startup}} + n \cdot t_{\text{bandwidth}}$.

We consider $O(n)$ algorithms which are linear in the size of data $n$, e.g. FEM matrix assembly for $n$ finite elements, sparse matrix multiply or components of a multigrid algorithm such as a grid transfer or smoother. The parallel computing time is $C_1 \cdot n/p$ for a partition of $n$ data onto $p$ processors. We call $v := n/p$ the volume. The runtime depends on the communication time. The data to be transferred is proportional to the separator or surface $s_j$ of the partition $s_j := \partial f(I_j) \setminus \partial \Omega$. Altogether we have

$$t = C_1 \frac{n}{p} + C_2(t_{\text{startup}} + s \cdot t_{\text{bandwidth}}).$$

(6)

This model suggest that we have to minimize the surface to volume ratio $s/v$ of the partition for a high parallel efficiency of

$$\text{efficiency} = \frac{1}{1 + \frac{C_2}{C_1} \left( \frac{1}{v} t_{\text{startup}} + s \cdot t_{\text{bandwidth}} \right)}.$$

(7)

While the lowest continuous surface to volume ratio is obtained for the sphere by

$$s = \sqrt{\frac{2d}{d-1} \pi^{d/2}} \frac{v^{(d-1)/d}}{1},$$

we usually deal with partitions aligned with the mesh. Hence the cube with $s = 2dv^{(d-1)/d}$ is of interest. In general we regard estimates of type

$$s \leq C_{\text{part}} \cdot v^{(d-1)/d}$$

(8)

with low constants $C_{\text{part}}$ as optimal.

7 Estimates for Space-Filling Curves

The estimate for the locality of a discrete space-filling curve $F$ we will use with $F : [1, \ldots, k^d] \mapsto [1, \ldots, k]^d$ is of type

$$\|F(x) - F(y)\|_2 \leq C \sqrt{|x - y|}.$$  

(9)

Gotsman and Lindenbaum give an upper bound $C = (d + 3)^{d/2}2^d$ for the Hilbert curve and tighter bounds for $C = 6\frac{d}{2}$ for $d = 2$ and $C = 23$ for $d = 3$, which has been improved. Analogous estimates have been derived for the Hilbert curve and the Peano curve. It turns out that a version of the Sierpinski curve, also called H-index gives even better constants.

Lemma 1: Given a connected discrete space-filling curve $F$ on a domain $[1, \ldots, k]^d$ and a partition $F([j, \ldots, j + v - 1])$ of $v$ nodes, the surface $s$ of the partition is bounded by (8). The constant $C_{\text{part}}$ depends on the curve.

The proof is based on (9) and the connectedness of the partition. It is sufficient to consider $s$ of the bounding box.
This lemma does not hold for curves of Lebesgue type which are also called bit-interleaving, because the discrete partitions tend to be disconnected. However, we can generalize the situation to unstructured and adaptively refined meshes by the following construction: We create an enumeration of a mesh by some heuristic in order to obtain a 'local' discrete space-filling curve. Then we perform mesh refinement by some geometric refinement rules. Each element $E_j$ of the coarse grid is substituted by several smaller elements $E_{j,k}$. The enumeration is changed such that it cycles through these new elements $E_{j,k}$ right after the elements $E_{j-1}$ or $E_{j-1,k}$. This leads in the limit to a continuous space-filling curve. Alternatively a standard continuous space-filling curve can be super-imposed onto the grid.

**Corollary 2:** Estimate (8) also holds for a space-filling curve partitioning of a (quasi-) uniform mesh by superposition of $f$ or mesh dependent construction of $f$.

Estimate (7) combined with corollary 2 gives a parallel efficiency for large problems of

$$\text{efficiency} = 1/(1 + \frac{C_2 C_{\text{part}} t_{\text{bandwidth}}}{C_1} \cdot \frac{p}{n^{1/d}}).$$

(10)

This implies optimal parallel efficiency for very large problems, $n \to \infty$. Estimate (10) holds for a code for the solution of partial differential equations in the steps of setting up an equation system, a single matrix multiply and a fixed number of Krylov iterations. Furthermore, using the same space-filling curve on all grid levels, this also holds for an additive multigrid implementation and for standard multigrid if we neglect terms $\log n \cdot t_{\text{startup}}$ proportional to the number of grid levels. For the scalability of a global PDE solver an $O(n)$ multigrid solver is essential. Solvers with higher than linear complexity may scale in $p$ like (10) but scale completely different in $n$. Under suitable conditions, the estimates can be generalized to adaptively refined grids.

### 8 Key-Based Addressing

Instead of linked lists or trees, we propose to use hash storage techniques. First we describe a key based addressing scheme. The entity (a node) is stored in an abstract vector, where it can be retrieved by its key. Furthermore it is possible to decide, whether a given key is stored in the table or not, and it is possible to loop over all keys stored in the vector. In order to reduce the amount of storage of the grid, we omit any pointers and use keys instead. For a (hyper-) cube shaped domain $\Omega = [0,1]^d$, we can use the coordinates of a node for addressing purposes. The coordinates (and the keys) of hierarchical son nodes and father nodes can easily be computed from the node’s coordinates. The computation of neighbor nodes requires special care, because it is not immediately clear, where to look for the node. Given a one-irregular grid with hanging nodes, for example, a neighbor node can be located in the distance of $h$ or $2h$ from the node with a local step-size $h$, see Figure 1 (right). In the worst case this results in two vector look-up operations, one in distance $h$ along a coordinate direction and, if it was unsuccessful, one look-up in distance $2h$. Similar key based addressing schemes can be obtained for other grid refinement procedures and for different domains.

Key based addressing does simplify the implementation of a sequential, adaptive code. Now, we generalize the concept of key addressing and hash tables to the parallel case.
The idea is to store the data in a hash table located on the local processor. However, we use global keys, so a ghost copy of the node may also reside in the hash table of a neighbor processor. Furthermore, we base the code on the space-filling curve partitions of the previous section. The position of a node on the space-filling curve, along with the known partition, defines the home processor of a node. Given a node on a processor, it is easy to determine to which processor the node belongs to. If a node occurs, which does not belong to the processor, it must be a ghost copy, and it is computable where to find its original.

The next idea is to combine the position on the space-filling curve with the hash key. The computation of the position on the curve can be computed for any given coordinate tuple. It is a unique mapping \([0, 1]^d \rightarrow [0, 1]\) similar to the mapping required for hash keys. The position can be used as a key. Furthermore, for the construction of the hash table, we need a hash function. This can be any mapping \([0, 1] \rightarrow [0, m]\) with a large integer number \(m\), preferably prime. Many cheap functions related to pseudo-random numbers will do here. Modifications of the hash function can improve the cache performance of the code: Space-filling curves introduce locality in the key addressing scheme, which is used for the parallelisation of the code. Exploiting the data locality once again on the local processor, one can optimize the usage of secondary disk storage and of the memory hierarchy of caches, which is difficult otherwise.

This framework for the parallelisation of adaptive codes originally has been invented for particle methods and has been generalized to programming environments for some grand challenge PDE projects. Multigrid methods have been considered in Ref. 29, 36.

### 9 Numerical Experiments

<table>
<thead>
<tr>
<th>level</th>
<th>nodes</th>
<th>1</th>
<th>4</th>
<th>4 uni proc</th>
<th>16</th>
<th>64</th>
<th>256</th>
<th>512</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>289</td>
<td>0.81</td>
<td>1.68</td>
<td>1.42</td>
<td>2.31</td>
<td>4.50</td>
<td>10.67</td>
<td>11.36</td>
</tr>
<tr>
<td>5</td>
<td>1089</td>
<td>9.51</td>
<td>4.02</td>
<td>3.74</td>
<td>3.40</td>
<td>5.35</td>
<td>10.66</td>
<td>24.79</td>
</tr>
<tr>
<td>6</td>
<td>4225</td>
<td>45.17</td>
<td>12.97</td>
<td>12.48</td>
<td>6.90</td>
<td>6.76</td>
<td>12.45</td>
<td>34.35</td>
</tr>
<tr>
<td>7</td>
<td>16641</td>
<td>267.7</td>
<td>53.08</td>
<td>51.77</td>
<td>18.28</td>
<td>14.15</td>
<td>18.15</td>
<td>23.59</td>
</tr>
<tr>
<td>8</td>
<td>66049</td>
<td>1722</td>
<td>302.4</td>
<td>289.3</td>
<td>72.92</td>
<td>29.20</td>
<td>33.10</td>
<td>32.58</td>
</tr>
<tr>
<td>9</td>
<td>263169</td>
<td>1891</td>
<td>1792</td>
<td>419.0</td>
<td>95.47</td>
<td>58.31</td>
<td>67.26</td>
<td></td>
</tr>
</tbody>
</table>

Table 1. Poisson problem, uniformly refined grids in two dimensions, timing of the multigrid solution on ASCI Blue Pacific.

As test cases for our approach, we consider the Poisson equation and a convection-diffusion problem. We use a finite difference discretisation, where the degrees of freedom are associated with the nodes, and the differential operator is defined on the edges connecting the nodes. In a similar fashion the sparse grid discretisation and the additive multigrid can be defined.

All numbers reported are scaled CPU times measured on a Cray T3E parallel computer with 300MHz Alpha 21164 processors (T3E-600) and on ASCI Blue Pacific (technology refresh, PowerPC 604e, 332 MHz). We use the portable message passing MPI programming interface and its intrinsic timing routines.
<table>
<thead>
<tr>
<th>time</th>
<th>processors</th>
<th>level</th>
<th>nodes</th>
<th>1</th>
<th>4</th>
<th>4 uni proc</th>
<th>16</th>
<th>64</th>
<th>256</th>
<th>512</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>1 4 4 uni proc</td>
<td>4.69</td>
<td>4.64</td>
<td>8.26</td>
<td>17.34</td>
<td>26.51</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>4913</td>
<td>55.85</td>
<td>53.65</td>
<td>17.19</td>
<td>12.63</td>
<td>20.65</td>
<td>33.76</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>35937</td>
<td>1088</td>
<td>1036</td>
<td>155.1</td>
<td>44.01</td>
<td>36.18</td>
<td>65.32</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>274625</td>
<td></td>
<td>3116</td>
<td>383.4</td>
<td>110.5</td>
<td>106.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2. Poisson problem, uniformly refined grids in three dimensions, timing of the multigrid solution on ASCI Blue Pacific.

<table>
<thead>
<tr>
<th>time</th>
<th>processors</th>
<th>nodes</th>
<th>1</th>
<th>4</th>
<th>16</th>
<th>64</th>
<th>128</th>
<th>256</th>
</tr>
</thead>
<tbody>
<tr>
<td>1089</td>
<td>5.08 1.27 0.72 0.64 0.84 1.30</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1662</td>
<td>5.85 2.01 0.97 0.72 0.86 1.33</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2745</td>
<td>10.7 3.26 1.37 0.85 0.94 1.38</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4834</td>
<td>20.3 5.84 2.01 1.08 1.08 1.46</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8915</td>
<td>39.8 10.9 3.38 1.42 1.26 1.56</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16948</td>
<td>78.5 39.7 5.68 2.08 1.66 1.78</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>32788</td>
<td>157 77.7 10.7 3.34 2.30 2.14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>64251</td>
<td>20.7 5.97 3.62 2.80</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>126810</td>
<td>10.9 6.14 4.12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3. Poisson problem, adaptively refined grids in two dimensions, timing of the multigrid solution on a Cray T3E-600.

<table>
<thead>
<tr>
<th>time</th>
<th>processors</th>
<th>nodes</th>
<th>1</th>
<th>4</th>
<th>16</th>
<th>64</th>
<th>128</th>
<th>256</th>
</tr>
</thead>
<tbody>
<tr>
<td>35937</td>
<td>291 85.6 29.6 11.2 7.61 5.94</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50904</td>
<td>423 129 41.0 14.8 10.1 7.17</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>89076</td>
<td>405 236 71.2 24.6 14.6 9.98</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>189581</td>
<td>154 49.7 29.2 17.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>460421</td>
<td>109 61.1 35.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1201650</td>
<td>142 77.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3251102</td>
<td>345 188</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4. Poisson problem, adaptively refined grids in three dimensions, timing of the multigrid solution on a Cray T3E-600.

<table>
<thead>
<tr>
<th>nodes</th>
<th>1</th>
<th>4</th>
<th>4 uni proc</th>
<th>16</th>
<th>64</th>
<th>256</th>
<th>512</th>
</tr>
</thead>
<tbody>
<tr>
<td>59049</td>
<td>210.6</td>
<td>89.07</td>
<td>85.97 33.91</td>
<td>12.13</td>
<td>20.95</td>
<td>32.3</td>
<td></td>
</tr>
<tr>
<td>452709</td>
<td>2989</td>
<td>1462</td>
<td>1385 556.6</td>
<td>198.1</td>
<td>63.13</td>
<td>37.77</td>
<td></td>
</tr>
<tr>
<td>2421009</td>
<td>2290</td>
<td>768.5</td>
<td>458.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 5. Poisson problem, uniformly refined sparse grids in ten dimensions, timing of the iterative solver on ASCI Blue Pacific.
Table 6. Convection-diffusion problem, uniformly refined sparse grids in three dimensions, timing of the iterative solution on a Cray T3E-600.

<table>
<thead>
<tr>
<th>level</th>
<th>nodes</th>
<th>1/h</th>
<th>processors</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>81</td>
<td>4</td>
<td>0.04</td>
</tr>
<tr>
<td>3</td>
<td>225</td>
<td>8</td>
<td>0.25</td>
</tr>
<tr>
<td>4</td>
<td>593</td>
<td>16</td>
<td>1.31</td>
</tr>
<tr>
<td>5</td>
<td>1505</td>
<td>32</td>
<td>10.0</td>
</tr>
<tr>
<td>6</td>
<td>3713</td>
<td>64</td>
<td>52.5</td>
</tr>
<tr>
<td>7</td>
<td>8961</td>
<td>128</td>
<td>158</td>
</tr>
<tr>
<td>8</td>
<td>21249</td>
<td>256</td>
<td>473</td>
</tr>
<tr>
<td>9</td>
<td>49665</td>
<td>512</td>
<td>1426</td>
</tr>
<tr>
<td>10</td>
<td>114689</td>
<td>1024</td>
<td>4112</td>
</tr>
</tbody>
</table>

Table 7. Convection-diffusion problem, uniformly refined sparse grids in three dimensions, timing of the iterative solution on ASCI Blue Pacific.

<table>
<thead>
<tr>
<th>level</th>
<th>nodes</th>
<th>1/h</th>
<th>processors</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>593</td>
<td>16</td>
<td>0.67</td>
</tr>
<tr>
<td>5</td>
<td>1505</td>
<td>32</td>
<td>3.97</td>
</tr>
<tr>
<td>6</td>
<td>3713</td>
<td>64</td>
<td>23.33</td>
</tr>
<tr>
<td>7</td>
<td>8961</td>
<td>128</td>
<td>71.92</td>
</tr>
<tr>
<td>8</td>
<td>21249</td>
<td>256</td>
<td>208.8</td>
</tr>
<tr>
<td>9</td>
<td>49665</td>
<td>512</td>
<td>619.0</td>
</tr>
<tr>
<td>10</td>
<td>114689</td>
<td>1024</td>
<td>1845</td>
</tr>
<tr>
<td>11</td>
<td>262145</td>
<td>2048</td>
<td>2184</td>
</tr>
<tr>
<td>12</td>
<td>593921</td>
<td>4096</td>
<td>1766</td>
</tr>
<tr>
<td>13</td>
<td>1335297</td>
<td>8192</td>
<td>888.1</td>
</tr>
</tbody>
</table>

Tables 1 and 2 show execution times and scaling for uniform grids in two and three dimensions with a multigrid solver. These times give the wall clock times for the solution of the equation system, on different levels of grids and on different numbers of processors. We assume a constant number of iterations within a nested iteration. We consider also adaptively refined grids of the unit square and the unit cube. The grids are refined toward the corner \( \vec{0} \). Tables 3 and 4 depict times in the adaptive grid refinement case of the Poisson equation in two and three dimensions.

\[-\Delta u + \vec{v} \cdot \nabla u = f\]  \hspace{1cm} (11)

A convection-diffusion problem (11) on an uniform, ten-dimensional sparse grid can be found in Table 5. Adaptively refined three dimensional sparse grid solvers are shown in Table 6 and 7, where both hardware platforms can be compared.

We obtain a scaling of about a factor 2 from one level to the next finer level, which means 2 times more unknowns on the next finer level. Increasing the number of processors speeds up the computation accordingly. Again we observe scalability of the algorithm. In order to use all processor efficiently the grid has to be fine enough, i.e. it has to possess a large number of unknowns.
10 Conclusion

In this paper we gave a survey of the basic ingredients of an efficient solver for self-adjoint elliptic PDEs, i.e. multilevel solvers, adaptive grid refinement and parallelisation. We focused on the interplay between these ingredients and tried to illustrate how they can be glued together into an adaptive parallel multilevel method. Here we proposed the application of hash storage techniques for data management and the use of space-filling curves for load balancing in the parallel version of the algorithms and presented a version of a parallel adaptive multilevel method based on these approaches.

Note finally that load balancing for adaptive multilevel solvers with space filling curves can be obtained (after slight modifications) in an analogous way and with analogous results also for the case of general unstructured grids. This is actual work in progress.

Acknowledgments

The computing resources were generously provided by the John von Neumann Institute for Computing (NIC) at Research Centre Jülich (Cray T3E-600) and the Institute for Scientific Computing Research (ISCR) of the Lawrence Livermore National Laboratory (ASCI Blue Pacific).

References


High Performance FEM Simulation via FEAST and Application to Parallel CFD via FEATFLOW

Christian Becker, Sven H. M. Buijssen, Susanne Kilian, and Stefan Turek

Institut for Applied Mathematics
University of Dortmund, 44227 Dortmund, Germany
E-mail: featflow@featflow.de

The aim of this project is the development of efficient and robust numerical methods for the simulation of PDEs, particularly for incompressible flows in industrial applications. Since corresponding configurations lead to huge systems of nonlinear equations with fully nonstationary behaviour, the CPU requirements are so high such that special data and matrix structures and hardware-oriented implementation techniques have to be realized. The resulting FEAST software is able to exploit a significant percentage of the potentially available computing power of more than 1 Gflop/s (per single processor) in combination with powerful FEM discretization and parallel multigrid solution techniques.

1 Description of the Project

The aim of this project is to develop efficient numerical methods and implementation techniques for the simulation of PDEs with special emphasis on complex flows which are described by the incompressible Navier-Stokes equations. These components have to be integrated into a CFD software package for industrial applications which is capable of predicting the flow in complex situations. Preliminary results of industrial partners based on simulations with different CFD codes and our own experience show that existing research codes and particularly commercial codes have severe problems to provide sufficiently reliable predictions for important flow quantities in the fully nonstationary case. Even on today’s supercomputers, the required CPU time is often too large while the results are still too inaccurate.

In Figure 1, the results are shown for a typical steady 3D turbulence calculation with the RNG-\(k-\epsilon\) model which costs 6.5 h CPU time on a SGI Origin2000 with 6 processors.

<table>
<thead>
<tr>
<th>Quantity</th>
<th>Exp.</th>
<th>Sim.</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Drag ((c_d))</td>
<td>0.165</td>
<td>0.317</td>
<td>92 %</td>
</tr>
<tr>
<td>Lift ((c_l))</td>
<td>-0.083</td>
<td>-0.127</td>
<td>53 %</td>
</tr>
</tbody>
</table>

Figure 1. Experiment vs. CFD simulation by DaimlerChrysler obtained with commercial software
for only about 500,000 grid cells. The significant differences between the results from simulation and experiment indicate that much finer computational meshes are needed which in turn will lead to even larger CPU time. Since these codes usually do not contain optimal multigrid solvers, the ratio between number of grid points and required CPU time increases more than linearly, such that 10 times more unknowns are expected to require at least 20 - 30 times more CPU time. The shown geometrical configuration is still quite simple and nonsteady flow behaviour is not considered at all, so it can be imagined that the simulation of realistic time-dependent flow behaviour is impossible for most of the existing CFD tools, even on recent supercomputers.

Based on the research codes FEAST\(^1\) and FEATFlow\(^{10}\), a new simulation tool for industrial flow configurations is being created which includes modern numerical and algorithmic components for adaptive control of the discretization in space and time, special FEM discretizations and highly efficient multigrid as well as nonlinear iteration techniques.\(^8\) Besides these mathematical tasks we additionally have to apply improved software technologies and implementation strategies which are adapted to modern processor design.

## 2 Computational Bottlenecks and the FEAST Software

One of the main components in iterative solvers are matrix-vector (MV) applications. Hereby, sparse concepts are the standard techniques in FEM codes (and others): Depending on the programming language, the matrix entries plus index informations are stored as long arrays, containing the 'nonzero entries' only. For an overview of such techniques, see for instance SPARSKIT\(^{11}\) and the literature cited therein. While this sparse approach can be applied for general meshes and arbitrary numberings of unknowns, no explicit advantage of (possible) highly structured parts can be gained. Consequently, a massive loss of performance compared to the possible peak rates may occur since – at least for large problems with more than 100,000 unknowns – no ‘caching in’ and ‘pipelining’ can be exploited. In this case, the higher cost of memory access will dominate the resulting MFlop/s rates.

To demonstrate this failure, we show some examples for the FEATFlow code which is one of the most efficient simulation tools for incompressible flow on general domains.\(^5\) We apply FEATFlow to a car configuration (ASMO2D) and measure the MFlop/s rates for the MV multiplication inside of the multigrid solver for the momentum equation.\(^8\) The results in Table 1 show that the computational performance is weak and depends strongly on the problem size and the kind of numbering – Two Level ordering, Cuthill-McKee, STOchastic - while the number of arithmetic operations and memory accesses are identical.

The FEAST software is conceptualized to combine such highly tuned linear algebra tools with sophisticated FEM simulation strategies.\(^9\) In contrast to many other approaches which often aim to develop flexible software for research or educational purposes, the FEAST software is designed for high performance applications with industrial background, particularly for Computational Fluid Dynamics (CFD). Consequently, our emphasis lies more on ‘efficiency’ and ‘robustness’ and less on aspects like ‘easy implementable’, ‘flexible’ or ‘most modern programming language’. One of the most important principles in FEAST is the consequent application of (recursive) Divide and Conquer strategies: The
solution of a ‘global’ problem is recursively split into smaller independent subproblems on patches as part of the complete set of unknowns. While on certain ‘anisotropic’ parts the usual sparse techniques may be applied (for instance, if local adaptivity is employed), we try to exploit the much higher performance possible on the other, highly structured patches. Consequently, the intention is to minimize the number of ‘sparse areas’ and to apply preferably most numerical linear algebra tasks on such ‘structured patches’ via the SparseBandedBLAS techniques. Hence, the major tasks for realizing such a simulation tool are:

1. Design of the ‘skeleton’ for the recursive splitting into local/global levels.
2. Implementation of the typical FEM facilities on the ‘low level’ patches.

In view of their typically excellent convergence behaviour, multigrid methods seem to be most suited for the solution of many PDEs. However, as the previous examples have shown, multigrid on general domains has often poor computational efficiency, at least if the implementation is based on the standard sparse techniques. As a result from our performance measurements, the realistic MFlop/s rates are often in the range of 1 MFlop/s only, even on modern high performance workstations. Moreover, the linear relationship between problem size and CPU time is hardly realizable, due to the problem-size dependent performance rates of the sparse components. Additionally, the robust treatment of complex mesh structures with locally varying details is often hard to achieve by typical ‘Black Box’ components.

Concerning parallelization some further serious problems occur: The parallel efficiency is often ‘bad’ and the achieved performance is far below peak, since the solution of the coarse grid problem leads to a large communication overhead. Furthermore, the relation between ‘local’ arithmetic operations and ‘global’ data transfer is poor, in general. Besides these computational aspects, the parallelization of the global smoothers – SOR, ILU – cannot be done efficiently because of their inherent recursive character. So, smoothing can only be performed blockwise, which may lead to a deterioration of the convergence rates. Further, the behaviour of such blockwise smoothing is hard to predict for complicated geometries with local and/or global anisotropies. Motivated by these facts, we developed the more general strategy ScARC for solving discretized PDEs aiming for the following goals:

The parallel efficiency shall be high due to a non-overlapping decomposition and a low communication overhead. The convergence rates are required to be independent of the mesh size \( h \), the complexity of the domain and the number of subdomains \( N \), and they shall be in the range of typical multigrid convergence rates (as \( \rho_{MG} \sim 0.1 \)). Further,
the method shall be easily implementable and uses only existing standard methods. The approach shall accomplish the treatment of complicated geometries with local anisotropies (large aspect ratios) without impairment of the overall convergence rates.

The underlying idea is to ‘hide recursively all anisotropies in single subdomains’ combined with ‘block Jacobi/Gauß-Seidel smoothing’ within standard multigrid. This approach is based on the numerical experience that these ‘simple’ block-oriented schemes perform well as soon as all occurring anisotropies are locally hidden, i.e., if the local problems on each block are solved (more or less) exactly. This procedure ensures the global robustness. On the other hand, this also means that the local solution quality in each block can significantly improve the global convergence behaviour. These ideas are combined with hierarchical data and matrix structures, which exploit tensorproduct-like meshes on each macro to achieve high performance rates for the linear algebra components in the local (multigrid) solvers. Consequently, all solution processes are recursively organized via sequences of more ‘local’ steps until the lowest level is reached, for instance a single macro with the described generalized tensorproduct mesh. The complete SCARC approach\textsuperscript{13} can be characterized as follows:

1. Scalable (with respect to ‘quality and number of local solution steps at each stage’)
2. Recursive (‘independently’ for each stage in the hierarchy of partitioning)
3. Clustering (for building blocks via ‘fixed or adaptive blocking strategies’)

### 3 Numerical Results

We perform some test calculations for the Pressure-Poisson problem. Figure 2 shows a typical macro decomposition (= coarse grid) for the 2D configuration NCC1701-D and a zoom of some macro element (= 1 quadrilateral) directly at the boundary of the vessel. During the mesh generation process, we have applied locally anisotropic refinement in normal direction towards the boundary such that boundary layers can be better resolved while at the same time the tensorproduct topology is locally preserved.

![Figure 2. Macro decomposition and anisotropically refined macro (3 refinement levels) for the NCC1701-D configuration](image)

Due to the local grid distortion, special multigrid components have to be employed which work efficient and robust w.r.t. such mesh anisotropies. In the framework of
SCaRC as generalized multigrid approach, particularly on parallel platforms, the corresponding local problems (on each refined macro of the coarse grid) are treated with special linewise Gauss-Seidel schemes as smoothing operations inside of an optimized local multigrid solver. Table 2 shows typical (parallel) multigrid rates for the solution of Pressure-Poisson problems within the framework of a nonstationary Navier-Stokes approach via discrete projection methods. The global multigrid convergence rates are more or less independent of the degree of (local) mesh distortion while the resulting CPU times, i.e. the MFlop/s rates, are not yet completely optimized.

The global multigrid convergence rates in Table 2 show the robustness and efficiency of this parallel SCaRC approach. The resulting CPU times in the right most table already indicate that high efficiency rates will be achieved, which will be further improved by optimizing the administration of the local multigrid solver and particularly by including the corresponding SPARSEBANDEDBLAS techniques: At the moment, the MFlop/s rates for the local multigrid solvers on each tensorproduct mesh are of order 25 MFlop/s and shall be improved by the optimized SPARSEBANDEDBLAS up to 50–100 MFlop/s per processor.

While the numerical components seem to be sufficiently efficient, the computational efficiency in terms of the MFlop/s rates still requires further improvement. Table 3 shows the recent efficiency rates for components from the optimized SPARSEBANDEDBLAS library; here, we can additionally distinguish between the cases of general and of constant matrix entries. It is remarkable that special algorithms and implementation techniques had to be developed for the TGS preconditioner on the Cray T90. It turns out that multigrid, even with this very robust smoother, can be very efficiently performed. Additionally, these results show that the actual MFlop/s rates can significantly depend on the local problem size due to different cache exploitation. This fact can be employed when applying the global SCaRC solver since the choice of the local number of macros and the corresponding problem size can be adapted to these effects.

4 Parallel Version of the Existing FEATFlow Solver

In the meantime, our work on a parallel version of FEATFlow for CFD problems has been completed. The numerical methods applied have been described by Turek and already been implemented sequentially within the FEATFlow package. Before giving some examples of problems with industrial background this implementation has recently been applied to, let us first briefly describe the overall mathematical
Table 3. MFlop/s rates for different problem sizes (#NEQ = number of grid points) for sparse indexed DAXPY (linear access, distributed access in parenthesis), matrix-vector applications (SBB), and complete multigrid solver (TGS smoother); additionally, we can distinguish between variable (V) and constant (C) matrix entries (Poisson problems on orthogonal meshes); special libraries are used on the Cray machines.

<table>
<thead>
<tr>
<th>2D case</th>
<th>N</th>
<th>DAXPY-El</th>
<th>SBB-V</th>
<th>SBB-C</th>
<th>MG-V</th>
<th>MG-C</th>
</tr>
</thead>
<tbody>
<tr>
<td>DEC 21264</td>
<td>65²</td>
<td>205 (178)</td>
<td>538</td>
<td>795</td>
<td>370</td>
<td>452</td>
</tr>
<tr>
<td></td>
<td>(667 MHz)</td>
<td>257²</td>
<td>224 (110)</td>
<td>358</td>
<td>1010</td>
<td>314</td>
</tr>
<tr>
<td>‘ES40’</td>
<td>1025²</td>
<td>78 (11)</td>
<td>158</td>
<td>813</td>
<td>185</td>
<td>401</td>
</tr>
<tr>
<td>Hitachi</td>
<td>65²</td>
<td>173 (82)</td>
<td>238</td>
<td>391</td>
<td>191</td>
<td>266</td>
</tr>
<tr>
<td>(375 MHz)</td>
<td>257²</td>
<td>143 (29)</td>
<td>243</td>
<td>388</td>
<td>198</td>
<td>260</td>
</tr>
<tr>
<td>‘SR8000’</td>
<td>1025²</td>
<td>144 (7)</td>
<td>226</td>
<td>390</td>
<td>200</td>
<td>267</td>
</tr>
<tr>
<td>AMD K7</td>
<td>65²</td>
<td>203 (195)</td>
<td>101</td>
<td>556</td>
<td>122</td>
<td>355</td>
</tr>
<tr>
<td>(850 MHz)</td>
<td>257²</td>
<td>29 (27)</td>
<td>78</td>
<td>241</td>
<td>72</td>
<td>166</td>
</tr>
<tr>
<td>‘ATHLON’</td>
<td>1025²</td>
<td>31 (10)</td>
<td>64</td>
<td>236</td>
<td>58</td>
<td>126</td>
</tr>
<tr>
<td>Cray T3E</td>
<td>65²</td>
<td>156 (41)</td>
<td>82</td>
<td>392</td>
<td>85</td>
<td>282</td>
</tr>
<tr>
<td>(666 MHz)</td>
<td>257²</td>
<td>47 (16)</td>
<td>73</td>
<td>294</td>
<td>70</td>
<td>172</td>
</tr>
<tr>
<td>‘SciLib’</td>
<td>1025²</td>
<td>30 (7)</td>
<td>47</td>
<td>283</td>
<td>52</td>
<td>142</td>
</tr>
<tr>
<td>Cray T90</td>
<td>65²</td>
<td>627 (500)</td>
<td>855</td>
<td>415</td>
<td>183</td>
<td>163</td>
</tr>
<tr>
<td>(440 MHz)</td>
<td>257²</td>
<td>698 (544)</td>
<td>933</td>
<td>887</td>
<td>276</td>
<td>258</td>
</tr>
<tr>
<td>‘SDTSOL’</td>
<td>1025²</td>
<td>699 (553)</td>
<td>939</td>
<td>1092</td>
<td>423</td>
<td>419</td>
</tr>
</tbody>
</table>

background. The underlying model are the instationary incompressible 3D Navier-Stokes equations. They are discretized by a finite element approach ($\tilde{Q}_1/\tilde{Q}_0$ ansatz). Stabilization of the convection term is done by applying an upwind scheme. For time discretization the Fractional-step-$\theta$ scheme is used. In order to realize adaptive time stepping, its solutions are compared with those from a Crank-Nicolson scheme. The nonlinear subproblems in each time step are split by applying a discrete projection method (which can be interpreted as a discrete analogue of schemes proposed by Chorin and Van Kan). Using a fixed point defect correction method the resulting nonlinear Burgers equations are linearized and finally solved by a parallel multigrid method. The remaining discrete Pressure-Poisson equations are solved using a parallel conjugate gradient scheme which is preconditioned by multigrid.

In order to parallelize the multigrid method the mesh is split into parallel blocks by a graph-oriented partitioning tool (see example in figure 8). Consistency with the sequential algorithm (MV application, restriction, prolongation) is guaranteed through local communication between at most two parallel blocks (this is possible because of the face oriented $\tilde{Q}_1/\tilde{Q}_0$ ansatz). As mentioned earlier, the inherent recursive character of global smoothers prevents a direct parallelization. Therefore, the global smoothing is replaced by smoothing within each parallel block only (block ILU smoother). To minimize the communication overhead solving the coarse grid problem, it is treated on a single processor with an optimized sequential algorithm. The cost are two global communications (setting up right side vector and propagation of solution vector).

4.1 An Industrial Problem

The program has been applied to a bunch of problems, one of them arising from steel industry. The behaviour of molten steel within a mould was supposed to be simulated. A
mould is a component of a slab continuous casting production set (Figure 3). It is one of the critical sections during production of slabs. Molten steel pours into the mould through a small nozzle in the upper part (Figure 5), solidifies due to cooling at the surrounding walls and leaves the mould at its bottom. The final aim is to simulate the complete production process including cooling down and electro-magnetic stirring. For now, these additional technical measures are neglected and we concentrate only on the simulation of the complex flow patterns within the mould.

Figure 4 shows the problem’s coarse grid. It only contains moderate aspect ratios ($AR = 30$). After four refinements we got a problem size of 29,706,240 unknowns in space. It took 2900 time steps to reach the endpoint $T = 36s$. Using 256 processors on the Cray T3E-1200 this simulation was performed within 20 hours. With 1 pre- and post-smoothing step for the linearized Burgers equations the average multigrid convergence rate was 0.03. For the discrete Pressure-Poisson equations the multigrid preconditioner performed 4 pre- and post-smoothing steps, leading to typical convergence rates for the preconditioned conjugate gradient method of 0.31.

Some snapshots of the resulting nonstationary non-periodic flow patterns can be seen in Figure 6 (isosurfaces of vertical velocity) and Figure 7 (particle tracing).

4.2 Deterioration of Solver Due to Block Smoothers

As explained, there is a price to pay for the ‘easy’ way of parallelizing the smoothing process by replacing global smoothing with block smoothers. A deterioration of convergence rates can be expected, but it is hard to predict the degree of deterioration for complicated geometries with local and/or global anisotropies. Therefore, a complex problem derived from automotive aerodynamics was studied: flow around a 3D sample car (denoted as ASMO3D). Because the industrial partner was/is interested in predictions on lift and drag coefficients on the car’s surface, a body-fitted mesh had to be used. As a result, the solvers had to cope with large areas of distorted hexahedrons and locally high aspect ratios.

The coarse grid consisted of 840 elements and was refined three times such that a

![Figure 3. Scheme slab continuous casting production set](image1)

![Figure 4. Coarse grid used. Four times uniformly refined it consists of 2,949,120 elements.](image2)

![Figure 5. Magnified view of the upper domain part with inflow nozzle](image3)
Figure 6. Flow pattern of molten steel within a mould. Visualization using isosurfaces of vertical velocity ($\Delta t = 5, 5s$)

Figure 7. Flow pattern of molten steel within a mould. Particle tracing visualization ($\Delta t = 3, 5s$)
A problem with 4,368,576 unknowns in space was gained. A time period $T = [0, 0.5]$ was simulated with different numbers of processors ranging from 1 to 64. For sample grid partitions into 8 and 64 parallel blocks respectively see Figure 8. Solving linearized Burgers equations required 1 pre- and post-smoothing step, discrete Pressure-Poisson equation as much as 8.

These scaling tests were performed on several architectures: Alphacluster ALiCE (Wuppertal), Cray T3E-1200, Linuxcluster (Heidelberg), Sun Enterprise 3500 (Dortmund). Unfortunately, the run time scaling for this problem (see Figure 9) is rather disappointing. Increasing the number of processors from 4 to 64 on the Cray T3E-1200 gives only a speedup of approximately 5. Scaling on the remaining architectures was even worse.

It turned out that this is not mainly caused by an increasing communication overhead, but by the fact that the average number of iterations solving the discrete Pressure-Poisson equation increases from 2.2 (1 processor) to 6.2 (64 processors). Trying other solving methods (classical multigrid, cg with additive instead of multiplicative multigrid preconditioning) and smoothers (Jacobi, SOR) gave even worse results. Thus, already at moderate aspect ratios the solution process suffers significantly from the deterioration of smoothing property due to the use of ‘simple’ block smoothers.

5 Outlook

The central point of this project is the development of mathematical components – FEM discretizations, adaptivity and multigrid solvers – and their realization in a software pack-
age which directly includes tools for parallelism and hardware-adapted high-performance in low level kernel routines. The code generation uses the new FEAST software in order to achieve highest computational efficiency. These software developments can be viewed as ‘basic research’ in the field of mathematical software for PDEs. It is the special goal in this project to realize and to optimize the SPARSEBANDED BLAS concepts for specific parallel computers (Cray T3E, Cray T90, Sun HPC machines, Linux cluster) and to adapt the mathematical components to complex configurations. Since the corresponding geometries may lead to severe mesh distortions, for instance near boundaries, special multigrid components for highly-stretched body fitted meshes have been developed which allow high numerical efficiency and robustness in the multigrid-like SCARC approach.

At the moment, we are completing the 2D Navier-Stokes solver which is based on conforming bilinear FEM and discrete projection methods. This code will be integrated into FEAST such that the parallel/sequential high-performance tools of the SPARSEBANDED BLAS will be directly available. Besides that, we will continue our work with the parallel 3D adaptation of the FEATFLOW solver which is presently applied on several parallel computers to prototypical configurations similar to the shown geometries. This parallel 3D code is our candidate for all further developments which aim to incorporate the high-performance FEAST techniques into this CFD tool in order to achieve highest computational efficiency on modern computers in combination with the ‘best’ numerical approaches.

References

11. SPARSKIT (by Y. Saad), www.cs.umn.edu/Research/arpa/SPARSKIT.

Peter Bastian, Klaus Johannsen, Stefan Lang, Sandra Nägele, Christian Wieners, Volker Reichenberger, Gabriel Wittum, and Christian Wrobel

IWR/Technical Simulation, University of Heidelberg
Im Neuenheimer Feld 368, 69120 Heidelberg, Germany

The program package $UG$ provides a software platform for discretizing and solving partial differential equations. It supports high level numerical methods for unstructured grids on massively parallel computers. Various applications of complex up to real-world problems have been realized, like Navier-Stokes problem with turbulence modeling, combustion problems, two-phase flow, density driven flow and multi-component transport in porous media. Here we report on new developments for a parallel algebraic multigrid solver and applications to an eigenvalue solver, to flow in porous media and to an simulation of Navier-Stokes equations with turbulence modeling.

1 Introduction

In many cases the modeling of physical and technical problems leads to a description by partial differential equations. Due to the complexity of the equations and the geometry involved often the mathematical problem can be treated only by numerical methods. Appropriate discretizations lead to large systems of (non-linear) equations to be solved. To make the numerical treatment feasible, advanced numerical methods in conjunction with High Performance Computing have to be applied. Unstructured grids, adaptivity, multigrid methods, and parallelism have proven to be an efficient approach. Unstructured grids are required by the complex geometries. Adaptivity has to be used to reduce the computational cost especially for three-dimensional simulations. Finally the use of both multigrid methods and High Performance Computing on massively parallel MIMD machines is indispensable to reduce the computational time.

To bring together the features mentioned above the program package $UG$ has been designed during the last decade. It provides a platform for the discretization and the solution of partial differential equations on parallel computers.

The effort for building up a new application (in lines of code) on top of the UG platform is illustrated in figure 1. The shown example is two-phase flow in porous media. The coding of the problem specific program part (pm) has only the extend of about 12% of the total amount. All other parts are problem independent and may be used for other problem class implementations as well. The used symbols have the following meaning:

- pm: two-phase flow in porous media
- ug: kernel library for (multi)grid management, numerical procedures and graphics,
- ddd: the parallel programming model (dynamic distributed data),

503
Figure 1. Effort to implement a parallel adaptive application in percent of lines of code (LOC).

- ppif: small and fast message passing interface (to shared memory on Cray T3E),
- dom: handling of the domain geometry,
- dev: devices drivers for X11, postscript and other devices.

This shows clearly the advantage of code reuse, when proper abstractions are introduced. Especially in the case of parallel adaptive applications a huge amount of work (and therefore development time) can be saved through the platform design. Furthermore the scientists involved in the problem class development do not need to treat detailed problems of parallel implementation, since only few lines of code have to be added to a problem class to get it running in parallel.

The kernel of UG has now reached a mature state and state-of-the-art numerical methods can be applied to complex and real-world problems. Applications to the Navier-Stokes equation with turbulence modeling, to combustion problems, to two-phase flow, density driven flow and multi-component transport in porous media and to flow in fractured rock have been realized, see also\textsuperscript{2,3}.

Nevertheless further work has to be done in various directions, as to mention

- improvement of dynamical load balancing\textsuperscript{4},
- development of new parallel numerical algorithms,
- implementation of new applications.
- improvement of stability and efficiency

In this paper we report on some of these aspects. Section 2 reports to an porous media application with special emphasis on parallel performance. In Section 3 recent developments of the parallel algebraic multigrid solver\textsuperscript{3} are discussed. Section 4 is devoted to the parallel solving of eigenvalue and plasticity problems. Finally in Section 5 an application to the Navier-Stokes equations with turbulence modeling is given.
2 Two-Phase Flow

This section evaluates the speedup when a fixed sized problem should be computed fast for productivity reasons. Fast computation of a given problem is for example advantageous, if parameter studies are necessary for optimization.

Two phase flow problems with phases consisting of liquids with different viscosities form fronts which tend to be unstable. These instabilities can be seen in real live or physical experiments as fingers. These fingering results from a statistical behaviour, where interactions between the two phases on a very small scale influence the structure of the viscous front on a visible level.

As test scenario we choose a two-phase flow problem, where water displaces oil in a cubic channel (see figure 2). Interesting are the different fingering patterns evolving from different viscosity parameters of the two phases. Since the resolution of the highly nonlinear effects at the front must be quite fine, the combination of parallelism and adaptivity tends to be an optimal approach. As load balancer a simple Recursive Orthogonal Bisection Method (RCB) is applied to distribute the coarse grid levels in flow direction. Thus important coupling information of the problem is kept processor local to avoid degradation of the solver's convergence rate. No load balancing need to be done during computation, since load remains equally distributed between the processors when the front moves.

Table 2 shows the speedup of the displacement example for the first 34 timesteps. The minimal amount of serial main memory needed is about 2 GB to allow the front to finger after about 100 timesteps. This corresponds to about 1 million unknowns which have to treated in each timestep. Results for one to eight processors are computed on
a IBM SP-256, from 8 to 256 processors on Cray T3E at NIC. Timings are scaled by
a constant factor, which is gained from the comparison of the 8 processor run on both
machines.

The nonlinear solution process scales quite nice, giving a speedup of 173 for 256 pro-
cessors. Speedup losses have two reasons: The first is the introduction of communication
in the linear solver, when going from one to two processors, speedup loss is 20%. Second
the convergence rates becomes with increasing processor count more worse. The cycle
time of one linear iteration step scale good, which shows a stable scaling of the machine
itself.

Grid adaption shows a scaling which is not comparable to that of the nonlinear solution.
A speedup factor of 12.6 could be achieved in the 256 processor configuration. Analyzation
of the losses has shown, that the grid manager itself has a better speedup behaviour, but the
underlying programming modell DDD has a bottleneck when the interfaces for communi-
cation are rebuild after grid adaption. This process can be speeded up by reimplementation
of the DDD interface module.

The total efficiency of the 256 processor run is 0.57, which can be treated as a quite
satisfying overall result for a fixed sized problem.

3 Parallel Algebraic Multigrid FAMG

Besides classical (or geometric) multigrid methods purely algebraic multigrid methods
(AMG) are desirable for many reasons; AMG considers only the stiffness matrix but not
any geometric information such as elements or coordinates. AMG has the potential to solve
complicated problems better than other methods. Even in a geometric multigrid AMG
can be applied namely as the coarse grid solver. A third point of interest is the coupling
of multigrid methods with already existing programs. The smallest possible interface—
the matrix itself—is already feasible for AMG and the solver has no interaction with the
geometry. A major drawback of many existing AMG methods is the missing or even
impossible parallelization.

The starting point for our parallel AMG is the filtering AMG (FAMG) by WAGNER5, 6.
The main idea is to choose for each node good parent pairs of nodes for eventual elimina-
tion which ensures a certain filtering condition and leads to exactness on a given subspace.
The best pairs are selected to eliminate the corresponding nodes; this parent nodes persist
on the next coarser grid level and restriction and prolongation matrix entries with individ-
ually calculated values are installed between the nodes and their parents. The recursive
application of this process yields a grid hierarchy on which standard multigrid algorithms
can be realized.

For parallelizing this FAMG several additional steps have to be done. Since the elim-
ination of a node influences its neighborhood, and thus the following selections for elimi-
nation, the cardinal point for the parallel FAMG will be to break up this sequential order
in a suitable way. Due to the locality of the direct influence of a node our approach will
divide the nodes into two classes: those which are influenced directly by nodes on other
processors and the rest. Whereas the latter can be eliminated locally on each processor,
the processing of the first ones needs a partly synchronization. Therefore a parallel graph
coloring method is used7. For further details see3.

Now we present first results of the new parallel FAMG. We examine the Laplace op-
erator on the unit square with a tensor product mesh (called structured case) and on the
shape of Lake St. Wolfgang (Austria) with an unstructured mesh from a grid generator. The structured mesh is distributed by a special load balancer to achieve tensor product like processor configurations, the unstructured mesh is distributed by recursive coordinate bisection (RCB). For anisotropic model problems the equation $\varepsilon u_{xx} + u_{yy} = f$ is solved with the anisotropy parameter $\varepsilon$.

Tab. 2 shows that the convergence rate is rather independent from the mesh width respectively the number of nodes. This is essential to solve large problems efficiently.

<table>
<thead>
<tr>
<th>unknowns</th>
<th>conv. rate</th>
<th>nr. it.</th>
<th>time [sec]</th>
<th>unknowns</th>
<th>conv. rate</th>
<th>nr. it.</th>
<th>time [sec]</th>
</tr>
</thead>
<tbody>
<tr>
<td>16641</td>
<td>0.053</td>
<td>7</td>
<td>6.6</td>
<td>14095</td>
<td>0.057</td>
<td>7</td>
<td>4.6</td>
</tr>
<tr>
<td>66049</td>
<td>0.054</td>
<td>7</td>
<td>9.7</td>
<td>55637</td>
<td>0.123</td>
<td>9</td>
<td>9.4</td>
</tr>
<tr>
<td>263169</td>
<td>0.054</td>
<td>7</td>
<td>17.3</td>
<td>221065</td>
<td>0.206</td>
<td>12</td>
<td>20.7</td>
</tr>
<tr>
<td>1050625</td>
<td>0.057</td>
<td>7</td>
<td>38.1</td>
<td>881297</td>
<td>0.246</td>
<td>14</td>
<td>51.7</td>
</tr>
<tr>
<td>4198401</td>
<td>0.064</td>
<td>7</td>
<td>110.1</td>
<td>3519265</td>
<td>0.256</td>
<td>14</td>
<td>137.2</td>
</tr>
<tr>
<td>16785409</td>
<td>0.068</td>
<td>7</td>
<td>295.3</td>
<td>14065217</td>
<td>0.258</td>
<td>14</td>
<td>391.6</td>
</tr>
</tbody>
</table>

Table 2. The convergence rate depends only weakly upon mesh width (structured grid on a 16x8 processor configuration, unstructured grid on 128 processors).

Fig. 3 shows that the solver is very robust with respect to the variation of the anisotropy parameter. The convergence rate is bounded by 0.15 even for 128 processors and the solution time is quite independent of this parameter; this holds from low load up to full load examples.

Next we evaluate the quality of the parallelization. We consider the speedup where the same problem is solved on different number of processors. Tab. 3 indicates a quite constant convergence rate and the solution time decreases nearly by the factor as the number of processors increases.

We have seen very promising features of the new parallel FAMG. Nevertheless, further model problems (e. g. jumping coefficients and convection) should be inspected to explore...
the benefits and the limitations of the new method. At the end real world problems should be solved.

<table>
<thead>
<tr>
<th>PE</th>
<th>structured isotropic conv. nr. time [sec]</th>
<th>anisotropic $10^{-6}$ conv. nr. time [sec]</th>
<th>unstructured conv. nr. time [sec]</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>0.058 7 220.6</td>
<td>10$^{-4}$ 2 140.7</td>
<td>0.237 13 235.8</td>
</tr>
<tr>
<td>16</td>
<td>0.067 7 135.8</td>
<td>0.053 7 123.2</td>
<td>0.194 12 179.6</td>
</tr>
<tr>
<td>32</td>
<td>0.062 7 81.4</td>
<td>0.046 6 70.7</td>
<td>0.241 13 97.6</td>
</tr>
<tr>
<td>64</td>
<td>0.059 7 58.3</td>
<td>0.043 6 43.7</td>
<td>0.240 13 78.4</td>
</tr>
<tr>
<td>128</td>
<td>0.057 7 38.1</td>
<td>0.044 6 31.8</td>
<td>0.246 14 51.7</td>
</tr>
</tbody>
</table>

Table 3. Speedup on different processor configurations (structured grid with 1 million unknowns, unstructured grid with 880000 unknowns)

4 A Flexible Finite Element Library

In UG, a general finite element library is included which supports various discretizations, and which is especially adapted to problems in nonlinear solid mechanics. Here, the applicability and flexibility of the finite element module is illustrated on the following model problem in eigenvalue computations.

Let $\Omega \subset \mathbb{R}^d$, $d = 2, 3$ be a domain. We solve the eigenvalue problem:

$\text{find } (w_i, \lambda_i) \in H^1(\Omega) \times \mathbb{R}$ such that

$$\int_\Omega \nabla w_i \cdot \nabla v \, dx = \lambda_i \int_\Omega w_i v \, dx, \quad v \in H^1(\Omega).$$

This corresponds to the eigenvalue problem

$$-\Delta w_i = \lambda_i w_i \text{ in } \Omega, \quad w_i \cdot n = 0 \text{ on } \partial \Omega$$

in the strong formulation; in particular, we impose Neumann boundary conditions. Thus, we know a priori the first trivial eigenpair $w_0 \equiv 1$ and $\lambda_0 = 0$. Note that this results in a singular stiffness matrix (which, of course, has severe consequences for the solution process).

The eigenmode approximations are computed by a block inverse iteration with a full Ritz-Galerkin orthogonalization in every step. Starting from initial guesses $\tilde{w}_1, \ldots, \tilde{w}_m$, we perform the following algorithm:

a) Ritz-Galerkin step: We form the small matrices

$$M = \left( \int_\Omega \nabla \tilde{w}_i \cdot \nabla \tilde{w}_j \, dx \right), \quad N = \left( \int_\Omega \tilde{w}_i \tilde{w}_j \, dx \right),$$

508
and we solve the generalized eigenvalue problem
\[ Mx_i = \lambda_i Nx_i, \quad x_i \in \mathbb{R}^m; \]
then, we define approximations by
\[ w_i = \frac{1}{x_i^T N x_i} \sum_{j=1}^{m} x_{ij} \tilde{w}_j, \quad i = 1, \ldots, m. \]

b) Block inverse iteration step: We update the space of test functions by solving
\[ \int_{\Omega} \nabla \tilde{w}_i \cdot \nabla v \, dx = \lambda_i \int_{\Omega} w_i v \, dx, \quad v \in H^1(\Omega), i = 1, \ldots, m \]
and return to a).

For the solution of the linear problems, we use a fixed number of parallel multiplicative multigrid cycles with Krylov acceleration and Gauß-Seidel smoother, where we project in every step the solution onto the space which is orthogonal to \( w_0 \).

As an example, we present the parallel results of the eigenmode computations on a domain representing the surface of a large lake (Bodensee in Germany). The shape of the eigenmode \( w_{11} \) is visualized in Fig. 4 (using GRAPE\textsuperscript{10}), the first 4 of 16 computed eigenvalues are listed in Tab 4.

<table>
<thead>
<tr>
<th>discretization</th>
<th>( P_1 )</th>
<th>( P_2 )</th>
<th>( P_2 )</th>
<th>( P_2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>elements</td>
<td>1620992</td>
<td>101312</td>
<td>405248</td>
<td>1620992</td>
</tr>
<tr>
<td>unknowns</td>
<td>816737</td>
<td>205745</td>
<td>816737</td>
<td>3254465</td>
</tr>
<tr>
<td>( \lambda_0 )</td>
<td>0.00000e-00</td>
<td>0.00000e-00</td>
<td>0.00000e-00</td>
<td>0.00000e-00</td>
</tr>
<tr>
<td>( \lambda_1 )</td>
<td>8.88113e-04</td>
<td>8.88274e-04</td>
<td>8.89473e-04</td>
<td>8.88088e-04</td>
</tr>
<tr>
<td>( \lambda_2 )</td>
<td>2.46599e-03</td>
<td>2.46597e-03</td>
<td>2.46582e-03</td>
<td>2.46576e-03</td>
</tr>
<tr>
<td>( \lambda_3 )</td>
<td>4.77479e-03</td>
<td>4.77490e-03</td>
<td>4.77461e-03</td>
<td>4.77447e-03</td>
</tr>
</tbody>
</table>

Table 4. Comparison of the numerical solution with different discretizations and different mesh sizes of the 4 lowest eigenvalues of the Bodensee.

Figure 4. The 11th eigenmode \( w_{11} \) of the Bodensee.
Table 5. Parallel performance of the eigenvalue computation (average time for the linear solver in every step) for $P_2$-elements. Due to the cpu-time-limit, not sufficiently many steps of the block inverse iteration could be performed on level 4 (32 proc.) and level 5 (128 proc.).

The parallel performance (see Tab. 5) underlines the efficiency of the method, which is asymptotically of optimal complexity and optimal parallel scalability (due to the large coarse grid problem with 6332 elements we obtain optimal multigrid efficiency not in the first refinement step). Since we use a second order discretization and millions of unknowns, the results are reliable due the monotone asymptotic convergence; by simple extrapolation, we expect an accuracy of at least 0.1%.

Nevertheless, fully reliable results can be obtained only by the computation of rigorous eigenvalue inclusions; this is realized using interval arithmetic, and first results (on more simple geometries) are already documented in\textsuperscript{11}.

A second quite different application is a plasticity experiment on a gear geometry. The calculation consists of 27 load steps and is performed using 128 Processors of the Cray T3E-1200 at NIC, see figure 5. In the context of the UG platform this example demonstrates the full functionality included in the software: multigrid consisting of tetrahedra, prims and hexahedra, grid adaption, and dynamic load balancing. This features are combined with a nonlinear solver using a multiplicative multigrid scheme and a discretization based on Prandl-Reuss Plasticity. Since this is a recently performed calculation no detailed timings have been done yet. Next steps will be the analyzation of speedup and scalability behaviour.

5 Navier-Stokes Equations and Turbulence Modelling

The Navier-Stokes library in UG uses a Finite Element based Finite Volume Method with colocated variables. Since a colocated scheme is not stable, a special stabilization scheme is applied which introduces a physical advection correction scheme to couple the momentum equation and the pressure equation. This results in a Laplacian term for pressure in the continuity equation scaled with the mesh size squared and therefore tends to zero as the grid is refined. This physical advection correction scheme called FIELDS was developed by Raw\textsuperscript{12}. The idea is to solve in each element a Finite Difference approximation of the linearized momentum equation at all integration points. The resulting integration
point velocities depend on all corner values of velocity and pressure. After insertion in
the continuity equation a pressure dependence and full coupling of all equations is gained.
This can be done independently at all elements and is therefore very advantageous for
parallelization.

A special problem class in the Navier-Stokes community is turbulence modelling. A
very promising way to simulate turbulent flow characteristics is the so called Large Eddy
Simulation (LES). In contrast to Reynolds averaged turbulence models (RANS), it is not
based on time averages but on local volume averages. This means that large structures have
to be resolved and only the small ones are modelled. In RANS methods all structures are
modelled with the difficulty that a suitable turbulence model is not easy to design. LES
models use the fact that small structures are nearly isotropic and more universal than large
structures and therefore modelling becomes simpler. But to reach the necessary resolution
very fine meshes have to be used and because of this parallel calculations are needed.

The averaged equations for LES are derived by applying filter operators (for example
a volume-average box filter) to the governing equations. For the momentum equation this
results in:

\[
\frac{\partial \bar{u}_i}{\partial t} + \frac{\partial}{\partial x_j} (\bar{u}_i \bar{u}_j) + \frac{\partial \bar{p}}{\partial x_i} - \frac{\partial}{\partial x_j} (\nu \left( \frac{\partial \bar{u}_i}{\partial x_j} + \frac{\partial \bar{u}_j}{\partial x_i} \right)) + \frac{\partial}{\partial x_j} \tau_{ij} = 0
\]

where an overbar denotes an average value. The subgrid scale stress tensor \( \tau_{ij} \) is modelled
by the dynamic model of Germano\textsuperscript{13}. This model, in contrast to the Smagorinsky model\textsuperscript{14}, for which a constant and universal model parameter is assumed, is truly local and hence it is able to reflect approximately local flow phenomena. Finally the model has the form:

\[
\tau_{ij} - \frac{1}{3} \delta_{ij} \tau_{kk} = -2C \Delta^2 |\mathbf{S}| \mathbf{S}_{ij}
\]

with \( \mathbf{S}_{ij} = \frac{1}{2} \left( \frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) \) and \( |\mathbf{S}| = \sqrt{2 \mathbf{S}_{ij}} \).

An interesting problem is the flow around an infinite long square cylinder as described in the workshop of Rodi et al.\textsuperscript{15}. The infinite dimension in axis direction can be numerically realized using periodic boundary conditions since the flow structures in that direction can be assumed to be almost periodic. A sketch of the domain is given in Fig. 6. For implementing periodic boundary conditions in UG each element at the periodic boundary and its partner at the opposite side have to be assigned to the same processor. For this reason the load balancing strategy RCB of CHACO has been modified to handle periodicity. An example for 8 processors can be seen in Fig. 6.

The Reynoldsnumber for this problem was 1000. The grid is built of 800000 hexahedrons which correspond to 3 million unknowns. The nonlinear system is linearized by a Quasi-Newton method and the resulting linear system is solved by the Kryloc subspace method.
method BiCGSTAB with multigrid as preconditioner. As smoother the incomplete LU factorization with $\beta$ modification was employed and the time solver was a diagonally implicit Runge-Kutta-method of second order. In Fig. 7 the solution of the velocity component in x-direction on a line through the midpoint of the cylinder in mean flow direction is shown at 6 successive points of time illustrating the complex behaviour of the flow.

6 Conclusions

With the program package $\textit{UG}$ we follow the strategy of combining advanced numerical methods with the advantages of high-performance computers like a Cray T3E. It has been shown that on the base of the parallel $\textit{UG}$ platform both the development of new, highly efficient numerical algorithms and the solution of complex problems is feasible. The obtained results show clearly that

- the computational time can be reduced significantly,
- for some problems parallel computers are necesssary, due the limitations of memory size on sequential machines,
- High Performance Computing becomes an enormouse factor of productivity, when treating real world problems.

Therefore the availability of HPC computing power will have a strong impact on the research related to practical applications of scientific and industrial importance.

Acknowledgments

The benchmark computations were performed with computer time provided by HLRZ in Jülich and HLRS in Stuttgart.

References

Already published:

**Modern Methods and Algorithms of Quantum Chemistry - Proceedings**  
Johannes Grotendorst (Editor)  
NIC Series Volume 1  
Winterschool, 21 - 25 February 2000, Forschungszentrum Jülich  
ISBN 3-00-005618-1, February 2000, 562 pages  
*out of print*

**Modern Methods and Algorithms of Quantum Chemistry - Poster Presentations**  
Johannes Grotendorst (Editor)  
NIC Series Volume 2  
Winterschool, 21 - 25 February 2000, Forschungszentrum Jülich  
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)  
NIC Series Volume 3  
Winterschool, 21 - 25 February 2000, Forschungszentrum Jülich  
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 Lövenich  
NIC Series Volume 5  
ISBN 3-00-006329-3, August 2000, 145 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)
NIC Series Volume 7
Workshop MPOOL, 18 May 2001, Budapest
ISBN 3-00-007968-8, June 2001, 160 pages

Europhysics Conference on Computational Physics - Book of Abstracts
Friedel Hossfeld, Kurt Binder (Editors)
NIC Series Volume 8
Conference, 5 - 8 September 2001, Aachen
ISBN 3-00-008236-0, September 2001, 500 pages

Quantum Simulations of Complex Many-Body Systems: From Theory to Algorithms - Lecture Notes
Johannes Grotendorst, Dominik Marx, Alejandro Muramatsu (Editors)
NIC Series Volume 10
Winter School, 25 February - 1 March 2002, Rolduc Conference Centre, Kerkrade, The Netherlands
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)
NIC Series Volume 11
Winter School, 25 February - 1 March 2002, Rolduc Conference Centre, Kerkrade, The Netherlands
ISBN 3-00-009058-4, February 2002, 85 pages

All volumes are available online at http://www.fz-juelich.de/nic-series/.