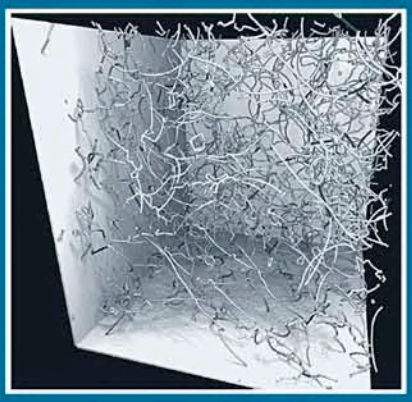
WOODHEAD PUBLISHING IN MATERIALS



Multiscale materials modelling

Fundamentals and applications

Edited by Z. Xiao Guo

CRC

WP

Multiscale materials modelling

Related titles:

Analysis and design of plated structures Volume 2: Dynamics
(ISBN 978-1-84569-116-5)
This second of two volumes on plated structures reviews dynamics, particularly vibration. Chapters review the behaviour of components such as beams and plates together with materials such as steel, ceramic-metal and other composites. Together with its companion volume, this important book summarises the wealth of research on understanding the behaviour of thin-walled (plated) components. It will be an invaluable reference for all those concerned with the design, construction and service life of plated structures.

Durability of composites for civil structural applications
(ISBN 978-1-84569-035-9)
This comprehensive book on the durability of FRP composites will make it easier for the practising civil engineer and designer to use these materials on a routine basis. It addresses the current lack, or inaccessibility, of data related to the durability of these materials, which is proving to be one of the major challenges to the widespread acceptance and implementation of FRP composites in civil infrastructure. The book should help further the acceptance of composites for civil structural applications by providing a source for practising engineers, decision makers, and students involved in architectural engineering, construction and materials, disaster reduction, environmental engineering, maritime structural technology, transportation engineering and urban planning.

Geosynthetics in civil engineering
(ISBN 978-1-85573-607-8)
Geosynthetics are essential to civil engineering and have a multitude of applications. The first part of the book looks at design principles for geosynthetics, their material properties and durability, and the range of national and international standards governing their use. Part II reviews the range of applications for synthetics as well as quality assurance issues. There are chapters on geosynthetic applications as filters, separators and barrier materials, in improving building foundations and landfill sites, and as limited design life materials.

Details of these books and a complete list of Woodhead's titles can be obtained by:

- visiting our web site at www.woodheadpublishing.com
- contacting Customer Services (e-mail: sales@woodhead-publishing.com;
- fax: +44 (0) 1223 893694; tel.: +44 (0) 1223 891358 ext. 130; address: Woodhead Publishing Ltd, Abington Hall, Abington, Cambridge CB21 6AH, England)

Multiscale materials modelling

Fundamentals and applications

Edited by
Z. Xiao Guo

Woodhead Publishing and Maney Publishing
on behalf of
The Institute of Materials, Minerals & Mining

CRC Press
Boca Raton Boston New York Washington, DC

WOODHEAD PUBLISHING LIMITED
Cambridge England

Woodhead Publishing Limited and Maney Publishing Limited on behalf of
The Institute of Materials, Minerals & Mining

Published by Woodhead Publishing Limited, Abington Hall, Abington,
Cambridge CB21 6AH, England
www.woodheadpublishing.com

Published in North America by CRC Press LLC, 6000 Broken Sound Parkway, NW,
Suite 300, Boca Raton, FL 33487, USA

First published 2007, Woodhead Publishing Limited and CRC Press LLC

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library.

Library of Congress Cataloging in Publication Data
A catalog record for this book is available from the Library of Congress.

Woodhead Publishing ISBN 978-1-84569-071-7 (book)
Woodhead Publishing ISBN 978-1-84569-337-4 (e-book)
CRC Press ISBN 978-0-8493-9110-1
CRC Press order number WP9110

The publishers' policy is to use permanent paper from mills that operate a sustainable forestry policy, and which has been manufactured from pulp which is processed using acid-free and elementary chlorine-free practices. Furthermore, the publishers ensure that the text paper and cover board used have met acceptable environmental accreditation standards.

Typeset by Replika Press Pvt Ltd, India
Printed by TJ International Limited, Padstow, Cornwall, England

Contents

Contributor contact details *ix*

Preface *xi*

1 The role of *ab initio* electronic structure calculations in multiscale modelling of materials 1

M. OB, Masaryk University, Czech Republic and Academy of Sciences of the Czech Republic

1.1 Introduction 1
1.2 Basic equations of electronic structure calculations 2
1.3 Illustrative examples 7
1.4 Conclusions 19
1.5 Acknowledgments 20
1.6 References 21

2 Modelling of dislocation behaviour at the continuum level 25

K. W. SCHWARZ, IBM Watson Research Center, USA

2.1 Introduction 25
2.2 Brief history 27
2.3 Implementation 28
2.4 Some current applications 44
2.5 Extensions of current discrete dislocation dynamics trends 54
2.6 Acknowledgments 57
2.7 References 57

3 Phase-field modelling of material microstructure 62

L.-Q. CHEN, Penn State University, USA

3.1 Introduction 62
3.2 Model description 63

3.3 Advantages and disadvantages 76
3.4 Recent developments and future opportunities 77
3.5 Acknowledgments 79
3.6 References 79

4 Mesoscale modelling of grain growth and microstructure in polycrystalline materials 84

D. Moldovan, Louisiana State University, USA and D. Wolf, Argonne National Laboratory, USA

4.1 Introduction 84
4.2 Molecular dynamics simulation of grain growth 87
4.3 Mesoscale simulation methodology 97
4.4 Validation of mesoscale simulations 103
4.5 Mesoscale simulation results 106
4.6 Summary and conclusions 114
4.7 Acknowledgments 117
4.8 References 117

5 Finite element and homogenization modelling of materials 121

J. Llorca, C. González and J. Segurado, Polytechnic University of Madrid, Spain

5.1 Introduction 121
5.2 Representative volume element 122
5.3 Homogenization techniques 125
5.4 Computational micromechanics 131
5.5 Multiscale coupling 136
5.6 Future directions 143
5.7 Acknowledgments 144
5.8 References 144

6 Grain–continuum modelling of material behaviour 148

M. O. Bloomfield and T. S. Cale, Rensselaer Polytechnic Institute, USA

6.1 Introduction 148
6.2 Representations and models 154
6.3 Grain-continuum approach 161
6.4 Grain-continuum examples 165
6.5 Opportunities 174
6.6 References 180

7	Coupled atomistic/continuum modelling of plasticity in materials	189
	R. E. Miller, Carleton University, Canada	
7.1	Introduction	189
7.2	Automatic adaption: the QC method	193
7.3	Kinematically identifying dislocations – the CADD method	202
7.4	Challenges and future trends	213
7.5	References	215
8	Multiscale modelling of carbon nanostructures	220
	T. Y. Ng, S. H. Yeak, and Y. X. Ren, Nanyang Technological University, Singapore; and K. M. Liew, City University of Hong Kong, Hong Kong	
8.1	Introduction to carbon nanotube dynamics	220
8.2	Overlap TB/MD multiscale model	221
8.3	Simulation results of carbon nanotubes under axial loading	227
8.4	Introduction to hydrogen interaction with carbon nanostructures	239
8.5	Hybrid calculations with multiscale ONIOM scheme	241
8.6	Chemosorption of hydrogen atoms onto carbon nanotubes	249
8.7	References	259
9	Multiscale modelling of structural materials	261
	D. Porter, QinetiQ, UK	
9.1	Introduction	261
9.2	Structural materials	262
9.3	Metals	264
9.4	Polymers	273
9.5	Ceramics	281
9.6	Time scales	282
9.7	Future trends	284
9.8	References	285
	Index	*288*

Contributor contact details

Editor

Z. Xiao Guo
Department of Materials
Queen Mary University of London
Mile End Road
London
E1 4NS
UK

email: x.guo@qmul.ac.uk

Chapter 1

M. ob
Department of Chemistry
Faculty of Science
Masaryk University
Kotlářská 2
CZ-611 37 Brno
Czech Republic

and

Institute of Physics of Materials
Academy of Sciences
of the Czech Republic
i kova 22
CZ - 616 62 Brno
Czech Republic

email: mojmir@ipm.cz

Chapter 2

K.W. Schwarz
IBM Watson Research Center
P.O. Box 218
Yorktown Heights
NY 10598
USA

email: kws@us.ibm.com

Chapter 3

Long-Qing Chen
Department of Materials Science
and Engineering
Penn State University
University Park
PA 16802
USA

email: lqc3@psu.edu

Chapter 4

D. Moldovan
Department of Mechanical Engineering
Louisiana State University
2504 CEBA Bldg
Baton Rouge
LA 70803
USA

email: moldovan@me.lsu.edu

Chapter 5

J. LLorca, C. González and J. Segurado
Department of Materials Science
Polytechnic University of Madrid
E. T. S. de Ingenieros de Caminos
28040 – Madrid
Spain

email: jllorca@mater.upm.es

Chapter 6

M. Bloomfield
Rensselaer Polytechnic Institute
Troy
NY 12180-3590
USA

email: bloomm2@rpi.edu

Chapter 7

R. Miller
Department of Mechanical and Aeronautical Engineering
Carleton University
1125 Colonel By Drive
Ottawa, ON
K1S 5B6
Canada

email: rmiller@mae.carleton.ca

Chapter 8

Ng Teng Yong
School of Mechanical & Aerospace Engineering
Nanyang Technological University
50 Nanyang Avenue
Singapore 639798

email: mtyng@ntu.edu.sg

Chapter 9

D. Porter
QinetiQ
Cody Technology Park
Ively Road
Farnborough
Hampshire
GU14 0LX
UK

email: dporter@qinetiq.com

Preface

Materials are fundamental building blocks of products and devices. Humans initially used materials that nature provided in finished or semi-finished forms such as flints, wood and natural fibres. This eventually progressed to engineering materials with sophisticated and heterogeneous structures due to both composition and processing. Continued harnessing of engineered materials and products in a timely and cost-effective manner requires the rapid development of new/improved processing techniques as well as in-depth understanding and accurate control of materials chemistry, processing, structure, property, performance, durability and, more importantly, their relationships. This scenario usually involves multiple length and timescales and multiple processing and performance stages, which are, sometimes, accessible only via multiscale modelling. In the past, materials modelling has contributed greatly to our understanding of materials science and to advances in various technologies. However, many of the activities are usually confined within rather separate disciplines or communities, e.g. applied mathematics, physics, chemistry, materials, engineering or medicine, each of which usually concentrates on isolated problems involving rather narrow scales or aspects of materials. It is therefore essential to bring together modelling expertise across all the length/timescales to develop multiscale-linking methodologies to fulfil future industrial demands. Recent years have seen rapid development of computational technologies, both in terms of hardware and software. It is now possible to attempt seriously multiscale modelling and simulation using even desktop computers, and to predict accurately complex materials behaviours via computational methods. Under the theme of 'Multiscale Modelling', a wealth of new results has appeared that are either specific to one given scale or establish connections between different scales. A psychological barrier has been broken, taking advantage of the progress in available computing power, which now allows some overlap between different simulation methods: the realisation that thermodynamics and purely continuum frames cannot solve all problems; and the present drive in nanostructures/nanotechnology and in energy/environmental issues.

This book aims to provide a guiding tool for both academic researchers, who are developing or wish to apply appropriate modelling methodologies for a specific phenomenon in materials science, and industrialists, who would like to gain a comprehensive knowledge of multiscale materials modelling for product and/or process design and optimisation. The chapters are contributed by internationally recognised experts in the field, and cover the spectrum of scales in modelling methodologies. I wish to take the opportunity to thank all the contributing scientists and the staff at Woodhead Publishing Limited for untiring assistance in bringing the book to publication.

Professor Z. Xiao Guo
Department of Materials
Queen Mary University of London

1

The role of *ab initio* electronic structure calculations in multiscale modelling of materials

M. O B, Masaryk University, Czech Republic
and Academy of Sciences of the Czech Republic

1.1 Introduction

Most, if not all, of the properties of solids can be traced to the behavior of electrons, the 'glue' that holds atoms together to form a solid. An important aim of the condensed matter theory is thus calculating the electronic structure (ES) of solids. The theory of ES is not ony helpful in understanding and interpreting experiments, but it also becomes a predictive tool of the physics and chemistry of condensed matter and materials science.

Many of the structural and dynamical properties of solids can be predicted accurately from *ab initio* (first-principles) electronic structure calculations, i.e. from the fundamental quantum theory. Here the atomic numbers of constituent atoms and some structural information are employed as the only pieces of input data. Such calculations are routinely performed within the framework of density functional theory in which the complicated many-body motion of all electrons is replaced by an equivalent but simpler problem of a single electron moving in an effective potential.

A general formulation of the quantum mechanical equations for ES including all known interactions between the electrons and atomic nuclei in solids is relatively simple, but we are still not able to solve these equations in their full generality. A great many approximations must be performed, which, in many cases, leads to a comprehensive solution. Its analysis brings us then some understanding of various phenomena and processes in condensed matter. The ES problem is computationally very demanding. This is why practical ES calculations in solids were rather rare prior to the availability of larger high-speed computers.

Since the 1980s, ES theory has exhibited a growing ability to understand and predict material properties and to use computers to design new materials. A new field of solid-state physics and materials science has emerged – computational materials science. This has achieved a considerable degree of reliability concerning predictions of physical and chemical properties and phenomena, thanks in large part to continued rapid development and availability

of computing power (speed and memory), its increasing accessibility (via networks and workstations), and to the generation of new computational methods and algorithms which this enabled. State-of-the-art ES calculations yield highly precise solutions to the one-electron Kohn–Sham equation for a solid and provide an understanding of matter at the atomic and electronic scale with an unprecedented level of detail and accuracy. In many cases, we are able not only to simulate experiment but also to design new molecules and materials and to predict their properties before actually synthesizing them. A computational simulation can also provide data on the atomic scale that are inaccessible experimentally. In contrast to semi-empirical approaches – the adjustable parameters of which are fitted to the properties of the ground state structure and, therefore, may not be transferable to non-equilibrium configurations – *ab initio* calculations are reliable far from the equilibrium as well.

In multiscale modelling of materials, the role of *ab initio* electronic structure calculations is twofold: (i) to study the situations where the electronic effects are crucial and must be treated from first principles and (ii) to provide data for generation of interatomic potentials. In this chapter, we will discuss both these aspects. Let us note that there exists a vast literature devoted to multiscale modelling of materials. Recent reviews may be found for example in[1,2] and in the *Handbook of Materials Modeling* edited by S. Yip[3], the latest developments are documented in the proceedings of various meetings on multiscale modelling of materials (the latest one took place in September 2006 in Freiburg, Germany[4]).

1.2 Basic equations of electronic structure calculations

In a solid where relativistic effects are not essential, we may describe the electron states by the non-relativistic many-electron Schrödinger equation

$$\mathscr{H}_{\mathrm{e},\{\mathbf{R}_\alpha\}}\,\Psi = E\Psi \qquad [1.1]$$

with the Hamiltonian

$$\mathscr{H}_{\mathrm{e},\{\mathbf{R}_\alpha\}} = -\sum_i \nabla_i^2 + \sum_i V_{\mathrm{e},\{\mathbf{R}_\alpha\}}(\mathbf{r}_i) + \sum_{i,j}{}' \frac{1}{|\,\mathbf{r}_i - \mathbf{r}_j\,|}, \qquad [1.2]$$

where $\{\mathbf{R}_\alpha\}$ are the instantaneous positions of the atomic nuclei, $\{\mathbf{r}_i\}$ denote positions of electrons, the $V_{\mathrm{e},\{\mathbf{R}_\alpha\}}(\mathbf{r}_i)$ is the potential experienced by the ith electron in the field of all nuclei at the positions $\{\mathbf{R}_\alpha\}$ with the atomic numbers Z_α, i.e.

$$V_{\mathrm{e},\{\mathbf{R}_\alpha\}}(\mathbf{r}_i) = -\sum_\alpha \frac{2Z_\alpha}{|\,\mathbf{r}_i - \mathbf{R}_\alpha\,|} \qquad [1.3]$$

and the last term in equation [1.2] represents the electrostatic electron–electron interaction (the prime on the summation excludes $i = j$). Let us note that here and throughout the chapter we use Rydberg atomic units with $\hbar = 1$, $2m_e = 1$ and $e^2 = 2$, where m_e and e denote the electron mass and charge, respectively.

1.2.1 Density functional theory

An important approach to the many-electron problem is the density functional theory (DFT), which was awarded by the Nobel Prize for Chemistry in 1998. In 1964, Hohenberg and Kohn[5] provided two basic theorems establishing formally the single-particle density $\rho(\mathbf{r})$ as a variable sufficient for a description of the ground state of a system of interacting electrons. According to their first theorem, the knowledge of the ground-state single-particle density $\rho(\mathbf{r})$ implicitly determines (to within a trivial constant) the external potential acting on the electron system. Since in turn the external potential fixes the many-body Hamiltonian $\mathscr{H}$, then, rather remarkably, the knowledge of $\rho(\mathbf{r})$ gives the entire Hamiltonian. Once the Hamiltonian is known from $\rho(\mathbf{r})$, all ground-state properties of the system are implicitly determined.

This is a great reduction of the many-electron problem as the single-particle density is a function of three variables only. All ground-state characteristics of the system in general and the total ground-state energy in particular may, therefore, be considered as functionals of only one function – the single-particle density $\rho(\mathbf{r})$.

According to the second theorem which has the form of a variational principle, the total energy of the N-electron system, $\mathscr{E}[\rho]$, is minimized by the ground-state electron density, if the trial $\rho(\mathbf{r})$ are restricted by the conditions

$$\rho(\mathbf{r}) \geq 0 \text{ and } N[\rho] \equiv \int \rho(\mathbf{r})\, \mathrm{d}^3\mathbf{r} = N.$$

Thus, the determination of the ground-state electron density and the total energy becomes extremely simple compared to the problem of solving the $3N$-dimensional Schrödinger equation: we just vary the density $\rho(\mathbf{r})$, a function of only three variables, regardless of the number of particles involved, until we find the minimum of $\mathscr{E}[\rho]$.

The DFT has emerged as an extremely powerful tool for analyzing a large variety of many-body systems as diverse as atoms, molecules, bulk and surfaces of solids, liquids, dense plasmas, nuclear matter and heavy ion systems. It is also a basis of all modern electronic structure calculations.

The DFT variational principle yields the Kohn–Sham equation having, formally, the same form as the one-particle Schrödinger equation

$$H_s\psi_i(\mathbf{r}) \equiv [-\nabla^2 + V_{\mathrm{eff}}(\mathbf{r})]\psi_i(\mathbf{r}) = \varepsilon_i\psi_i(\mathbf{r}), \qquad [1.4]$$

Here H_s is the effective one-electron Hamiltonian and $\psi_i(\mathbf{r})$ are one-electron

wavefunctions, ε_i are corresponding eigenenergies and V_{eff} is the effective potential for electrons which is, in general, non-local. It may be expressed as

$$V_{\text{eff}}(\mathbf{r}) = V_{\text{ext}}(\mathbf{r}) + \int \frac{2\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\,\mathrm{d}^3\mathbf{r}' + \frac{\delta E_{\text{xc}}[\rho]}{\delta\rho(\mathbf{r})}, \qquad [1.5]$$

where $V_{\text{ext}}(\mathbf{r})$ is the external potential due to atomic nuclei and external fields and $E_{\text{xc}}[\rho]$ is the so-called exchange-correlation energy functional containing the non-classical part of the electron–electron interaction and the difference between the kinetic energy of interacting and non-interacting electron systems[6]. The one-particle density is given by

$$\rho(\mathbf{r}) = \sum_{i=1}^{N} |\psi_i(\mathbf{r})|^2, \qquad [1.6]$$

where the sum is over the N lowest (occupied) one-electron energy states.

Equations [1.4]–[1.6] must be solved self-consistently, i.e. the density $\rho(\mathbf{r})$ must correspond to the correct effective potential $V_{\text{eff}}(\mathbf{r})$. The ground-state energy is then given by

$$\mathscr{E} = \sum_{i=1}^{N} \varepsilon_i - \iint \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\,\mathrm{d}^3\mathbf{r}\,\mathrm{d}^3\mathbf{r}' - \int V_{\text{xc}}(\mathbf{r})\,\rho(\mathbf{r})\mathrm{d}^3\mathbf{r} + E_{\text{xc}}[\rho] \qquad [1.7]$$

with

$$V_{\text{xc}}(\mathbf{r}) = \frac{\delta E_{\text{xc}}[\rho]}{\delta\rho(\mathbf{r})} \qquad [1.8]$$

being the so-called exchange-correlation potential.

The most important problem in the DFT consists in the fact that the exact energy functional (its exchange-correlation part E_{xc}) is not known. For realistic calculations, various approximations were used. The simplest of them is the local-density approximation (LDA) proposed already by Kohn and Sham[7]. The $E_{\text{xc}}[\rho]$ is written as

$$E_{\text{xc}}[\rho] = \int \rho(\mathbf{r})\,\varepsilon_{\text{xc}}[\rho(\mathbf{r})]\,\mathrm{d}^3\mathbf{r}, \qquad [1.9]$$

where $\varepsilon_{\text{xc}}[\rho(\mathbf{r})]$ is the exchange-correlation energy per particle in a homogeneous system of density ρ. Here an inhomogeneous system is replaced by a piece-wise homogeneous system.

The LDA has been remarkably successful in describing the ground-state properties of a large range of physical systems. It has proved to be surprisingly powerful in both a wide variety of cohesive properties and band structure calculations. This is the basis of its current acceptance and widespread utilization.

Nevertheless, in systems with strong density gradients, the LDA is not a very good approximation. For example, fundamental band gaps in semiconductors and insulators are typically underestimated by 40 %. LDA fails to reproduce the correct ground state in iron and cannot be applied in so-called heavy-fermion systems.

There have been several attempts to improve upon the LDA. In the generalized gradient approximation (GGA) (see e.g.[8] and references therein) an expression similar to [1.9] is used, but with $\varepsilon_{xc}[\rho]$ replaced by a local function of the density and the magnitude of its gradient, $\varepsilon_{xc}[\rho, | \nabla\rho |]$. The weighted-density approximation (WDA)[9] includes true non-local information through Coulomb integrals of the density with model exchange-correlation holes. The self-interaction corrected (SIC) methods[10,11] try to remedy LDA by explicitly removing the self-interaction terms in the electrostatic and exchange-correlation energies. The SIC–LDA approach better describes inner core electrons and is successful in treating materials with *f*-electrons[12,13] and transition metal oxides[14]. The so-called GW approximation[15] is formulated by means of a perturbation expansion of the one-particle Green function and approximates the electron self-energy by the product of an electron propagator (G) and a screened Coulomb interaction (W). The conceptual simplicity of this method is an advantage, but it is difficult to remove uncontrolled approximations. Finally, let us mention a very effective scheme called LDA + U[16]. It includes the on-site Coulomb interaction (U) for the description of correlation effects in localized *d*- and *f*-bands and seems to be appropriate e.g. for Mott insulators[16,17]. Regrettably, it seems that none of the proposed methods for going beyond the LDA leads to systematically improved results in a parameter-free manner.

1.2.2 Methods of electronic structure calculations for perfect solids

Solving the one-electron problem [equations [1.4]–[1.6]] is greatly simplified if we can suppose that the arrangement of the nuclei is periodic, i.e. that the effective potential obeys the periodicity condition $V_{eff}(\mathbf{r} + \mathbf{T}) = V_{eff}(\mathbf{r})$, where $\mathbf{T}$ is an arbitrary translation vector of the crystal lattice. A direct consequence of this condition is the famous Bloch theorem stating that each solution of equation [1.4] may be labelled by a vector $\mathbf{k}$ such that $\psi_{\mathbf{k}}(\mathbf{r}) = \exp(i\mathbf{k} \cdot \mathbf{r})u_{\mathbf{k}}(\mathbf{r})$, where $u_{\mathbf{k}}$ is a function having the same period as the crystal lattice, i.e. $u_{\mathbf{k}}(\mathbf{r} + \mathbf{T}) = u_{\mathbf{k}}(\mathbf{r})$. Therefore, it is sufficient to find the wave function $\psi_{\mathbf{k}}(\mathbf{r})$ in the so-called primitive cell, which is the smallest region generating the entire crystal by repetition in the space. Due to the periodicity condition, the region of $\mathbf{k}$-vectors may be limited to a primitive cell of the reciprocal lattice, known as the first Brillouin zone (BZ)[18].

To solve the Kohn–Sham equation [1.4], we expand the one-electron wave function as

$$\psi_{n\mathbf{k}}(\mathbf{r}) = \sum_i c_{i,n\mathbf{k}} \chi_{i\mathbf{k}}(\mathbf{r}), \qquad [1.10]$$

where the basis functions $\chi_{i\mathbf{k}}(\mathbf{r})$ satisfy the Bloch condition and form a complete set, and n is a counting index (band index). For the coefficients $c_{i,n\mathbf{k}}$ we obtain (from now on, we drop the subscript s in the one-electron Hamiltonian H_s)

$$\sum_j [\langle \chi_{i\mathbf{k}} \mid H \mid \chi_{j\mathbf{k}} \rangle - \varepsilon_{n\mathbf{k}} \langle \chi_{i\mathbf{k}} \mid \chi_{j\mathbf{k}} \rangle] \, c_{j,n\mathbf{k}} = 0, \qquad [1.11]$$

where

$$\langle \chi_{i\mathbf{k}} \mid \chi_{j\mathbf{k}} \rangle = \int_\Omega \chi^*_{i\mathbf{k}}(\mathbf{r}) \chi_{j\mathbf{k}}(\mathbf{r}) \, \mathrm{d}^2\mathbf{r},$$

$$\langle \chi_{i\mathbf{k}} \mid H \mid \chi_{j\mathbf{k}} \rangle = \int_\Omega \chi^*_{i\mathbf{k}}(\mathbf{r}) \, H \, \chi_{j\mathbf{k}}(\mathbf{r}) \, \mathrm{d}^3\mathbf{r}, \qquad [1.12]$$

Ω being the volume of the primitive cell. The energies $\varepsilon_{n\mathbf{k}}$ are determined by the well-known secular equation

$$\det [\langle \chi_{i\mathbf{k}} \mid H \mid \chi_{j\mathbf{k}} \rangle - \varepsilon_{n\mathbf{k}} \langle \chi_{i\mathbf{k}} \mid \chi_{j\mathbf{k}} \rangle] = 0. \qquad [1.13]$$

The Bloch theorem enables us to calculate the electronic wave functions and electron energies by effectively block-diagonalizing the matrix of the Hamiltonian, with each block (corresponding to a particular $\mathbf{k}$) having a manageable size. The size of each block is the number of selected basis orbitals per atom, multiplied by the number of atoms in the unit cell. The blocks are smallest when there is only one atom per unit cell; in this case, some methods yield a block size as small as 9×9, corresponding to one s orbital, three p orbitals and five d orbitals.

Various methods used in ES calculations may be distinguished according to the choice of the basis functions $\{\chi_i\}$. The better we choose them (according to the character of the problem), the smaller is the number needed for a description of the one-electron wave functions $\psi_{n\mathbf{k}}$. One uses augmented (APW) or orthogonalized (OPW) plane waves, linear muffin-tin orbitals (LMTO), linear combination of atomic (LCAO), Gaussian (LCGO) and augmented Slater-type (LASTO) orbitals, augmented spherical waves (ASW), etc. The Korringa–Kohn–Rostoker (KKR) method proceeds by the use of the Green function of the Kohn–Sham equation [1.4] and is also called the Green function (GF) method. The pseudopotential approach, applied mostly to solids containing no d- or f-elements, is also widely used. A detailed description of these methods may be found in many books and articles, e.g.[18–22], to name but a few.

After choosing an appropriate basis, the system of equations [1.4]–[1.6] is solved iteratively to selfconsistency, i.e. the electron density $\rho(\mathbf{r})$ must generate the effective one-electron potential $V_{\text{eff}}(\mathbf{r})$. The quality and speed of the convergence of such calculations is related not only to the choice of a suitable basis, but also to the sophistication of the iterative process, where as a plausible input atomic-like potentials are usually employed and input and output potentials or densities are appropriately mixed before starting a new iteration. Sometimes hundreds of iterations are needed, e.g. in metallic materials with a high peak in the density of states alternating above and below the Fermi energy, or in most surface problems.

1.3 Illustrative examples

As an illustration of materials science problems which may be investigated by *ab initio* electronic structure calculations, we present a study of magnetism of iron and its changes during phase transformations and a calculation of theoretical tensile strength in metals and intermetallics. Further, we show how the *ab initio* results may be employed to generate semi-empirical interatomic potentials for large-scale computer simulations of defect configurations. The last example demonstrates how computer modelling may be applied in positron annihilation studies of nanocrystalline materials.

1.3.1 Structure and magnetism of iron and iron overlayers

The behaviour of iron at tetragonal deformation is intimately connected with the properties of iron overlayers on various metallic substrates[23–25]. We performed *ab initio* calculations of ES and total energy of iron along tetragonal (Bain's) deformation paths at various volumes[26] and used them to understand and predict the structure and magnetic ordering of iron overlayers on (001) substrates.

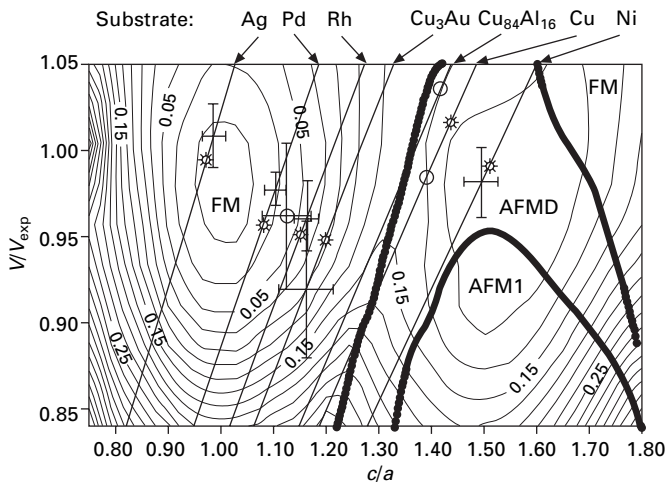
We start with the bcc structure and consider it as a tetragonal one with the c/a ratio equal to 1. Subsequently, we perform a tetragonal deformation (uniaxial deformation along the [001] axis), i.e. we change the c/a ratio and the structure does indeed become tetragonal. However, at $c/a = \sqrt{2}$, we arrive at the fcc structure, which again has cubic symmetry. The points $c/a = 1$ and $c/a = \sqrt{2}$ correspond to the only higher-symmetry structures along the tetragonal deformation path and, therefore, symmetry-dictated extrema of the total energy may be expected here[27,28].

We calculate total energy of iron along the tetragonal deformation paths keeping the atomic volume constant; the region of atomic volumes studied extends from $V/V_{\text{exp}} = 0.84$ to $V/V_{\text{exp}} = 1.05$ (V_{exp} is the experimental atomic volume). We include non-magnetic (NM), ferromagnetic (FM) and two

antiferromagnetic states, namely the single-layer antiferromagnetic one (AFM1), in which the (001) planes have alternating magnetic moments (↑↓↑↓ …), and the double-layer one (AFMD), where the pairs of (001) planes have alternating magnetic moments (↑↑↓↓ …).

For the electronic structure calculations, we used the full-potential linearized augmented plane waves (FLAPW) WIEN97 code described in detail in[29]. For the exchange-correlation energy, we employed the GGA. Further details of the calculations may be found in[26].

The total energy of iron as a function of volume and tetragonal deformation may be seen in Fig. 1.1. Here we show only those states the energies of which are the lowest for a given configuration. We can clearly see the 'horseshoes' dividing the plane into the AFM1, AFMD and FM regions. The global minimum energy is in the FM region at $c/a = 1$, $V/V_{\text{exp}} = 0.985$, which corresponds to the bcc structure. The calculated equilibrium volume is about 1.5 % lower than the experimental value, which may be considered as a very good agreement.



1.1 Total energy of iron (per atom) as a function of *c/a* and volume relative to the FM bcc equilibrium state energy calculated within the GGA. Only states with minimum energy are shown. The contour interval is 20 meV/atom. Thick lines show the FM/AFMD and AFMD/AFM1 phase boundaries. The straight lines correspond to constant lateral lattice parameters of various (001) substrates, as described in the text. The crosses composed from the vertical and horizontal error bars centered at those straight lines represent the structures of Fe films on the corresponding substrates found experimentally. One of these crosses is out of the line; it stands for the experimental structure of Fe/CuAu(001) films found in[35] and its center is denoted by an open circle. The other two open circles show the experimental structures where no error bars were given. Smaller open circles combined with asterisks represent the theoretical results found in[26].

The total energy profile of bulk iron presented in Fig. 1.1 enables us to predict easily the lattice parameters and magnetic states of iron overlayers at (001) substrates. Let us suppose that the pseudomorphic iron overlayers adopt the lattice dimensions of the substrate in the (001) plane and relax the interlayer distance (characterized by c/a). If the lattice constant of an fcc substrate is equal to a_{sub} then, in the coordinates $x = c/a$, $y = V/V_{exp}$ and $z = E$ (the total energy of iron per atom), the surfaces corresponding to a fixed a_{sub} in the (001) planes are the planes $y = kx$, where $k = (\sqrt{2}/8)(a_{sub}^3/V_{exp})$. The configuration and magnetic state of iron overlayers on a (001) substrate correspond to the energy minimum constrained to this plane, provided that the effect of the substrate/overlayer interface is not very strong. In Fig. 1.1, these planes for different values of a_{sub} are displayed by straight lines together with available experimental results and our theoretical predictions.

The experimental point for Fe films on Ag(001) is taken from[30]. It corresponds to a slightly distorted bcc structure and lies in the FM region, in agreement with experimental findings[31]. Our theoretical point is within the experimental limits. A similar situation takes place for Fe films on Pd(001)[32] and Rh(001)[33].

As with $Cu_3Au(001)$ substrate, there are two experimental findings for thin films. One of them[34] lies at the straight line for the Cu_3Au substrate in Fig. 1.1, and our theoretical point is within the experimental limits. The other one[35] is somewhat shifted from the Cu_3Au line. The structure of the films corresponds again to a tetragonally strained FM bcc phase, in agreement with experiment[35].

Fe films on $Cu_{84}Al_{16}(001)$ are reported to be fcc with the volume of 12.15 Å[36]. The corresponding point lies very close to the phase boundary between the FM and AFMD ordering in the fcc region. This is also in accordance with experimental findings – up to 4 ML, high-spin FM state is reported, for higher thicknesses a low-spin and/or AFM phase was found[36]. Here we do not have enough calculated results to find out the theoretical point, but it will probably not be too much higher than $(c/a,\ V/V_{exp}) = (1.44, 1.05)$.

Another system close to the FM/AFMD phase boundary is Fe films on Cu(001). This fact confirms a conclusion of Ref.[25] that FM and AFMD phases are energetically almost degenerate along the line corresponding to the lattice constant of Cu and somewhat favored over the AFM1 phase, and that this is closely connected with a variety of magnetic states found in the Fe films on Cu(001). Our theoretical point lies in the AFMD region and is not very far from the experimental result[37] (see Fig. 1.1).

Finally, the straight line for Fe films on Ni(001) is, for lower volumes, close to the AFMD/AFM1 phase boundary. However, the experimental point[38] is distinctly in the AFMD region, and our theoretical prediction lies again within the experimental limits. In reality, however, magnetic polarization due to FM Ni substrate may induce FM order in the film[39].

Let us note that the AFMD structure may be considered to be a close approximation of the spin-spiral state with $\mathbf{q} = (2/\pi/a)(0, 0, 0.6)$ (a is the lattice constant) found as a ground state of the fcc iron[40]. Recent calculations of Tsetseris[41] with non-collinear magnetic moments show that the inclusion of non-collinearity does not change the picture presented above too much.

1.3.2 Theoretical tensile strength in metals and intermetallics

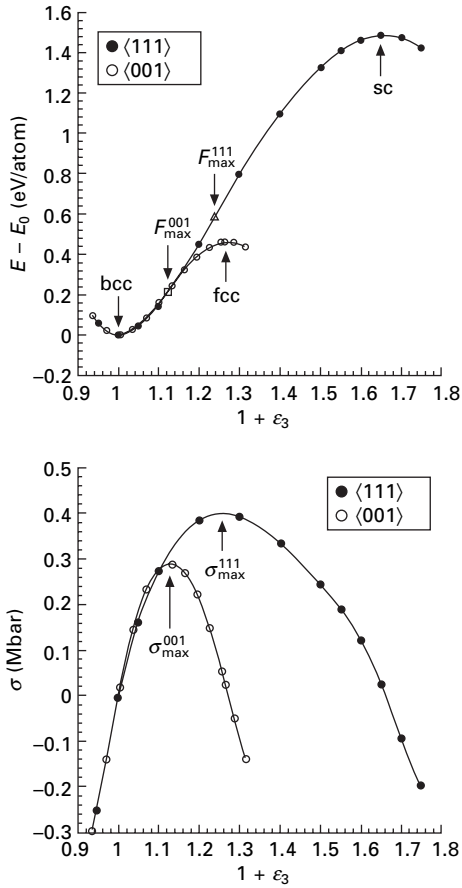
Recently, theoretical calculations of strength of materials became possible using *ab initio* ES calculations. The first paper dealing with ideal tensile strength from first principles was that of Esposito *et al.*[42], using unrelaxed structures in Cu. Paxton *et al.*[43] and Xu and Moriarty[44] calculated shear strength for unrelaxed shear deformation. Since then several further *ab initio* calculations of properties of the systems far from the ground state have been carried out, exploring their stability, etc., but the results were not employed to obtain the strength.

Probably the first *ab initio* simulation of a tensile test, including the relaxation in perpendicular direction to the loading axis, was performed by Price *et al.*[45] for uniaxial loading of TiC along the [001] axis. Later, our group at the Institute of Physics of Materials in Brno in collaboration with the group at the University of Pennsylvania, Philadelphia, initiated systematic *ab initio* studies of theoretical strength and stability in metals and intermetallic compounds under extreme conditions. In[46], we obtained the theoretical tensile strength for [001] and [111] uniaxial loading in tungsten. The calculated results compared very well with the experiment performed on tungsten whiskers by Mikhailovskii *et al.*[47]. Further, we calculated ideal tensile strength in NiAl[48] and Cu[49]. These results established a basis for further calculations of ideal strength. Li and Wang computed the ideal tensile strength in Al[50] and in β-SiC[51]. The theoretical group at the University of California at Berkeley calculated ideal shear strength in Al and Cu under fully relaxed conditions[52], performed a thorough theoretical analysis of the problem of strength and elastic stability[53] and, amongst others, verified our values of ideal tensile strength for tungsten[54].

Recently, we have simulated a tensile test in prospective materials for high-temperature applications, namely in transition metal disilicides $MoSi_2$ and WSi_2 with $C11_b$ structure, including the calculation of tensile strength for [001] loading and the analysis of bonds and their changes during the test[55,56]. A table summarizing most *ab initio* values of theoretical tensile strength for various materials is given in[57]. An extensive review of semi-empirical and *ab initio* calculated values of uniaxial and hydrostatic tensile strengths as well as of shear strengths may be found in[58]. *Ab initio* calculations of theoretical tensile and shear strength are also reviewed in the article[59],

which is, however, limited mostly to the results obtained in the author's group.

In this chapter, let us illustrate the calculation of theoretical tensile strength in the case of tungsten[46]. The strain-dependence of the total energy E of tungsten during a tensile test with loading along the crystallographic axes [001] and [111] is shown in Fig. 1.2. As tungsten is elastically nearly isotropic ($C_{44} \approx C'$), the energy curves for small deformation are nearly identical. However, a strong anisotropy for higher deformations is seen in Fig. 1.2.



1.2 Total energy, E, per atom (upper panel) and applied stress, σ, acting on the face of the body-centered cell perpendicular to the loading axis (lower panel) vs $1 + \varepsilon_3$ for uniaxial deformation of tungsten loaded along [001] and [111] crystallographic directions, respectively. E_0 is the ground-state energy of the bcc tungsten, ε_3 is the strain along the loading axis. The positions of inflexion points corresponding to the maximum forces, F^{001}_{max} and F^{111}_{max} (maximum stresses), are indicated. The extrema of the curves correspond to higher-symmetry structures (sc stands for simple cubic).

It was shown by Craievich *et al.*[27] that, for the volume-conserving tetragonal deformation (Bain's) path, the total energy exhibits symmetry-dictated extrema for structures of higher symmetry, namely for bcc and fcc in this case. The same is true for the Bain's path corresponding to uniaxial loading[23,60]. The total energy of tungsten has a minimum at the bcc structure and a maximum at the fcc structure, with an inflexion point in between which corresponds to the maximum of applied stress (see Fig. 1.2) and, if any other instability of the material does not occur before reaching the inflexion point, also to the theoretical tensile strength of 29 GPa (for a more detailed discussion of possible instabilities, see e.g.[46,61]).

The trigonal deformation curve (for loading along the [111] axis) passes from the bcc structure to simple cubic (sc) structure[28,62,63], again with an inflexion point in between. As the structural energy difference $E_{sc} - E_{bcc}$ = 1492 meV/atom is about three times higher than the $E_{fcc} - E_{bcc}$ = 463 meV/atom, the E vs $1 + \varepsilon_3$ curve for the [111] loading must rise much higher and further than that for the [001] loading (see Fig. 1.2). Consequently, the inflexion point is obtained at higher strain and with a higher value of stress (40 GPa) than for the tetragonal deformation.

Thus, a marked anisotropy of calculated ideal tensile strength in the [001] and [111] loading directions may be understood in terms of structural energy differences of nearby higher-symmetry structures occuring at the deformation path. It is in contrast with a nearly perfect elastic isotropy of tungsten at small deformations. Both values of theoretical tensile strength are comparable in magnitude with the shear moduli $C' \approx C_{44}$ (C'_{exp} = 163 GPa).

Calculated tensile strengths agree quite well with the experimental value of 24.7 ± 3.6 GPa obtained by Mikhailovskii *et al.*[47]. They measured the tensile stress in tungsten whiskers grown along the [110] direction. The experimental value is slightly lower than the calculated results, not very far from our theoretical strength of 29 GPa for [001] loading direction. We have also calculated the E vs $1 + \varepsilon_3$ curve for the [110] loading. However, this curve is even steeper than that for the [111] loading and the stress at the inflexion point is very high ($\sigma^{th}_{110} \approx 54$ GPa). Probably the material breaks down due to some other instability before reaching the inflexion point, similar to the finding for Al[64].

1.3.3 Generation of interatomic potentials for large-scale computer simulations of defect configurations

While significant progress has been made in first-principles ES calculations, there is still a wide range of problems for which we have to resort to simpler schemes of total energy evaluation. Namely, first-principles calculations rapidly become intractable if the ES and atomic configuration are to be determined in a selfconsistent way for a system having more than about 500–1000 atoms

in the repeat cell. Instead, frequently semi-empirical tight-binding models are employed. They can reproduce important features of the underlying ES and yield plausible atomic configurations[65–67]. In even more complicated cases, when thousands of atoms need to be relaxed, phenomenological models are usually used. In these treatments, the total energy is taken as a chosen function of atomic positions. Examples are pair potentials[68], many-body central-force potentials (embedded atom method[69], Finnis–Sinclair (FS) potentials[70]), glue model[71], and empirical potentials for covalent solids[72].

Both tight-binding approaches and phenomenological models contain adjustable parameters which are usually determined to reproduce some properties of the equilibrium ground state, and it is assumed that they are transferable to low-symmetry configurations such as regions of crystal defects. However, there is no guarantee that parameters fitted to equilibrium state will be applicable in non-equilibrium configurations. A significant improvement is to include into the fitting procedure some high-energy configurations calculated by first-principles methods. This may enhance the accuracy and transferability of these parameters considerably since other regions of configurational space that are not accessible experimentally are taken into account.

In a way, semi-empirical approaches are 'less fundamental' than the *ab initio* ES calculations. They should capture the physics necessary for the phenomena studies, but should not try to be all-embracing. For example, many-body central-force potentials[69,70] cannot describe directional bonds. However, they are sufficient in most close-packed structures, where angularly dependent forces are not important. In those cases, they are capable by discovering phenomena which are not purely crystallographic in nature, as e.g. the appearance of the bcc structure at some grain boundaries in copper[73]. When directional bonds are important, as e.g. in semiconductors or in bcc metals, we have to use some tight-binding model or even *ab initio* calculations. But if significant changes in the electronic structure occur with the change of atomic configuration (e.g. onset or disappearance of ferromagnetism), then only the first principles electronic structure calculations can reveal the underlying physics. Therefore, the choice of the appropriate method depends on the phenomenon studied and information sought.

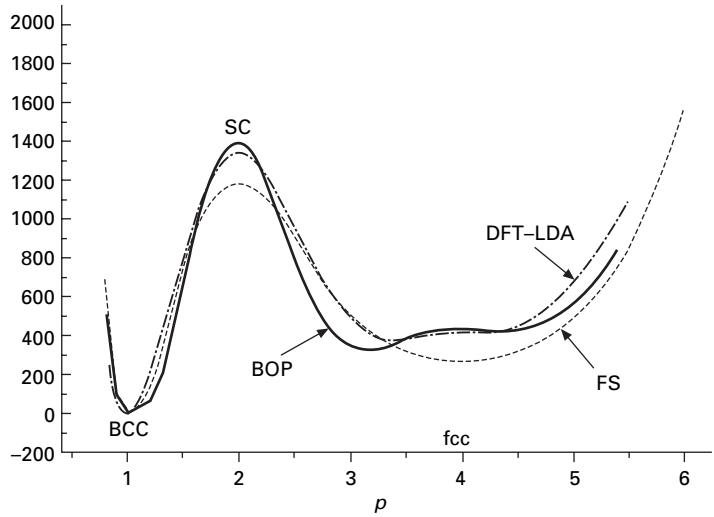
As an example of a successful modelling using the central-force potentials for a binary system, we present a combined *ab initio* and atomistic study of segregation of Bi to the grain boundaries in Cu[74,75]. Here many-body FS potentials were used and, due to lack of experimental information, the Cu–Bi potential has been fitted so as to reproduce the equilibrium density, the bulk modulus and the tetragonal shear modulus calculated for a hypothetical $L1_2$ compound Cu_3Bi from first principles. It was a crucial step in determining appropriate potentials for the Cu–Bi system. These potentials were validated by comparison between the calculated structures and the images obtained by

high-resolution electron microscopy. Now, *ab initio* results are currently used for constructing many-body central-force potentials in various systems (see e.g.[76] and the references therein).

To improve the semi-empirical description by inter-atomic interactions, an approach based on bond-order potentials (BOPs)[77] has been developed. It is, essentially, a tight-binding scheme decomposing the bond energy explicitly in terms of the contribution of the individual bonds which an atom makes with the neighbouring atoms. This is, of course, an approximation as this energy has a many-body character. The adjustable parameters are fitted using both the results of *ab initio* ES calculations and the macroscopic properties of the ground state. After fitting, again, the whole scheme must be carefully tested to ensure, for example, that the correct ground state is reproduced by the calculations (i.e. no configuration has a lower energy than the ground state), structural energy differences are given satisfactorily, etc. The best testing is performed on some configurations which are far from equilibrium. It turns out that the energy profiles along certain transformation paths are very useful for this purpose. They preserve a small unit cell and, therefore, the total energies along those paths may be calculated both *ab initio* and by using a 'less fundamental' scheme. In this way, the *ab initio* results constitute 'benchmarks' for any 'less fundamental' approach.

Figure 1.3 shows the total energy profile of molybdenum along the trigonal deformation path which connects continuously the bcc ($p = 1$), simple cubic (sc, $p = 2$) and fcc ($p = 4$) structures[28]. Here the calculations employing the BOPs[79,80] are compared with those performed *ab initio* using the FLAPW method[29] and also FS type central-force many-body potential[78]. The extrema for $p = 1$, 2 and 4 are dictated by symmetry[28]. However, additional minima found in the vicinity of $p = 4$ for Mo are specific to this material; they were also found for Nb[79] and W[28], but not for Ta[81]. These additional minima are well reproduced by BOPs. However, the central-force potentials exhibit larger deviations. They lead to a local minimum for the fcc structure in Mo while *ab initio* calculations and BOPs suggest that in Mo the fcc structure corresponds to a local maximum with subsidiary minima on each side of this maximum. Clearly, this feature is rather subtle since only small energy differences are involved, and it demonstrates that BOPs are capable of capturing such features, even though the corresponding structures are very far from the equilibrium bcc lattice.

After such testing, the developed BOPs may be used, with a high level of confidence, to simulate atomic configurations of extended defects, such as grain boundaries, dislocations, etc., where only relatively simple situations with a not very large number of independent atoms can be treated from first principles[82,83]. Recently, such a study employing a quantum-mechanically derived BOP has been performed to model the core structure and glide of the screw dislocations in molybdenum[80]. On the basis of atomistic simulations



1.3 Comparison of calculations of the energy (meV/atom) of molybdenum deformed along the trigonal deformation path using BOPs, *ab initio* electronic structure calculations (DFT-LDA) and Finnis–Sinclair type central-force potentials (FS). The atomic volume is equal to the experimental value.

using BOP for iridium, a mechanism of athermal dislocation cross slip has been proposed and the origin of brittle cleavage in iridium has been explained[84,85]. Very recently, a BOP for $MoSi_2$ has been constructed and tested[86,87].

1.3.4 Positron annihilation studies of nanocrystalline materials

Positrons are used to investigate so-called open-volume defects in solids, such as vacancies and their clusters, dislocations, grain boundaries, etc., where the electron density is 'less-than-average'. Such defects are attractive for positrons as the repulsion of the atomic nuclei is weaker there, and the positron can relax into the additional volume due to the defect with considerable energy advantage. In this way, positron samples the 'free volume' in the solid and its wavefunction is usually strongly localized in the region of the defect[88–90]. It turns out that the response of the positron is defect-specific and, therefore, different types of defects can be distinguished.

Probably the most common positron annihilation technique used in defect studies is the positron lifetime spectroscopy (PLS). In open-volume defects, the positron lifetime increases considerably (for example, in monovacancies it is about 1.6 times higher than in the defect-free bulk material). To associate

measured lifetimes with particular defects, positron lifetime calculations are often indispensable. At present, reliable models of many types of defects based on atomistic simulation are available. It is therefore advisable to investigate positron trapping at defects with realistic atomic configurations which may appear in real materials, rather than at elementary defects with simplified geometry as has been done frequently so far.

In this section we demonstrate a suitable computational technique applied to a model sample of nanocrystalline Ni (the details are published elsewhere[91]). The formula to calculate the positron lifetime τ reads[92]

$$1/\tau = \pi r_0^2 c \int n_e(\mathbf{r}) n_p(\mathbf{r}) \gamma[n_e(\mathbf{r})] d^3\mathbf{r} \qquad [1.14]$$

where r_0 is the classical electron radius, c is the velocity of light, n_e and n_p are, respectively, electron and positron densities, and γ stands for the so-called enhancement factor which describes the pileup of electrons around positrons. The enhancement factor is closely connected with the electron–positron (e–p) correlation function and increases with decreasing electron density. The positron lifetime also increases with decreasing electron density and saturates at a value of 500 ps. This fact somewhat limits the potential of PLS in the case of large vacancy clusters (voids) because the lifetime corresponding to positrons trapped at such clusters changes only minimally with their size. Nevertheless, there are many defects where this limitation does not apply.

Atomic superposition method

The electron density to be used in equation [1.14] and the positron crystal potential are usually obtained from *ab initio* electronic structure calculations. However, for systems having large number (thousands) of non-equivalent atoms, this procedure is unmanageable. A method devised to perform positron calculations in such systems is the atomic superposition technique (ATSUP)[93]. It is a non-selfconsistent technique which makes use of the atomic densities and electron Coulomb potentials of all atomic species in the system considered in order to approximate the electron density and Coulomb potential in the defect region by numerical superposition on a three-dimensional (3D) mesh, according to the atomic configuration of the system, supposing its periodicity. The positron potential is constructed as a sum of two parts: $V_p = - V_e^{Coul} + V^{corr}$, where V_e^{Coul} is the electron Coulomb potential originating from the superposition of Coulomb atomic potentials (the negative sign is due to the opposite charge of positrons) and V^{corr} is the positron correlation potential which depends on the electron density n_e. The Schrödinger equation for positrons is then solved numerically on the 3D mesh using the conjugate gradient method[93]. As an output of the calculations, we get the positron

wavefunction and energy as well as the positron lifetime and momentum distribution of annihilation photons characteristic for the defect studied. The calculated positron lifetime may be compared with experimental results. A great advantage of the ATSUP technique consists in its high computational speed; therefore, quite large atomic configurations may be treated.

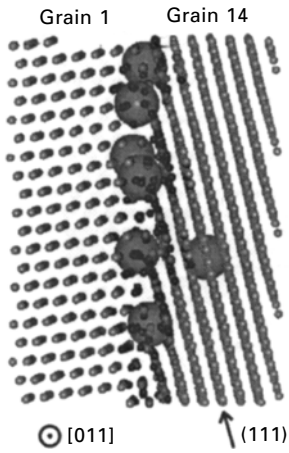
Some model (virtual) samples of materials may contain an excessive number of atoms (even hundreds of thousands or more). Such samples are too large to be considered as one cell for positron calculations, and smaller parts ('cuts') containing regions of interest have to be examined. However, such cuts (generally parallelepipeds) are not periodic and the ATSUP technique has to be modified. Explanations of the changes needed are given in[94]. Here we mention the basic points only.

First, a shallow potential well is created for positrons by adjusting the positron potential at the sides of the cut and, in this way, physically reasonable behavior of the positron wave function in this region is achieved. Second, as the cut may contain more positron traps, we extend the procedure for solving the Schrödinger equation so that it is possible to look for more positron states. We also employ a multigrid approach to speed up the calculations.

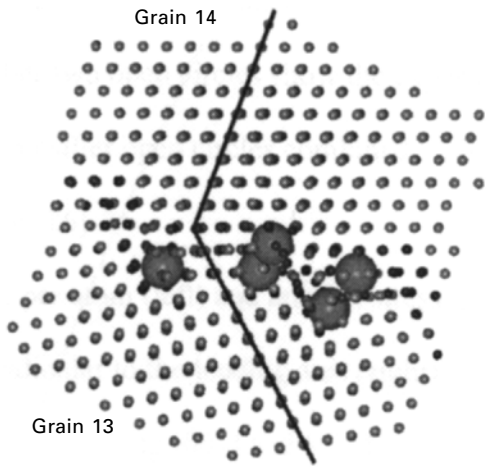
Defects in nanocrystalline Ni

In this study, we investigated the nanocrystalline sample of Ni described in detail by Van Swygenhoven *et al.*[91]. It was constructed as follows: First, the simulation cell volume was filled with nanograins (seeds) with random location and crystallographic orientation. Subsequently, the grains were allowed to grow until they touched each other. The sample was then relaxed for 50–100 ps at 300 K using molecular dynamics with second-moment (tight-binding) potentials according to Cleri and Rosato[95]. As a result, a computer sample of nanocrystalline Ni was obtained. It contained about 1.2 million atoms and 15 grains, the average grain size being 12 nm. Its average density amounted to 96 % of that of the perfect bulk Ni. Many high-angle grain boundaries were present. However, the sample did not contain any large vacancy clusters and only small free volumes (up to the size of one vacancy) associated with grain boundaries (GBs) were found.

So far, we have examined only several GBs and one triple junction (a common point of three or more GBs). Table 1.1 contains the results of our lifetime calculations for several positron lowest states at selected GBs (they are identified by the grain numbers of adjacent grains). Most calculated lifetimes are in the interval of 117–126 ps except for the GB 13–14. This GB exhibits slightly lower lifetimes (110–118 ps) than the triple junction and other GBs as it is close to a twin GB and has a slightly denser atomic arrangement and, therefore, smaller amount of free volume (Figs 1.4 and 1.5). All these values exceed the calculated bulk positron lifetime in Ni



1.4 A part of the grain boundary 1–14. The sites of positron localization are marked by large spheres. The vacancy is located a little bit aside from the GB. The orientation of some atomic planes is also indicated.



1.5 A part of the grain boundary between grains 13 and 14 together with the sites of positron localization. A pair of two twin atomic lines is indicated by solid lines.

(100 ps) only slightly and correspond to rather shallow positron traps. The sites of positron localization (maxima of the positron density) are marked by large spheres in Figs 1.4 and 1.5. In the case of GB 1–14, we have found a vacancy (with a positron lifetime of 162 ps) close to the GB (Fig. 1.4).

Up to now, we have analyzed a small fraction of GBs existing in the sample and found one vacancy only (see Table 1.1). This would correspond

Table 1.1 Calculated positron lifetimes for selected grain boundaries and triple junction (last column) in the computer sample of nanocrystalline Ni

Grain boundary	1–7	1–14	13–14	3–7–15
Lifetime (ps)	118–126	117–123, 162	110–118	119–125

to a vacancy concentration of about 1 ppm which may not be too far from reality.

The experimental lifetime spectra on nanocrystalline Ni[96] exhibit two components. Depending on the sample preparation procedure the first one ranges from 151 to 178 ps and its intensity varies from 43 to 80 %. This component is attributed to vacancy-like defects which are supposed to be associated with grain boundaries[97]. It seems, therefore, that in the measured samples positrons are trapped predominantly by defects having a size of about one vacancy.

The longer lifetime component found in measured samples varies from 300 to 400 ps. It corresponds to larger vacancy agglomerates, which do not appear in the computer sample due to the manner of its construction (see above and[91]), and we do not discuss it here in detail. Other examples of positron calculations and experiments in nanocrystalline materials may be found, e.g., in[94,98].

With regard to the comparison of theoretical calculations with experimental data, we concluded that the measured samples contain vacancies and vacancy clusters whereas the computer sample seems to be rather dominated by shallow traps (that have not been yet observed experimentally), though vacancy-like defects are also present. The real nanocrystalline samples certainly contain many shallow positron traps as there is a large volume fraction of GBs in such samples. From the point of view of PLS, the comparison of theory and experiment is complicated by the fact that the trapping efficiency for these shallow traps is not known; it seems that it is quite low (compared, e.g., to that for vacancies), which prevents these traps from being observed by PLS.

It would be desirable, of course, to construct a computer sample which would be closer to the measured samples. At the least a realistic description of larger vacancy agglomerates is needed. Furthermore, the influence of different preparation conditions on the structure of GBs would also be worthy of study. For these purposes, our sample may be considered as a good start. It may be expected that, by changing the conditions of computer simulations and improving the description of interatomic interactions, we will arrive, by trial and error, at more and more realistic computer samples.

1.4 Conclusions

The significance of *ab initio* (first-principles) electronic structure calculations lies in the high reliability of predictions of new properties and phenomena.

There are no adjustable parameters and well-defined approximations are introduced on the most fundamental level. Many macroscopic material characteristics may be determined in this way, e.g. cohesive energy, elastic constants, some strength characteristics, magnetic susceptibility, transport coefficients, etc. Specifically, in this chapter we have discussed the total energies of iron as a function of volume and tetragonal distortion and found the minimum energies and borderlines between various magnetic phases. The calculated contour plot has been used for understanding and prediction of lattice parameters and magnetic states of Fe films on various metallic substrates; our theoretical results are in a very good agreement with available experimental data. Further, we have shown that, using *ab initio* electronic structure calculations, the theoretical tensile strength of single crystals may be reliably determined and its anisotropy may be understood in terms of structural energy differences of nearby higher-symmetry structures occuring at the deformation paths.

The first-principles calculations may also be used for 'measurements in the computer'. Thus, computer simulations can substitute a real experiment and, more importantly, provide data on an atomic scale that are not accessible experimentally.

The information available from the first-principles calculations allows us both to test and construct simpler models which, in turn, may be used in very extensive atomic level studies, while the state-of-the-art first-principles calculations can only be made for a relatively small number of atoms (less than ≈ 500–1000). These simpler models employ an energy description using approximations pertinent to the level of physics investigated and concentrate on mechanisms and functional forms of physical relations. However, extensive tests of these 'less fundamental' models must always be carried out using higher levels of approximations. To achieve full understanding of the phenomena studied, it is often imperative to combine simpler methods with the first-principles calculations on the one hand and experiment on the other. This approach is being increasingly applied in multiscale modelling of materials[3,4].

1.5 Acknowledgments

I would like to thank Dr Ligen Wang, Dr Martin Friák and Dr Jan Kuriplach whose results represent important parts of this chapter. I am indebted to Dr Marc Cawkwell and Professor Václav Vitek for making some of their results for $MoSi_2$ available prior to publication. This research was supported by the Grant Agency of the Czech Republic (Projects No. 202/03/1351 and 202/06/1509), by the Grant Agency of the Academy of Sciences of the Czech Republic (Project No. IAA1041302), and by the Research Projects AV0Z20410507 and MSM0021622410.

1.6 References

1. A.E. Mattsson, P.A. Schultz, M.P. Desjarlais, T.R. Mattsson and K. Leung, *Modeling Simul. Mater. Sci. Eng*. **13**, R1 (2005).
2. J.A. Moriarty, L.X. Benedict, J.N. Glosli, R.Q. Hood, D.A. Orlikowski, M.V. Patel, P. Söderlind, F.H. Streitz, M. Tang and L.H. Yang, *J. Mater. Res*. **21**, 563 (2006).
3. *Handbook of Materials Modeling*, ed. S. Yip, Dordrecht, Berlin, Heidelberg, New York: Springer, 2005.
4. *Proceedings of the Third Int. Conf. on Multiscale Materials Modeling*, ed. P. Gumbsch, Stuttgart: Fraunhofer IRB Verlag, 2006.
5. P. Hohenberg and W. Kohn, *Phys. Rev*. **136**, B864 (1964).
6. R.G. Parr and W. Yang, *Density-Functional Theory of Atoms and Molecules*, Oxford, New York: Oxford University Press, 1989.
7. W. Kohn and L.J. Sham, *Phys. Rev*. **140**, A1133 (1965).
8. J.P. Perdew, J.A. Chevary, S.H. Vosko, K.A. Jackson, M.R. Pederson, D.J. Singh and C. Fiolhais, *Phys. Rev. B* **46**, 6671 (1995).
9. O. Gunnarsson and R.O. Jones, *Phys. Scr*. **21**, 394 (1980).
10. J.P. Perdew, *Chem. Phys. Lett*. **64**, 127 (1979).
11. A. Zunger, J.P. Perdew and G. Oliver, *Solid State Commun*. **34**, 933 (1980).
12. Z. Szotek, W.M. Temmerman and H. Winter, *Phys. Rev. Lett*. **72**, 1244 (1994).
13. A. Svane, *Phys. Rev. Lett*. **72**, 1248 (1994).
14. A. Svane and O. Gunnarsson, *Phys. Rev. Lett*. **65**, 1148 (1990).
15. L. Hedin, *Phys. Rev*. **139**, A796, (1965).
16. V.I. Anisimov, J. Zaanen and O.K. Andersen, *Phys. Rev. B* **44**, 943 (1991).
17. V.I. Anisimov, I.V. Solovyev, M.A. Korotin, M.T. Czyżyk and G.A. Sawatzky, *Phys. Rev. B* **48**, 16929 (1993).
18. J.M. Ziman, *Principles of the Theory of Solids*, London: Cambridge University Press, 1972.
19. H. Eschrig, *Optimized LCAO Method and the Electronic Structure of Extended Systems*, Berlin: Akademie-Verlag, 1988; Berlin Heidelberg: Springer, 1989.
20. T.L. Loucks, *Augmented Plane Wave Method*, New York, Amsterdam: Benjamin, 1967.
21. *Quantum Theory of Real Materials*, eds J.R. Chelikowsky and S.G. Louie, Boston, Dordrecht, London: Kluwer, 1996.
22. I. Turek, V. Drchal, J. Kudrnovský, M. ob and P. Weinberger, *Electronic Structure of Disordered Alloys, Surfaces and Interfaces*, Boston, London, Dordrecht: Kluwer, 1997.
23. P. Alippi, P.M. Marcus and M. Scheffler, *Phys. Rev. Lett*. **78**, 3892 (1997).
24. S.L. Qiu, P.M. Marcus and H. Ma, *J. Appl. Phys*. **87**, 5932 (2000).
25. D. Spi ák and J. Hafner: *Phys. Rev. B* **61**, 16129 (2000).
26. M. Friák, M. ob and V. Vitek: *Phys. Rev. B* **63**, 052405 (2001).
27. P.J. Craievich, M. Weinert, J.M. Sanchez and R.E. Watson, *Phys. Rev. Lett*. **72**, 3076 (1994).
28. M. ob, L.G. Wang and V. Vitek, *Comput. Mat. Sci*. **8**, 100 (1997).
29. P. Blaha, K. Schwarz and J. Luitz: WIEN97, Technical University of Vienna, 1997 (improved and updated Unix version of the original copyrighted WIEN-code, which was published by P. Blaha, K. Schwarz, P. Sorantin and S.B. Trickey, *Comput. Phys. Commun*. **59**, 399 (1990)).
30. H. Li, Y.S. Li, J. Quinn, D. Tian, J. Sokolov, F. Jona and P.M. Marcus, *Phys. Rev. B* **42**, 9195 (1990).

31. M. Stampanoni, A. Vaterlaus, M. Aesschliman and F. Meier, *Phys. Rev. Lett*. **59**, 2483 (1997).
32. J. Quinn, Y.S. Li, H. Li, D. Tian, F. Jona and P.M. Marcus, *Phys. Rev. B* **43**, 3959 (1991).
33. A.M. Begley, S.K. Kim, F. Jona and P.M. Marcus, *Phys. Rev. B* **48**, 1786 (1993).
34. B. Feldmann, B. Schirmer, A. Sokoll and M. Wuttig, *Phys. Rev. B* **57** 1014 (1998).
35. B. Schirmer, B. Feldmann and M. Wuttig, *Phys. Rev. B* **58**, 4984 (1998).
36. W.A. Macedo, F. Sirotti, G. Panaccione, A. Schatz, W. Keune, W.N. Rodrigues and G. Rossi, *Phys. Rev. B* **58**, 11534 (1998).
37. S.H. Lu, J. Quinn, D. Tian, F. Jona and P.M. Marcus, *Surf. Sci*. **209**, 364 (1989).
38. G.C. Gazzadi, P. Luches, A. di Bona, L. Marassi, L. Pasquali, S. Valeri and S. Nannarone, *Phys. Rev. B* **61**, 2246 (2000).
39. B. Schirmer and M. Wuttig, *Phys. Rev. B* **60**, 12945 (1999).
40. L.M. Sandratskii, *Adv. Phys*. **47**, 91 (1998).
41. L. Tsetseris, *Phys. Rev. B* **72**, 012411 (2005).
42. E. Esposito, A.E. Carlsson, D.D. Ling, H. Ehrenreich and C.D. Gelatt, Jr: *Phil. Mag*. **41**, 251 (1980).
43. A.T. Paxton, P. Gumbsch and M. Methfessel, *Phil. Mag. Lett*. **63**, 267 (1991).
44. W. Xu and J.A. Moriarty, *Phys. Rev. B* **54**, 6941 (1996).
45. D.L. Price, B.R. Cooper and J.M. Wills, *Phys. Rev. B* **46**, 11368 (1992).
46. M. ob, L.G. Wang and V. Vitek, *Mat. Sci. Eng. A* **234–236**, 1075 (1997).
47. I.M. Mikhailovskii, I. Ya. Poltinin and L.I. Fedorova, *Fizika Tverdogo Tela* **23**, 1291 (1981) (English translation in *Sov. Phys. Solid State* **23**, 757 (1981)).
48. M. ob, L.G. Wang and V. Vitek: *Phil. Mag. B* **78**, 653 (1998).
49. M. ob, L.G. Wang and V. Vitek, *Kovové materiály* (Metallic Materials) **36**, 145 (1998).
50. W. Li and T. Wang, *J. Phys.: Condens. Matter* **10**, 9889 (1998).
51. W. Li and T. Wang, *Phys. Rev. B* **59**, 3993 (1999).
52. D. Roundy, C.R. Krenn, M.L. Cohen and J.W. Morris, Jr, *Phys. Rev. Lett*. **82**, 2713 (1999).
53. J.W. Morris, Jr and C.R. Krenn, *Phil. Mag. A* **80**, 2827 (2000).
54. D. Roundy, C.R. Krenn, M.L. Cohen and J.W. Morris, Jr, *Phil. Mag. A*, **81**, 1725 (2001).
55. M. Friák, M. ob and V. Vitek, in *High-Temperature Ordered Intermetallic Alloys IX*, Mater. Res. Soc. Symp. Proc. vol. 646, Warrendale, PA: Materials Research Society, 2001, paper N4.8.
56. M. Friák, M. ob and V. Vitek, *Phys. Rev. B* **68**, 184101 (2003).
57. M. ob, M. Friák, D. Legut, J. Fiala and V. Vitek, *Mat. Sci. Eng. A* **387–389**, 148 (2004).
58. J. Pokluda, M. Černý, P. andera and M. ob: *J. Comput.-Aided Mat. Design* **11**, 1 (2004).
59. J.W. Morris, Jr, in *Handbook of Materials Modeling*, ed. S. Yip, Dordrecht, Berlin, Heidelberg, New York: Springer Verlag, 2005, p. 2777.
60. F. Milstein and B. Farber, *Phys. Rev. Lett*. **44**, 277 (1980).
61. F. Milstein and S. Chantasiriwan, *Phys. Rev. B* **58**, 6006 (1998).
62. J.A. Moriarty, *Phys. Lett. A* **131**, 41 (1988).
63. S. Fox and H.J.F. Jansen, *Phys. Rev. B* **53**, 5119 (1996).
64. D.M. Clatterbuck, C.R. Krenn, M.L. Cohen and J.W. Morris, Jr, *Phys. Rev. Lett*. **91**, 135501 (2003).

65. W.A. Harrison, *Electronic Structure and Properties of Solids*, New York: Dover Publications, 1989.
66. A.T. Paxton, *J. Phys. D: Appl. Phys*. **29**, 1689 (1996).
67. D.R. Bowler, M. Aoki, C.M. Goringe, A.P. Horsfield and D.G. Pettifor, *Modelling Simul. Mater. Sci. Eng*. **5**, 199 (1997).
68. V. Vitek, *Mat. Res. Soc. Bull*. **21**, (2) 17 (1996).
69. A.F. Voter, in *Intermetallic Compounds*, vol. 1, eds J.H. Westbrook and R.L. Fleischer, Chichester, New York: Wiley, 1995, p. 77.
70. M.W. Finnis and J.E. Sinclair, *Phil. Mag. A* **50**, 45 (1984).
71. F. Ercolessi, E. Tosatti and M. Parinello, *Phys. Rev. Lett*. **57**, 719 (1986); *Surf. Sci*. **177**, 314 (1986).
72. M.I. Baskes, *Phys. Rev. B* **46**, 2727 (1992).
73. C. Schmidt, F. Ernst, M.W. Finnis and V. Vitek, *Phys. Rev. Lett*. **75**, 2160 (1995); C. Schmidt, M.W. Finnis, F. Ernst and V. Vitek, *Phil. Mag. A* **77**, 1161 (1998).
74. D.E. Luzzi, M. Yan, M. ob and V. Vitek, *Phys. Rev. Lett*. **67**, 1894 (1991).
75. M. Yan, M. ob, D.E. Luzzi and V. Vitek, *Phys. Rev. B* **47**, 5571 (1993).
76. X.D. Dai, Y. Kong, J.H. Li and B.X. Liu, *J. Phys.: Condens. Matter* **18**, 4527 (2006).
77. D.G. Pettifor, *Bonding and Structure of Molecules and Solids*, Oxford: Oxford University Press, 1995.
78. G.J. Ackland and R. Thetford: *Phil. Mag. A* **56**, 15 (1987).
79. M. Mrověc, V. Vitek, D. Nguyen-Manh, D.G. Pettifor, L.G. Wang and M. ob, in *Multiscale Modelling of Materials*, eds V.V. Bulatov, T. Diaz de la Rubia, R. Phillips, E. Kaxiras, N. Ghoniem, Mater. Res. Soc. Symp. Proc. vol. **538**, Warrendale, PA: Materials Reseach Society, 1999, 529.
80. M. Mrověc, D. Nguyen-Manh, D.G. Pettifor and V. Vitek, *Phys. Rev. B* **69**, 094115 (2004).
81. L.G. Wang and M. ob, unpublished.
82. M. Yamaguchi, Y. Nishiyama, M. Shiga, H. Kaburaki and H. Matsuzawa: in Proceedings sub [4], 733.
83. L. Ventelon and F. Willaime: in Proceedings sub [4], p. 805.
84. M. Cawkwell, D. Nguyen-Manh, C. Woodward, D.G. Pettifor and V. Vitek, *Science* **309**, 1059 (2005).
85. M. Cawkwell, D. Nguyen-Manh, D.G. Pettifor and V. Vitek, *Phys. Rev. B*, **73**, 064104 (2006).
86. M. Cawkwell, M. Mrověc, D. Nguyen-Manh, D.G. Pettifor and V. Vitek, in *Integrative and Interdisciplinary Aspects of Intermetallics*, eds M.J. Mills, H. Inui, H. Clemens and C.L. Fu, Mater. Res. Soc. Symp. Proc. vol. **842**, Warrendale, PA: Materials Reseach Society, 2005, paper S2.8.
87. M. Cawkwell, V. Vitek, to be published.
88. *Positron Solid State Physics*, eds W. Brandt and A. Dupasquier, Amsterdam: North Holland, 1983.
89. *Positron Spectroscopy of Solids*, eds A. Dupasquier and A.P. Mills, Jr., Amsterdam: IOS Press, 1995.
90. H. Krause-Rehberg and H.S. Leipner, *Positron Annihilation in Semiconductors: Defect Studies,* Berlin, Heidelberg: Springer, 1999.
91. H. Van Swygenhoven, D. Farkas and A. Caro, *Phys. Rev. B* **62**, 831 (2000); H. Van Swygenhoven and P. Derlet, *Phys. Rev. B* **64**, 224105 (2001).
92. M.J. Puska and R.M. Nieminen, *Rev. Mod. Phys*. **66**, 841 (1994).

93. M.J. Puska and R.M. Nieminen, *J. Phys. F: Metal Phys*. **13**, 333 (1983); A.P. Seitsonen, M.J. Puska and R.M. Nieminen, *Phys. Rev. B* **51**, 14057 (1995).
94. J. Kuriplach, S. Van Petegem, D. Segers, C. Dauwe, M. Hou, E.E. Zhurkin, H. Van Swygenhoven and A.L. Morales, in *Structure and Mechanical Properties of Nanophase Materials – Theory and Computer Simulations vs. Experiment*, eds D. Farkas, H. Kung, M. Mayo, H. Van Swygenhoven and J. Weertman, Mat. Res. Soc. Symp. **634** Warrendate, PA: Materials Research Society, 2001, paper B3.8.
95. F. Cleri and V. Rosato, *Phys. Rev. B* **48**, 22 (1993).
96. S. Van Petegem, F. Dalla Torre, D. Segers and H. Van Swygenhoven, *Scripta Met*. **48**, 17 (2003).
97. H.E. Schaefer and R. Würschum, *Phys. Lett. A* **119**, 370 (1987).
98. J. Kuriplach, S. Van Petegem, M. Hou, E.E. Zhurkin, H. Van Swygenhoven, F. Dalla Torre, G. Van Tendeloo, M. Yandouzi, D. Schryvers, D. Segers, A.L. Morales, S. Ettaoussi and C. Dauwe, in *Positron Annihilation (Proc. 12th Int. Conf. on Positron Annihilation, Munich, Germany, 2000)*, eds W. Triftshäuser, G. Kögel and P. Sperr, Uetikon-Zürich: Trans Tech Publ., 2001. Appeared as *Mater. Sci. Forum* **363–365** (2001).

2

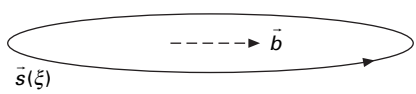
Modelling of dislocation behaviour at the continuum level

K. W. SCHWARZ, IBM Watson Research Center, USA

2.1 Introduction

A dislocation is a line defect in a crystal, marking the edge of a surface across which the atoms have been displaced by a lattice vector $\vec{b}$, and have then been allowed to rebond (Fig. 2.1). When such a line moves, it leaves behind additional surface across which relative atomic positions have changed irreversibly, as opposed to elastic distortion, where the atomic bonds are not broken, and where atomic positions recover when the forcing term is removed. Because it can occur in atomistically small increments, dislocation motion is nature's main mechanism for the irreversible (*plastic*) deformation of crystalline solids, and *discrete dislocation dynamics* (DDD) may be characterized as the attempt to describe the many aspects of plasticity in terms of a collection of dislocation lines moving and interacting in certain prescribed ways.

The theory of dislocations as filamentary topological line defects in an elastic continuum is well developed, and provides a satisfactory description of the elastic fields far away from the line. At the line itself, however, the filament idealization becomes unphysical and must be supplemented by a description of the dislocation *core*, the nature of which determines many of the detailed properties of a dislocation. Heuristic non-singular core models can be constructed in the context of the continuum approximation by distributing the source strength (i.e. the Burgers vector $\vec{b}$) over a finite core region. In



2.1 Schematic of a dislocation in a continuous medium. The material on the negative side of a surface bounded by the curve $s(\xi)$ is displaced by the *Burgers vector* $\vec{b}$, where the positive side is related to the line direction by the right-hand rule. The discontinuity in displacement across the surface creates a line source of elastic strain, the properties of which are independent of how the surface is defined.

reality, the core region of a physical dislocation extends only over a few interatomic spacings, and is thus properly treated at the atomistic level. Nevertheless, the continuum theory has value – once the detailed properties of a dislocation have been understood from atomistic theory (or measured experimentally), they can be fed as parameters into the continuum line model to provide a formalism of great predictive power. The numerical implementation of this formalism in the context of modern computers defines the field of DDD *simulation*, the aim of which is to establish honest connections between fundamental dislocation theory, theoretical models, and experiments.

The movement of a dislocation in the presence of a stress field σ_{ij} changes the elastic energy of the crystal. Thus, one can define an elastic force per unit length $\vec{f} = (\vec{b} \cdot \sigma) \times \tilde{s}'$ acting on* the dislocation, where $\vec{s}' = \mathrm{d}\vec{s}/\mathrm{d}\xi/|\mathrm{d}\vec{s}/\mathrm{d}\xi|$ is the local tangent to the dislocation. At any given time step, the typical DDD simulation will calculate the spatially varying stress σ_{ij} acting on the dislocations in the problem, determine the local elastic forces, and then move the dislocations in accord with some response model. When iterated, this process generates the evolution of the dislocation configuration. Assuming that the response model has either been obtained from atomistic theory or introduced phenomenologically, and that the applied stress field is known, the repeated calculation of the stress field generated by the dislocations themselves is the primary numerical challenge.

In principle (and given unlimited computer resources), continuum DDD simulations can be done at a high level of accuracy, taking careful account of contributions from dislocations far away, interface corrections, crystal anisotropy, and so on. This may very well be the future of the field. The situation today, however, is that many of the physical parameters entering into a DDD simulation are known only very approximately. Our discussion therefore places primary emphasis on making useful predictions within the current approximate context, and less on the elaborate formalisms needed to calculate smaller, for now largely irrelevant, effects.

The literature on dislocations is vast and impressive, and the reader is referred to the excellent standard texts[1–4] for background information. In addition to these classics, there is an excellent new book[5] and another forthcoming[6] focussing specifically on the simulation of dislocations. Our more modest aim here is to provide a snapshot of current activity in numerical DDD modeling, in particular as it refers to three-dimensionally curving dislocations. Thus, Section 2.2 reviews only briefly the relatively recent rise

*Since a dislocation is a displacement field, not a mechanical object, this force must be interpreted in the most general sense as the change in the energy of the elastic field with respect to a change in the dislocation configuration. Similarly, the dislocation velocity describes the translation of the displacement-field pattern, not the motion of the atoms in the core.

of DDD simulation as a useful tool. The field is still in flux and is developing rapidly. We therefore find it appropriate to devote considerable space in Section 2.3 to a discussion of the main technical issues that have been in play regarding the numerical implementation of the continuum theory. Section 2.4 gives an overview of a few of the current applications of DDD simulation. The concluding section comments on future research directions, in particular on the need to improve the coupling between the continuum DDD approach and the atomistic regime.

2.2 Brief history

There are two relatively easy ways of calculating the stress field produced by a dislocation filament: Brown's formula[7,8] and the Peach–Koehler formalism[9]. Brown's formula expresses stress in terms of a line integral simple enough to admit of a number of interesting analytical solutions, as detailed by Lothe[10]. It has the virtue of applying to anisotropic media in general, but can only determine the stresses in the plane of a planar dislocation configuration. The Peach–Koehler equation, on the other hand, can be used to calculate stresses everywhere from an arbitrary dislocation configuration, but applies only to isotropic media. The first significant DDD simulations were carried out by Bacon[11] and Foreman[12], who applied Brown's formula to the relaxation by glide of a single dislocation. Development along these lines was reported in several pioneering papers[13–16] and reached its highest level of sophistication in the work of Duesbery and collaborators[17–19]. At this time, Kubin, Canova, and coworkers[20–22] introduced the use of the Peach–Koehler formalism as part of an ambitious effort to address problems involving large numbers of dislocations interacting in three dimensions. Their seminal work has stimulated a remarkably rapid expansion in the number of different DDD codes that have appeared, in the technical sophistication of the algorithms, and in the variety of physical problems to which they have been applied.

The code described in Refs[20–22] simulates dislocation dynamics on a discrete lattice. Most earlier[11–19] and later workers have found it preferable to represent dislocations by a *nodal* description, that is as a series of points through which the dislocation passes (Fig. 2.2). One of the first 3D codes to take this approach was that of Zbib and coworkers[23,24], who approximate a dislocation as a piecewise continuous chain of straight segments connecting the nodal points. In addition, these authors expanded the idea, already proposed in Refs[20–22] of using 'rules' to incorporate strong local interactions such as dislocation annihilation, jog creation, and junction formation into DDD simulations[25]. The depiction of a dislocation as literally consisting of straight-line filaments chained end to end has the difficulty that the stress diverges at the corners where the segments meet, and the simple cutoff regularization adopted in[23,24] led to numerical instabilities when line segments less than

about 10 nm in length were used. A different 3D nodal code developed concurrently by Schwarz[26–29] avoided this difficulty by applying the Peach–Koehler formalism to continuously curving dislocations having a core of finite thickness. In addition to making it possible to simulate dislocation behavior down to the nanometer scale, the implementation of Refs 26–29 was the first to be fully parallelized to run on a modern supercomputer.

A reformulation of the problem in the language of the finite-element method[30–32] has since evolved into yet another simulation program[31,33,34]. Ghoniem and collaborators[35,36] have further refined this approach by using cubic splines fitted to the nodal points to define generalized coordinates. While mathematically quite formidable, their *parametric method* is claimed to have the advantage of requiring fewer points to describe smoothly curving dislocations than a straight-line interpolation scheme.

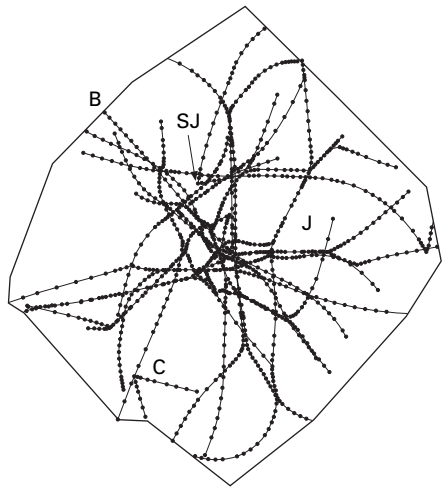
The new century has seen an acceleration of activity, leading to numerous enhancements in the capabilities of the DDD codes described above, and to a plethora of new applications. Perhaps best exemplifying the rapid pace of development is the recent effort of a group headed by Bulatov and Cai[37,38] to develop a new parallel DDD code capable of running efficiently on *tens of thousands* of processors. This ambitious project is specifically aimed at very large 'grand challenge' calculations such as the realistic prediction of strain hardening and dislocation patterning in metals. A third parallel code, also aimed at large-scale calculations, is being developed by Ghoniem and coworkers[39].

2.3 Implementation

The development of a DDD simulation program involves a number of technical and practical aspects which will be outlined in the present section. We will concentrate on issues that currently seem to be generating the most discussion, neglecting several interesting but more specialized subjects such as 2D systems of dislocations[40–42], anisotropy[43–45], and dislocation–inertia effects[46,47]. The typical object of our scrutiny (Fig. 2.2) thus consists of a collection of three-dimensionally curving dislocations, moving in a friction-limited way under the influence of applied stresses, and interacting via their stress fields with themselves, with each other, and with any boundaries that may be present.

2.3.1 Discretization and motion

Although all nodal codes specify the dislocation configuration by a sequence of points (referred to variously as mesh points, nodes, knots, or tracking points), they adopt slightly different procedures for interpolating between the points and for assigning velocities to the points. Let us describe a dislocation line by a parametric equation $\vec{s} = \vec{s}(\xi, t)$, and assume that its field translates



2.2 A sample DDD calculation of moderate size, representing a dislocation network as it might appear in a microcrystal. Strongly interacting structures and special points include junctions (J), bound crossed states (C), superjogs (SJ), and boundary termination points (B), as well as numerous more complicated states involving several dislocations interacting strongly. The dots represent a typical choice of nodal points.

with a well-defined velocity at every value of the running parameter ξ. We further suppose that this *local line velocity* is related to the elastic force per unit length acting on the dislocation by some general response model

$$\mathrm{d}\vec{s}/\mathrm{d}t = \Lambda(\vec{f}) \qquad [2.1]$$

Although $\Lambda(\vec{f})$ can in principle be a very general function of the state of the dislocation and its environment, it is usually assumed to have the form of a mobility model: that is, the line velocity is determined locally by a balance between the elastic force per unit length $\vec{f}$ and a dissipative force linear in $\mathrm{d}\vec{s}/\mathrm{d}t$, which arises from thermally activated motion over the Peierls barrier, electron/phonon drag, and other microscopic mechanisms[1,48]. If dislocation inertia effects are to be included, equation [2.1] must be replaced by an equation second order in time.

The most obvious way of defining the motion[29] is to assign to each node the velocity given by equation [2.1]. A somewhat different method is to concentrate the elastic and drag forces at the nodal points. That is, one calculates the change in the energy as well as the work done against the dissipative forces when a particular nodal point is displaced, and then moves the point with the velocity for which these balance[37]. Thus:

$$\left.\begin{aligned} \frac{\mathrm{d}\vec{s}_i}{\mathrm{d}t} &= M(\vec{f}_i) \\ \vec{f}_i &= -\frac{\partial E}{\partial \vec{s}_i} \end{aligned}\right\} \qquad [2.2]$$

While it requires additional assumptions about the behavior of the dislocation between the nodes to define $\vec{f}_i$ and the mobility function $M(\vec{f}_i)$ (the inverse of the local drag coefficient), such an implementation has the significant advantage that special points such as junction branch-points, superjog corners, and boundary termination points can be included in a natural way[37]. Yet a third popular approach has been to pick the node velocities so as to enforce the balance between the elastic and the dissipative forces over each dislocation loop as a whole. In the simple version[30,31,33], line segments are linearly interpolated between the nodes, the local forces are integrated over all of the segments, and the total elastic and irreversible contributions are equated to determine the effective nodal-point velocities via a matrix diagonalization. In the more general version developed by Ghoniem and coworkers[35,36], cubic splines are used to interpolate between the nodes, the mobility constant is replaced by a mobility tensor, and generalized coordinates and forces are defined to represent a given dislocation loop. These algorithms are typically presented in the language of the finite-element method, although they have no mechanical content beyond that discussed here.

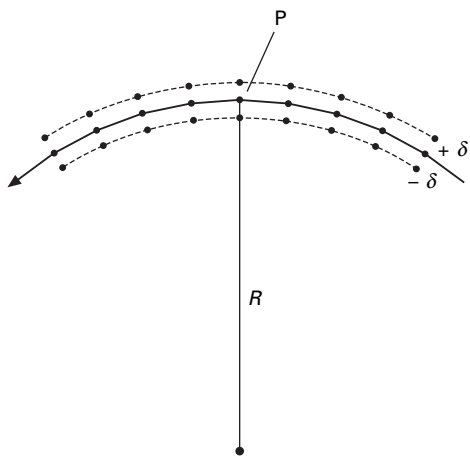
Once it has been decided how to assign an instantaneous velocity to each of the nodes, forward Euler differencing or some more sophisticated explicit–implicit time-discretization method[49,50] can be applied to evolve the line configuration. Surprisingly, elementary forward Euler differencing seems to work quite well*, producing convergent behavior and results which more or less appear to agree from program to program. Whether more refined methods perform more efficiently remains to be investigated.

In the absence of any detailed numerical analyses, a common-sense way of aiming for stability and accuracy is to require that the distance moved per time step by any nodal point be small compared to the distance between it and its neighbors. This means that the maximum time step is related to the minimum scale to be resolved (i.e. the minimum allowed node spacing $d_{\min}$) by $\Delta t \propto d_{\min}$. If this node spacing is applied throughout, the number of nodes also scales as $1/d_{\min}$, implying that the computational cost scales *at least* as rapidly as $1/d_{\min}^2$. To mitigate these costs, all modern nodal codes follow the strategy introduced in Ref.[29] of adaptively modifying the nodal-point density on the line, reducing it where the line curvature is low, and increasing it in regions of strong interaction, where the local line configuration varies rapidly.

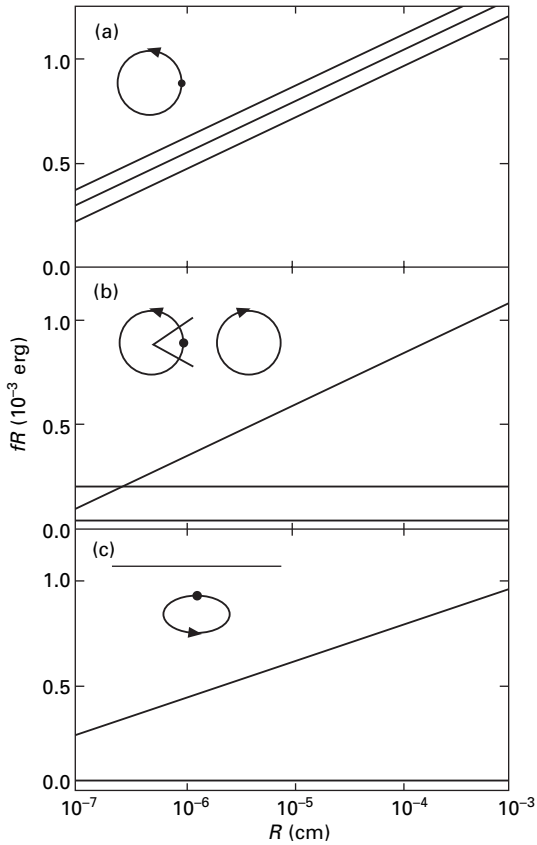
*This is unlikely to be true in situations for which inertial effects are important.

2.3.2 Regularization of the self-interaction

The field that any curved, infinitely concentrated topological line defect exerts on itself is logarithmically divergent (Fig. 2.3). This infinite self-interaction is artificial, and it must be regularized by taking into account the physical nature of the core region. Although for DDD simulations it is obviously desirable to connect this task to the microscopic scale in terms of a Peierls–Nabarro model or a more detailed atomistic calculation, current regularization schemes are all based on replacing the idealized dislocation filament by a finite distribution of the dislocation source strength (i.e. the Burgers vector). Any such regularization leads to a self-interaction proportional to $(1/R)$ $[\ln(8R/a_0) + O(1)]$[10,51], where R is the radius of curvature of the filament, a_0 is a measure of the width of the core distribution (the core 'radius'), and $O(1)$ signifies terms of order 1. The exact interpretation of a_0 and of the $O(1)$ term depends on the details of the assumed source distribution. In particular, the connection of a_0 to the microscopic physics has so far been limited to the assumption that $a_0 \simeq |\vec{b}|$. Because this could easily be off by a factor of two, *DDD simulations can at this time claim only limited absolute accuracy, no matter what specific core model is assumed*. As illustrated by Fig. 2.4a,



2.3 Regularization of the self-interaction. The line integral for the stress at P diverges as the integration along the dislocation approaches P. To avoid this, the dislocation strength is distributed over a finite core or band to produce a finite local contribution to the stress at P. The figure shows the simplest such procedure, where the field acting at P is taken to be a weighted average of the fields generated from the two split halves. Note that connecting the points by straight lines, arcs of a circle, or splines will yield slightly different results for finite $d_{\min}$. As implied in the figure, the finite core treatment must be carried out for a distance along the line which is large compared to the core radius. Other dislocations and the parts of the line farther away contribute an additional *non-local* stress at P.



2.4 Contributions to the glide force per unit length acting at the indicated point on an fcc glide loop. The loop radius R functions as a measure of the physical length scale. (a) The values of the force exerted by an isolated loop on itself are shown assuming $a_0 = b/2$ (top line), $a_0 = b$ (middle line), and $a_0 = 2b$ (bottom line). The range roughly indicates the current core error. (b) The contributions to the force per unit length from the local arc (slanted line), from the rest of the ring containing the indicated point (upper horizontal line), and from a neighboring ring (lower horizontal line). The local contribution is seen to dominate non-local contributions except at the smallest scales. (c) The contribution to the force per unit length from the dislocation loop itself (slanted line), and from the surface correction calculated according to Section 2.3.4 (lower line). The image force is a small correction unless the loop approaches the surface very closely. (a) and (b) are viewed along the (111) glide plane normal, (c) is viewed along the (110) direction horizontal to the (001) surface. Note that the linear variation of the non-local and image contributions with scale has been removed by plotting fR. The local term has an additional ln $(8R/a_0)$ variation which reduces its dominance at smaller scales.

the implied uncertainty in the self force ranges from ± 10 % at micron scales to ± 20 % at nanometer scales. We shall refer to this currently unavoidable uncertainty, which fortunately is still small enough so that DDD simulations can be quantitatively useful, as the *current core error*. It should eventually be possible to reduce the current core error in particular systems, perhaps by fitting a_0 to microscopic calculations or to experiments. Until then, however, the heroic efforts generally required to determine stress fields to better than 10 % accuracy are not worth undertaking.

The highly non-trivial issues involved in defining a particular core distribution and computing the associated total force per unit length have recently been analyzed in impressive detail by Cai *et al*.[52]. Historically, the earliest regularization method was the purely heuristic procedure proposed by Brown[53] of splitting the line in half in the glide plane, as shown in Fig. 2.3, and taking the self-stress to be the average of the two at P. A more serious analysis was provided by Gavazza and Barnett[51], who considered a distribution which essentially puts all of the source strength on the surface of a hollow tube of radius a_0 surrounding the dislocation center. They found that the Brown procedure was not entirely correct, but could be made correct by multiplying the field from the inner and outer arc by $1 - \delta/2R$ and $1 + \delta/2R$, respectively. This *modified* Brown splitting procedure is easy to implement computationally and, although its application has only been justified for smooth curves lying in a single glide plane, it is used in most calculations. A different distribution, the so-called *standard core* model, in which the source strength is distributed across a flat ribbon lying in the glide plane, has been proposed by Lothe[10]. Being much more difficult to implement numerically, however, the standard core model has not been used for DDD simulations. Very recently, Cai *et al*.[52] have obtained relatively simple explicit expressions for the case where the source strength is distributed axisymmetrically as a specific function $f(r^2/a_0^2)/a_0^3$, where a_0 is again a core size parameter, and the functional form is given numerically. This remarkable result is also well suited for computation[37], and its application to non-planar dislocations is rigorously justified.

The differences between these three regularization schemes may turn out to have some significance when matching to the atomistic core structure becomes a reality. Until that time, there is not much reason to choose one over the others, especially since DDD simulations are in any case limited by the current core error.

We conclude this section by mentioning two finite-core implementation details of proven practical importance. First, in evaluating the self-force acting at a particular point P, it is customary to apply the distributed-core model only near P, since the contributions from regions far away are adequately approximated by the simpler singular-core expressions. In making a separation between a local (regularized) contribution and a non-local contribution, it is

crucial[29] to extend the finite-core model over a distance an order of magnitude larger than the core size. Failure to do this results in numerical instability, and in particular the common[1] analytical procedure of simply cutting off the singular line integral a distance a_0 from P is computationally unacceptable. Secondly, to evaluate the self-force at a nodal point requires some information about the line configuration between the points. Most implementations simply connect the dots by straight lines. More sophisticated approaches have been to fit a circle to each node and its two neighbors[29], or to fit a cubic spline to each node and its two nearest neighbors[36]. Except for special points, the line can be assumed to be a smooth curve. For smooth regions, it is to be expected that interpolation with a circle or a spline will provide higher accuracy; equivalently, it should require significantly fewer nodal points to attain a given accuracy.

2.3.3 Approximations and 'rules'

The task of evolving dislocation configurations like that shown in Fig. 2.2 can quickly become gargantuan unless approximations are made. Fortunately, problems involving curving dislocations often present a clear hierarchy of effects that makes it possible to begin with rather elementary models, and then to move on to the more computationally challenging refinements as the problem becomes better understood, or as the current core error is reduced. The contributions to the force acting at a particular point P on a dislocation line can be loosely categorized as (i) the applied stress field driving the dislocation motion, (ii) the local self-stress generated by the line in the neighborhood of P in accord with whatever regularization procedure is adopted, (iii) the non-local field arising from the rest of the dislocation configuration, and (iv) the 'image' corrections arising from the presence of surfaces and interfaces. The simple estimates shown in Figs 2.4b and c demonstrate that the last two of these contributions are typically rather small compared to the first two*, basically because the force per unit length at P is proportional to $(1/R)$ $(\ln (8R/a_0) + O(1))$, while the contribution from an element (or an image) a distance Δ away from P is proportional to $1/\Delta$.

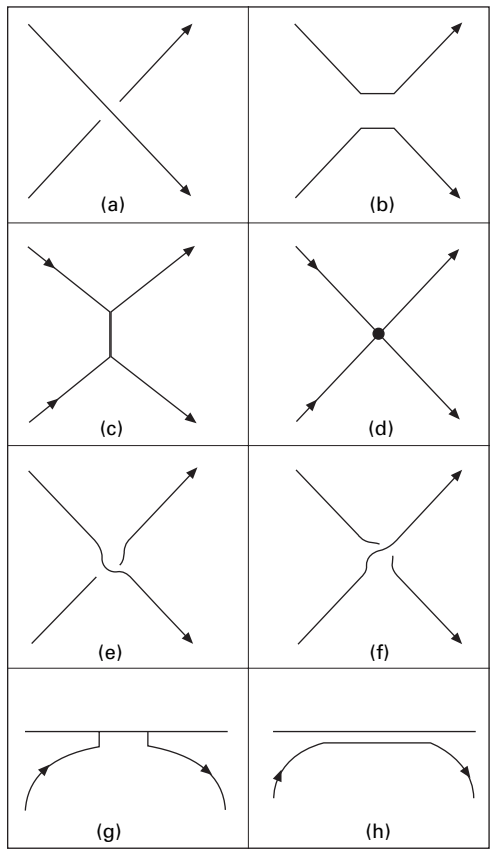
While the above implies that non-local or image contributions will be rather weak over most of a typical 3D dislocation configuration, it also implies that such contributions can never be neglected if another dislocation

*We emphasize that this presupposes that curvature plays an important role, and that the dislocations far away are randomly organized or (better yet) arranged into low-energy structures[54]. It is in fact not difficult to find exceptions, such as systems of straight parallel dislocations interacting two-dimensionally, or problems involving large-scale coherent dislocation arrays, such as the misfit array in plastically relaxed strained layers.

(or an interface) should happen to approach the dislocation to within a distance $\Delta \sim R/\ln(8R/a_0)$. In the case of two lines, such a *strong interaction* event often leads to a very complicated evolution, the interactions first affecting the local curvatures, which then in turn affect the further evolution of the two lines. Typically, two strongly interacting lines will end up either repelling each other locally, or undergoing an attractive dynamical instability leading to local annihilation or the formation of a bound state such as a junction, a bound crossed state, or a dipole. The analogous effects of an interface on an encroaching dislocation will be a local repulsion or attraction. In the typical 3D problem (Fig. 2.2), strong interactions will certainly occur as the dislocations move around three-dimensionally and try to cross other dislocations or to impinge on an interface. Although they happen only sporadically, strong interactions play a crucial qualitative role in reshaping the topology of the dislocation tangle via line–line reconnection, branched-structure formation, and line–surface reconnection. Accordingly, if one wishes to treat non-local and image forces as small corrections to the applied and the local 'line-tension' forces, one needs to come up with a way of including the topology-changing effects of the occasional strong-interaction event.

With a properly designed DDD code, it is possible to follow each strong interaction to smaller and smaller scales, with the nature of the final configuration becoming obvious well before atomistic scales are reached. Indeed, in a striking demonstration of the power of the elastic continuum approximation, Shenoy *et al.*[55] have shown that a suitably refined DDD simulation, which allows for dissociation of the core and for anisotropy, yields a Lomer–Cottrell lock structure which is in almost perfect agreement with that obtained by atomistic calculations. Obviously, this is not practical for the situation shown in Fig. 2.2, since the required scale refinements, and consequent reduction in time step, lead to a drastic slowing of the simulation each time a strong interaction needs to be resolved. The strategy adopted by DDD codes to get around this difficulty is to replace the details of the strong interactions with approximate 'rules' that mimic their effects as closely as possible.

The replacement of strong interactions by rules is illustrated in Fig. 2.5. The initial idea of applying rules[20–22] was expanded and refined by Rhee *et al.*[25] who utilized traditional qualitative estimates to develop a complete set of criteria for the processes illustrated, as well as a model for the important process of cross-slip. While their work refers to bcc crystals, it is just as applicable to the fcc case, and has formed the basis for further developments. Later, it was shown by Wickham *et al.*[56] that realistic criteria for the application of the junction-forming rule can be obtained by applying classical junction analysis to the initial configuration of the two strongly interacting dislocations. These authors also discovered that most attractive interactions do not in fact create junctions at all, but rather result in the formation of 'bound crossed



2.5 Rules commonly found in DDD simulations. The encounter (a) is turned into a line–line reconnection (b) if the Burgers vectors are the same. If the two lines are originally on intersecting glide planes, two superjogs are created by this process. If the Burgers vectors are different, a junction (c), a bound crossed state (d), or a locally repulsive interaction (e) is invoked, depending on the initial configuration. The repulsive lines are allowed to pass through each other (f) when the force becomes large enough. A line approaching an attractive (e.g. free) surface is reconnected to the surface as in (g). A line approaching a repulsive (e.g. rigid) surface is prevented from entering the surface as in (h).

states' (Fig. 2.5d). This work has been further analyzed and extended to the bcc case by Madec and collaborators[57,58].

To extend the case-by-case approach illustrated in Fig. 2.5 to more complicated strong interactions (some of which may be seen in Fig. 2.2) leads to rapidly escalating difficulties in implementation. A more recent focus, therefore, has been on replacing the rules shown in Fig. 2.5, which apply only to isolated strong interactions, by a *primitive* rule set that will

automatically approximate the proper behavior of strongly interacting dislocations in arbitrarily complicated situations such as the simultaneous interaction of several lines, multiple junctions, and so on. Schwarz[59] has formulated and tested such a rule set for his implementation in which each nodal point can have only two branches. Here, structures such as junctions and crossed points are handled by coupling nodal points together dynamically. Simple stability tests are used to create, merge, or destroy the coupled points, which are then moved using nodal dynamics. Cai *et al.*[38], whose implementation allows for nodes with more than two branches, have independently formulated a somewhat similar set of rules allowing for the creation, splitting, and merging of nodal points having more than two arms. The main difference seems to be that the coupled-point formulation treats a junction as two distinct interacting dislocations, whereas in the multibranch approach the two dislocations simply combine to make another branch with its own properties. Both primitive rule sets seem to approximate arbitrarily complicated multi-dislocation interactions in a satisfactory manner, although the approach of Ref.[38] appears to be simpler to implement.

2.3.4 Non-local and image fields

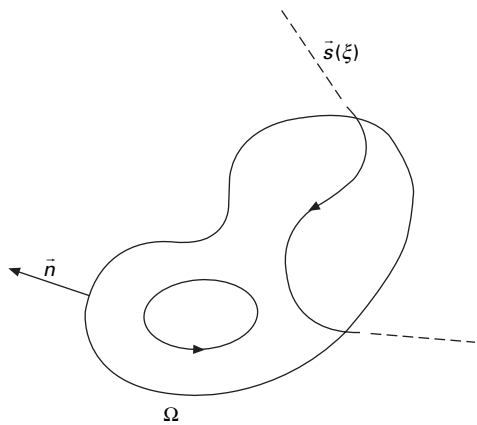
One can take the view that a satisfactory approximation is obtained if, for any point P on the dislocation, one evaluates only the applied and local fields, and perhaps the non-local fields generated by the lines in the immediate neighborhood of P. If this approximation is supplemented by the rules illustrated in Fig. 2.5, the remaining (and much more more tedious to compute) non-local and image contributions are then usually corrections of order 10 %. On the theoretical side, one can argue that this is comparable to the current core error and hence not worth worrying about; on the experimental side, few observations of dislocation behavior can be interpreted to even this degree of accuracy. Nevertheless, methods for the accurate calculation of the non-local and the image fields are of interest in their own right, and will become essential as DDD simulations attain greater absolute accuracy. Thus, considerable attention has been paid to developing these methods.

Finding the non-local stress field acting at a particular point P is a straightforward matter of summing over the fields of the line elements connecting those nodal points not involved in the evaluation of the local contribution. The main issue is efficiency, since a naive summation for all N nodal points requires N^2 field contributions to be evaluated. To facilitate the calculation, therefore, the dislocation content of regions far away from P is usually represented in terms of a multipole expansion. The initial formulation of this idea[24] has by now been superseded by very sophisticated adaptions[39,60] of the fast multipole method[61], which carry out the summation in an optimized way, turning an $O(N^2)$ problem into an $O(N)$ calculation with well-defined

error bounds. Even with such impressive machinery, however, the evaluation of the non-local contribution makes up the bulk of the computational load of large DDD simulations. To justify such an effort, it has been asserted that significant spurious effects (such as artificial pattern formation) can arise from small errors in the long-range fields or from the imposition of a long-range cutoff. More recently, it has been suggested[62,63] that patterning *arises* from the action of long-range dislocation interactions. Such ideas appear to originate in the study of 2D problems; in the 3D world exemplified by Fig. 2.2, they seem neither credible[64] nor supported by experience[57,65]. At this stage, therefore, it seems unlikely that the relatively small fields generated by distant dislocations can significantly affect 3D dislocation behavior.

Image effects are more difficult to evaluate, and in particular the proposition that they lead to only modest corrections is somewhat counter-intuitive. For one thing, a dislocation can actually intersect a boundary, in which case the image fields acting on the dislocation become infinite. For another, it seems natural to expect that image forces should play a dominant role for dislocations in nanoscale systems, for which the surface to volume ratios become very large. The issue is particularly important for the study of small confined structures. That is, it will be computationally very demanding to carry out DDD simulations in nanocrystals, *micro-electromechanical* structures (MEMS), or microelectronic devices if image effects are in fact dominant.

Although the problem of straight dislocations in the presence of a plane surface or interface has been studied extensively[66–68], the more general problem presented by curved dislocations and non-planar boundaries is a difficult one and, with few exceptions[69], only free surfaces have been considered in the context of DDD simulations. As illustrated in Fig. 2.6, the problem can then

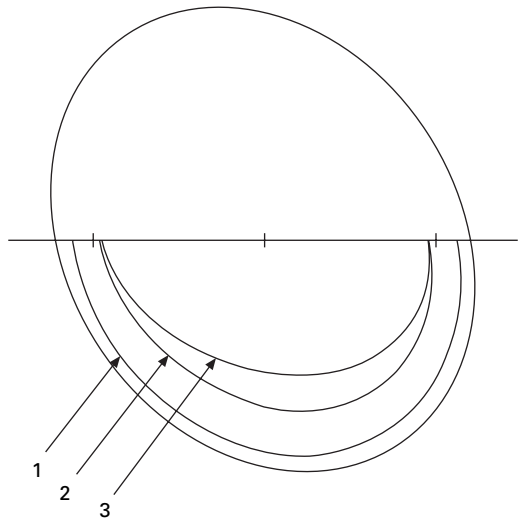


2.6 Schematic of the free-surface image problem. The stress field σ_{ij}^{∞} from continuous dislocations in the infinite medium is supplemented by a stress field $\sigma_{ij}^{\text{image}}$ which guarantees zero traction on the surface Ω.

be stated rather simply. One first finds the stress field σ_{ij}^{∞} which would be generated by the dislocations in an infinite medium. Since dislocations in an infinite medium must be continuous, this requires the artificial extension of any dislocations intersecting the surface, as shown, where the arbitrariness resulting from the fact that the continuation can have any shape is removed once the image fields have been added. In order to satisfy the free-surface boundary condition $\sigma_{ij}n_j = 0$ on the surface Ω, the traction $\sigma_{ij}^{\infty}n_j$ exerted by σ_{ij}^{∞} must now be cancelled by adding the effects of an equal and opposite traction. The problem, therefore, is to find the additional 'image field' $\sigma_{ij}^{\text{image}}$ created inside the body by a surface-traction $-\sigma_{ij}^{\infty}n_j$. Finite-element methods[33,70,71], boundary-element methods[72], and methods based on distributions of surface loops[73,74] have been implemented as a general way of solving this problem. While these methods can in principle be applied to arbitrary free surface shapes and to anisotropic media, they require diagonalization of a large matrix many times during a typical DDD simulation. Thus, they are well suited for calculating slowly varying mean fields over simple boundaries, but are rather impractical for dealing with dislocations close to or intersecting the boundary[71].

In contrast to the general free-surface problem, the special case of an isotropic medium with an infinite plane free surface can be treated relatively easily in terms of the Boussinesq–Cerruti Green's function formalism[75–78]. In this special case, it is computationally feasible to keep accurate track of image corrections, even for the singular case of a dislocation intersecting the boundary. Such studies provide insight into the importance of image corrections more generally. For example, Fig. 2.4c shows the calculated image force relative to the self-force for a loop of radius R, the center of which is a distance $2R$ from the surface. Clearly, the image correction is relatively small, on the scale of the current core error, and (as discussed in the preceding section) a strong interaction does not in fact occur unless $\Delta < R/\ln(8R/a_0)$. The behavior of a dislocation loop approaching a free surface is therefore well approximated by ignoring the surface until the loop approaches to within a distance $R/\ln(8R/a_0) \simeq O(R/10)$. Once it is closer than this, the rule pictured in Figs 2.5g and 2.5h is applied, and the dislocation experiences either a surface reconnection or a local repulsion, depending on the nature of the surface. The more extreme case of a curved dislocation actually intersecting a surface has been studied in detail by Liu and Schwarz[79]. They have found that the main effect of the image field is to maintain the angle of entry of the dislocation into the surface at a particular[80] value*. All other effects are at the current core error level of accuracy, as shown by the quantitative study illustrated in Fig. 2.7.

*Approximation schemes which try to produce this behavior in a more or less sophisticated way can be applied[29,79,81].



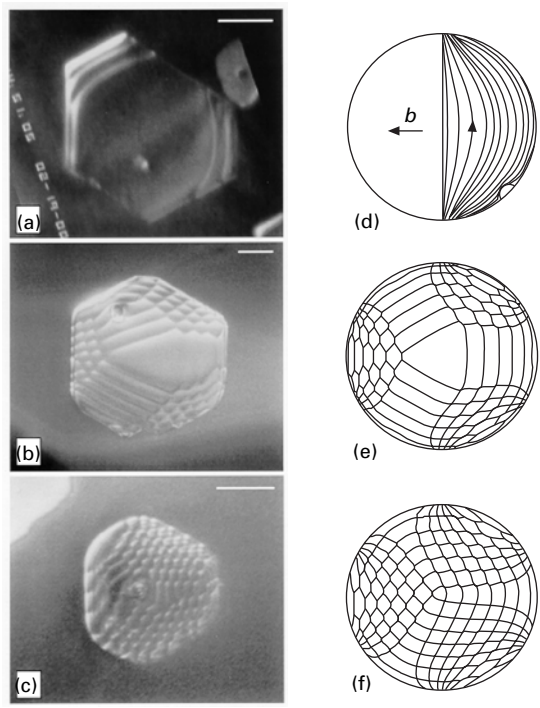
2.7 Nanometer-scale stationary loop configuration of a (111) $\frac{1}{2}$ [0$\bar{1}$1] dislocation intersecting a (001) free surface. The loop is subjected to a biaxial strain $\varepsilon_{xx} = \varepsilon_{yy} = 1.2 \times 10^{-2}$, and is viewed along the glide-plane normal. The complete loop is the stationary configuration in an infinite medium; half-loop 1 is computed taking accurate account of the image corrections via a highly refined B–C calculation[79]; half-loop 2 ignores image corrections except those acting directly at the intercept points; half-loop 3 ignores image corrections entirely. The scale markers are in units of 10 nm.

What about our intuitive notion that the image forces become very large in nanostructures? This is indeed the case. However, the line tension forces *also* increase as the scale is reduced and the curvatures become higher. As a result, the image *corrections* remain relatively small even down to the nanoscale. Similar qualitative behavior may be expected for other kinds of interfaces or other interface geometries, the implication being that image effects in general lead to rather modest corrections even in nanoscale structures and even for dislocations intersecting the boundaries*. A striking example of what one can accomplish with DDD simulations even if image corrections are ignored entirely[83] is shown in Fig. 2.8.

2.3.5 Periodic boundary conditions

In many of the larger calculations, e.g. those simulating the mechanical properties of bulk metals or the plastic relaxation of a strained layer, one is

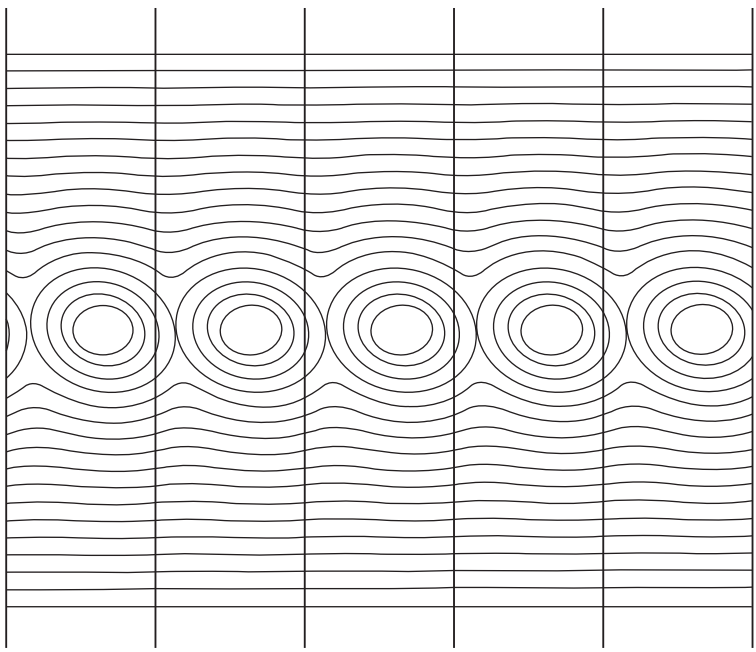
*An interesting apparent exception to this is the finding of Wang *et al.*[82] that the image force can have a strong effect on the cross-slip of a dislocation intersecting the surface.



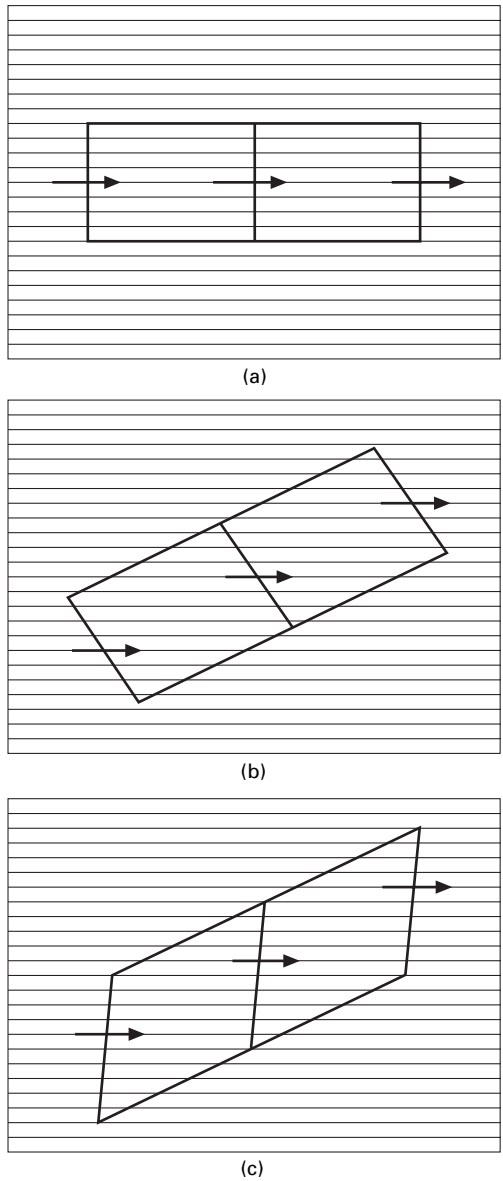
2.8 (111) plan-view of dislocations in $CoSi_2$ islands as observed by *in situ* transmission electron microscopy, and as predicted by a DDD simulation in which the island geometry is approximated as a truncated cone and image corrections are ignored completely. The pictures were taken in a **g** = <220> dark field condition for the Si substrate, showing the dislocations as lines of strong bright and dark contrast. The scale bar on each picture represents 250 nm. (a) Initial entry of the dislocations, recorded at the growth temperature of 850 °C. (b) Observed configuration after holding the specimen at 850 °C for 30 minutes. The island shows a more symmetrical and developed dislocation array, although the center remains unrelaxed. (c) A second, smaller island, this time fully dislocated, recorded after cooling the specimen. The small circular features within each island are pinholes which form during growth. (d) Series of position snapshots showing the calculated evolution of a single $b = \frac{1}{2}$ [100] dislocation loop introduced at the island edge. The motion of the dislocation slows down drastically as it grows into the center. (e) Calculated transient configuration resulting from the introduction of six sets of dislocation loops, each containing the three possible $\frac{1}{2}$ [110] Burgers vectors for the (111) interface plane. (f) Fully relaxed pattern predicted by the simulations. Additional loops introduced into the island will not grow. The location or order in which the dislocation loops are introduced does not affect the final configuration.

not interested in the effect of boundaries *per se*. Rather, the goal is to eliminate or at least clearly identify the effects arising from the finite size of the calculation. One approach is to place a real boundary such as a free surface around the computational sample, and to probe only the properties near the center. This approach presumably yields correct results when the measured properties become independent of the sample size, but is extremely inefficient. A more elegant approach which has received quite a bit of attention is the use of periodic boundary conditions. As was pointed out long ago in a different context[84], the use of periodic boundary conditions in a tangle code poses the danger illustrated in Fig. 2.9. That is, a loop growing through a rectilinear boundary can undergo a strong interaction with its own replica. In the illustration, the loop reconnects to its replicas to form two infinite lines. The dynamics of two lines is very different from that of a loop, and this kind of topological contamination can lead to incorrect results even when it is far from obvious that it is taking place.

An obvious way of minimizing self-interaction errors is to rotate the periodicity box as shown in Fig. 2.10b, so that the replica enters on a different



2.9 Sequential snapshots of a dislocation loop growing under periodic boundary conditions when it can re-enter on the same glide plane. By reconnecting to itself it goes over into two infinite lines, which behave in a manner completely different from that of a loop in an infinite system.



2.10 Periodic boundary conditions in DDD simulations. (a) The replica re-enters on the same glide plane. The figure is essentially Fig. 2.9 viewed edge on. (b) Rotation of the periodicity box allows the replica to enter on a harmless glide plane. (c) Definition of the periodicity box edges in terms of lattice vectors permits an infinity of satisfactory solutions, making it easy to avoid self-interactions for many slip systems at once.

glide plane[85]. Since the line must re-enter on an equivalent glide plane, the possible rotation angles form a discrete set, and the idea is to choose a value which avoids strong self-interactions for as many passes through the box as possible. It is not in general possible to apply this trick to more than one glide-plane normal, and a more flexible solution[86] is to define the edges of the periodicity box in terms of suitable lattice vectors (Fig. 2.10c). This offers an infinity of choices that automatically satisfy the re-entrant condition for an arbitrary crystal structure, making it quite easy to choose the periodicity box such that multiple slip systems evolve with minimal self-interactions. The added flexibility comes at a cost: the periodicity box is now defined by a non-orthogonal triad of vectors, and the algorithmic implementation is significantly more complicated.

It is worth pointing out that the use of periodic boundary conditions embeds certain periodicities by fiat, via the initial conditions. This may make it difficult to interpret computed dislocation patterns, when and if DDD simulations reach the point where such structures actually become manifest. Certainly one would expect any structure on the scale of the periodicity box to reflect the effects of the artificially applied periodicity. Thus it may turn out that a technique which reinjects dislocations in a more random manner will be required to produce convincing results.

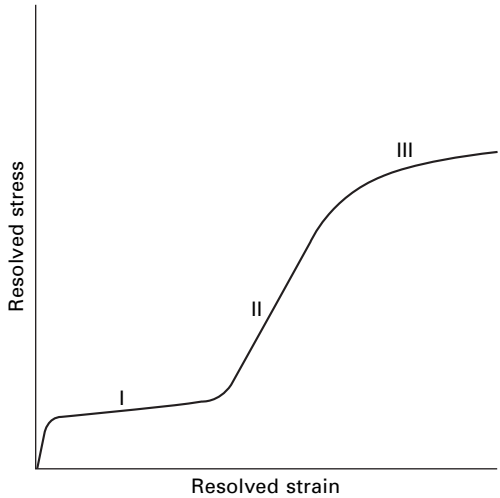
When using periodic boundary conditions, it is in principle possible to evaluate long-range fields exactly by summing the periodic image of each element over all of space. Cai *et al.*[87] have calculated such sums for orthorhombic periodicity volumes. By storing the results in a lookup table, the calculation can be made very efficient. Their method has not yet been extended to lattice-vector periodicity volumes.

2.4 Some current applications

As the usefulness of DDD simulation becomes manifest, the technique is being applied to an ever-expanding range of systems. A complete survey is beyond the range of a short review, and would in any case be obsolete by the time that this article appears in print. This section is therefore limited to a few topics reflecting the author's particular interests, and chosen to illustrate the potential range of applications for DDD simulations.

2.4.1 Bulk plasticity

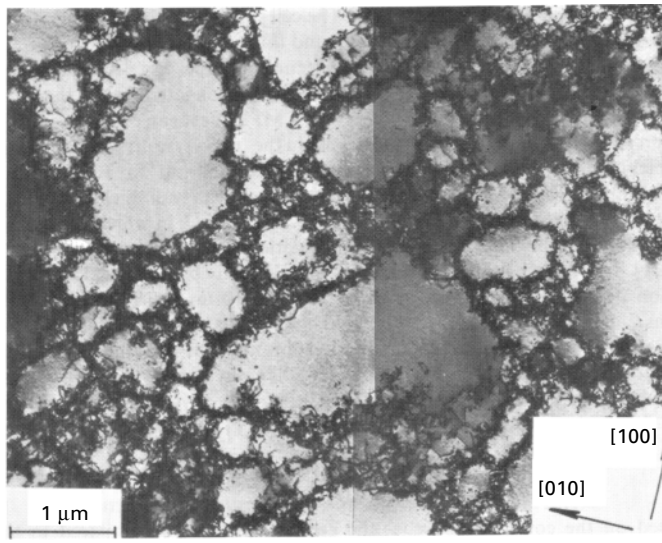
Although the mechanical properties of crystalline materials exhibit great variability, some aspects of the idealized deformation behavior shown in Fig. 2.11 are often observed[54,88,89]. After an initial elastic response, the material yields and enters an easy glide regime (I) where dislocations move primarily on a single set of glide planes (single slip), growing and multiplying in a



2.11 The idealized single-crystal strain-hardening behavior.

relatively unimpeded fashion. The transition to the hardening regime (stage II) is marked by the activation of a secondary slip system, leading to a rapid increase (storage) in the density of dislocations capable of interfering with the motion of the primary slip dislocations via the strong interactions of junction and crossed-state formation, and of superjog formation due to annihilation between dislocations on intersecting glide planes. In the dynamic recovery stage III, the annihilation of these stored dislocations due to cross-slip and climb processes takes on increasing importance, and the hardening rate decreases. These stages are accompanied by the development of an internal heterogeneous dislocation structure progressing from the formation of dipole and multipole bundles (stage I), to sheet-like structures (stage II), to two- and three-dimensional cellular patterns of dislocation walls (early stage III), to mosaic structures resembling the polycrystalline state (late stage III). The elucidation of these phenomena has long been considered the Holy Grail of metallurgy, and has been the prime motivator in the development of large-scale DDD simulation codes.

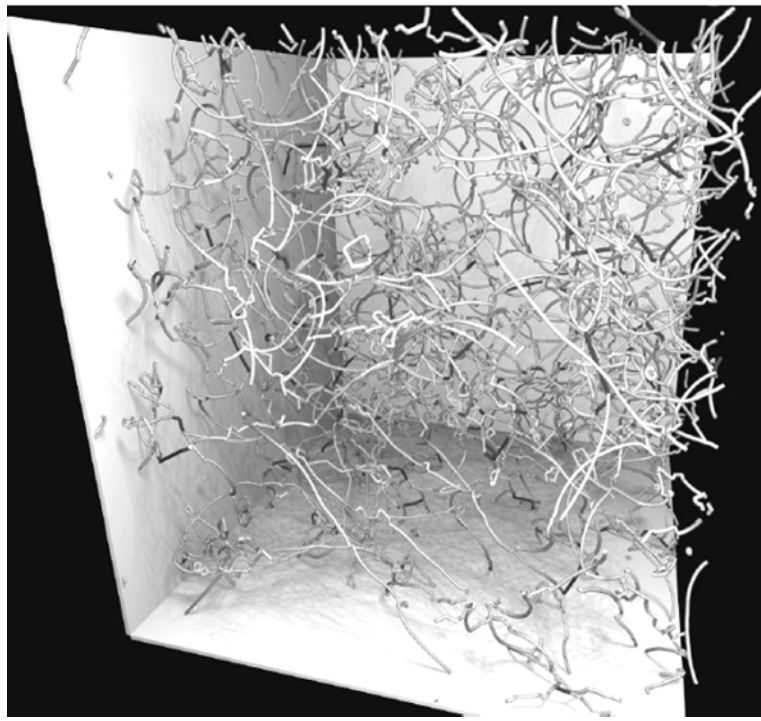
The computational cost of evolving a late-stage patterned structure such as that shown in Fig. 2.12 is formidable. The number of nodal points required to describe the structure is of order $\rho D^3/d_{\min}$, where ρ is the line-length density, D is the size of the sample box, and $d_{\min}$ is the minimum scale to be resolved. If the maximum displacement during a time step is of order $0.01d_{\min}$, it takes $100D/d_{\min}$ steps for a dislocation to cross the sample. For typical values $\rho = 10^{14}$ m^{-2}, $D = 10\ \mu$, and $d_{\min} = 0.01\ \mu$, one is then facing the task of evolving ten million nodal points for hundreds of thousands of time steps. While most of the current codes have produced impressive-looking pictures of dislocation tangles, it seems that only a massively parallel, grand challenge



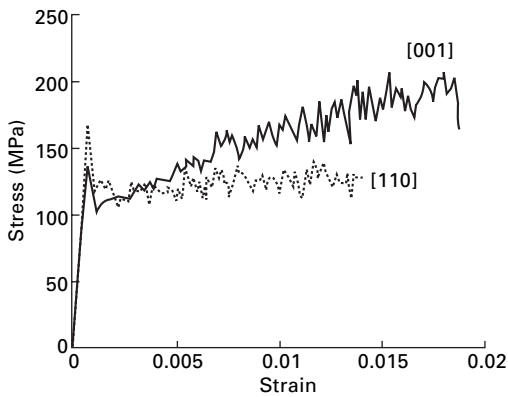
2.12 TEM micrograph[90] of the dislocation cell structure in a Cu crystal deformed in tension along the [001] direction. The resolved shear strain is 0.52.

approach of the type currently being pursued at Lawrence Livermore National Laboratories[37,38] offers any hope of obtaining a structure like that shown in Fig. 2.12 by brute-force simulation. Figures 2.13 and 2.14 show the most recent results obtained as of this writing[91].

A more indirect course is to use DDD simulations to seek insight into individual hardening processes, and to evaluate the parameters entering classical hardening models. Particularly noteworthy here are recent efforts to determine the effectivess of interactions between intersecting slip systems in impeding dislocation motion, data which are usually expressed in terms of a hardening matrix. The mechanism-oriented approach has led to several qualitatively new discoveries. As mentioned earlier, it has recently been found that dislocations with differing Burgers vectors are more likely to form bound crossed states than to form the extended junctions usually considered[56–58]. Multi-junctions in which more than two dislocations combine to form structures with multiply-branched endings also seem to appear naturally during the simulations, and their existence has secretly been confirmed experimentally[92]. Perhaps most significantly for hardening theory, it has been found that the creation of superjogs by the annihilation of intersecting dislocations[93] is at least as potent a hardening mechanism as junction formation[94,95]. Similar observations have been made in the context of layer plasticity both on the basis of experiment[96] and DDD simulations[97]. Insights into the mechanisms leading to patterning or its effect on work-hardening, however, are still lacking.



2.13 Simulated dislocation microstructure in a 10 μm cubed cell under periodic boundary conditions[91]. Single-crystal Mo under uniaxial tension at a strain rate 1 s^{-1}.

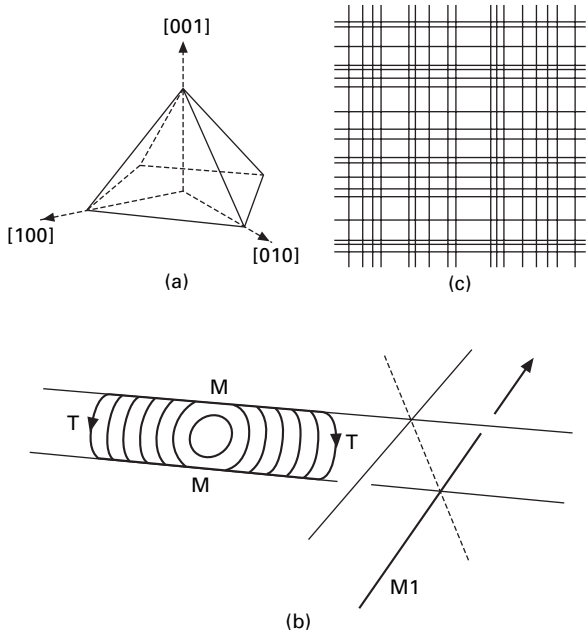


2.14 Stress–strain curves for different loading directions, showing anisotropy in strain hardening rate[91]. Single-crystal Mo under uniaxial tension at a strain rate 1 s^{-1}

2.4.2 Film plasticity

The mechanical behavior of thin films[98,99] has significant technological implications in microelectronics, electro-optics, magnetic recording, and many other areas. Because of its strong connections to technology, research in film plasticity has traditionally been carried out by two largely separate communities focused on semiconductor multilayers and on thin metal films, respectively. Strain-relaxed SiGe layers on Si substrates, for example, are used to provide a new kind of crystalline substrate enabling the development of novel microelectronic devices. Such semiconductor multilayers are easily grown in single-crystal form, and the prototypical phenomenon of interest is the one-time plastic relaxation of a single-crystal strained layer on an unstrained substrate. Copper or aluminum films, on the other hand, serve as the electrical interconnects between devices, and must maintain their integrity under repeated thermal-stress cycling. Furthermore, they typically grow only in polycrystalline form, so that the prototypical phenomenon of interest is the plastic response of a polycrystalline film to cyclic loading. Although film plasticity involves at least as many complications as bulk plasticity, it is computationally more accessible: dislocation patterning of the type observed in bulk plasticity does not seem to occur[100], and compared to bulk calculations the number of nodal points required for a simulation is less by a factor of $h/D << 1$, where h is the film thickness and D the sample size.

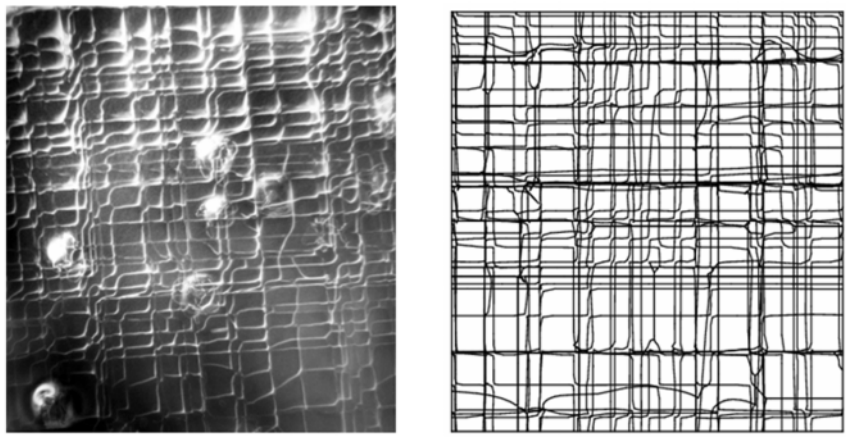
The plastic relaxation of a single-crystal strained layer is mediated by the motion of dislocations threading the layer, leaving behind an array of misfit dislocations in the interface (Fig. 2.15). In order to be useful for further processing, the strained layer must relax as completely as possible, leaving as few dislocations threading the layer as possible. While this goal has been a major focus of research and development efforts for many decades, theory for even this relatively simple problem has been limited to the phenomenological application of two basic ideas. The first is the fact that a minimum critical stress must exist in a layer before a dislocation can grow into it[101–103]. Since the critical stress is inversely proportional to the layer thickness, thin layers are harder to deform plastically. The second idea is that misfit dislocations can block the propagation of threading dislocations[104,105], and thus prevent the strained layer from relaxing to its lowest energy state. DDD simulations have proven very useful in going beyond this elementary picture and, as shown in Figs 2.16 and 2.17, have in fact been able to predict not only the observed degree of strain relaxation, but also the dislocation patterns obtained in well-controlled layer relaxation experiments[106]. Indeed, in this application DDD simulations have reached the point of serving as a practical engineering tool for substrate development. More fundamentally, such simulations have identified several additional mechanisms limiting layer plasticity. Thus, bulk hardening mechanisms such as junction formation and



2.15 Typical geometry for the relaxation of a capped fcc layer. (a) The pyramid shows the crystallographic directions appropriate to the fcc structure. The base of the pyramid is parallel to the plane of the layer, which coincides with the (001) plane of the crystal. The faces of the pyramid define the {111} set of glide planes, while the edges coincide with the <011> set of allowed Burgers vectors. (b) Same-perspective view of two intersecting glide planes, (111) and $(\bar{1}11)$, in the capped layer. A loop introduced on the (111) plane is shown at several steps of its evolution leading to two independently propagating arms T connected by misfit segments M at the interfaces. One of the threading arms is about to encounter a blocking misfit dislocation M1 on a $(\bar{1}11)$ glide plane. (c) Plan view of an idealized relaxed layer, showing the typical cross-hatched pattern of misfit dislocations that relieve the strain.

superjog creation are found to play a significant role, and the interactions between dislocations on parallel glide planes are found to be at least as important as blocking in determining the lowest strain to which a given layer can relax[97].

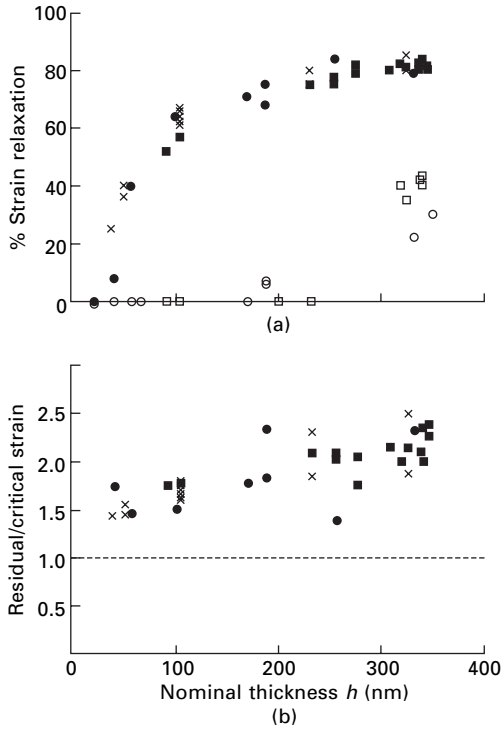
The plastic behavior of metal films is much less well understood. A typical cyclic loading curve[107,108] has the form of a hysteresis curve as shown in Fig. 2.18. Here the differential thermal expansion of film and substrate leads to a cyclic application of strain to the film as the temperature is raised and lowered. Such curves are sensitive to grain size, to grain orientation, to previous strain history, and to the nature of the film surfaces. At higher temperatures, the plastic response is further complicated by the emergence



2.16 (a) Plan view TEM image of the final dislocation configuration of a 104 nm $Si_{0.8}Ge_{0.2}$ strained layer grown epitaxially on a Si substrate. The layer is allowed to relax by first injecting dislocations via helium implantation, and then annealing. The round features are coalesced-helium structures that have formed below the SiGe/Si interface. (b) Representative sample of the dislocation configuration predicted by DDD simulation for the same thickness and applied strain. The scale is the same as that of Fig. 2.16a.

of grain-boundary diffusion as an important relaxation mechanism. The characteristic hysteretic irreversibility seen in Fig. 2.18 may reflect the dynamics of an internal dislocation network as it responds to loading and unloading, or it may reflect the conditions under which grain boundaries emit and absorb dislocations, or it may reflect stress relief by grain-boundary diffusion, but most likely it represents a complicated mixture of all three.

Even an exploratory DDD simulation of a single grain requires one to introduce models for the nucleation, annihilation, and transmission of dislocations by the grain boundaries. A promising approach to including transmission of dislocations through grain boundaries has been explored by De Koning *et al.*[109,110], while Hartmaier *et al.*[111,112] have proposed a unified DDD model in which grain-boundary diffusion is pictured in terms of the climb of discrete dislocations in the boundary. So far, however, DDD simulations have explored only the cyclic loading of a pre-existing dislocation structure in a grain with hard walls[113–115] and in a single-crystal layer with impenetrable top and bottom interfaces[116,117]. The results have been encouraging, appearing to exhibit the effects of grain size (Hall–Petch effect), and the hysteresis arising from the formation of internal dislocation configurations stabilized by junctions and superjogs (Bauschinger effect). But observations of actual dislocation behavior in grains clearly show dislocations being emitted and absorbed by the grain boundaries[118–120], and

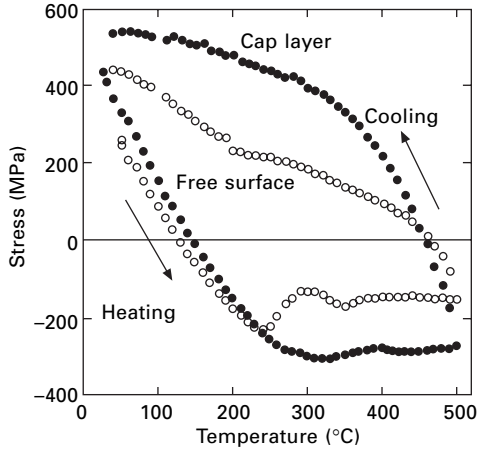


2.17 (a) The solid symbols show the strain relaxation percentage $100(1 - \varepsilon_{final})/\varepsilon_{initial}$ observed in helium-implanted and annealed $Si_{0.8}Ge_{0.2}$ layers. Circles refer to UHVCVD-grown layers, squares to RTCVD-grown layers. The open symbols show the behavior of layers that were not implanted. The x symbols show residual strains predicted by DDD simulations. (b) The data for implanted samples compared with the theoretical results when plotted in the form $\varepsilon_{final}/\varepsilon_{crit}$.

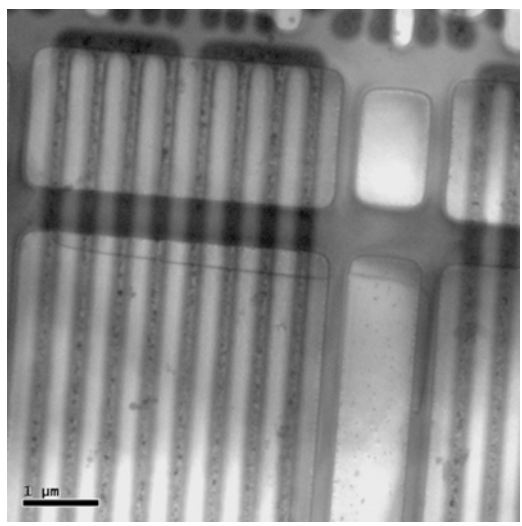
are thus not necessarily in accord with such hard-wall models. Espinosa and collaborators[121] have formulated a somewhat more promising model, in which a grain boundary can act as a source capable of emitting a limited number of dislocations. This model has had some success in predicting grain-size effects in films. The interaction of dislocations with grain boundaries is currently a fashionable topic, and the development of a model suitable for DDD simulations is an urgent priority.

2.4.3 Microelectronic devices

Although the subject has been largely ignored by the dislocation community, the nucleation and migration of dislocations has sporadically been a very serious problem for the microelectronics industry[122]. The core of a dislocation can getter impurities, causing it to act as an intermittent short in a logic or



2.18 Stress-evolution as a function of temperature for Cu films with a thickness of 0.5 μm. The film stress was measured by the wafer-curvature technique. Solid circles refer to a film with a capping layer, open circles to a film with a free surface. From Ref.[108], data from Ref.[107].



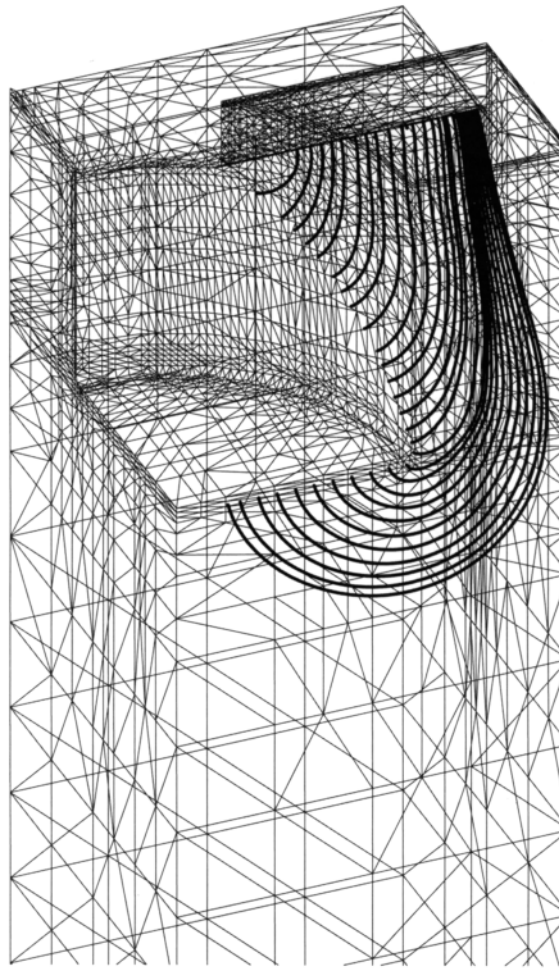
2.19 Dislocation in a section of a partially processed wafer.

memory device. Figure 2.19 shows an area of a chip which has developed such a 'killer defect' at an early stage of processing, rendering the entire wafer useless. The stakes are currently being raised, as developers explore a variety of ingenious ways of using high stresses *per se* as a way of improving device performance[123]. In such applications, dislocations are not merely a

problem to be eliminated, they present an unavoidable complication. Failure to predict and control their behavior can carry huge financial penalties.

The microelectronics application is not as specialized as it may seem, since the issues it raises also arise in treating dislocations in nanocrystals, MEMS structures, epitaxial islands, and other nanoscale confined systems. In all instances, the challenge is to understand the nucleation and evolution of just a few dislocations in a very complicated environment. For this kind of problem a DDD simulation must be able to propagate dislocations in an arbitrarily defined geometry, and to accept arbitrarily complicated stress fields from a concurrent or ancillary finite-element calculation (Fig. 2.20). In addition, the microelectronics problem tends to involve a variety of different interfaces between crystalline and non-crystalline materials. Thus, the calculation of image effects is prohibitively complicated, and one is forced to approximate them as discussed in Section 2.3.4. As we have been at pains to emphasize (and as illustrated by Fig. 2.8), this is not necessarily a bad approximation. Indeed, it is found in practice that the DD simulation is the easy part of the problem: a greater limitation is the determination of the stress fields in a particular device after it has undergone processes such as deposition, implantation, oxidation, etching, and chemical-mechanical polishing. At present, these stress fields are only crudely estimated by current process-modeling codes, so that agreement on the order of that seen in Fig. 2.21 has actually come as an encouraging surprise.

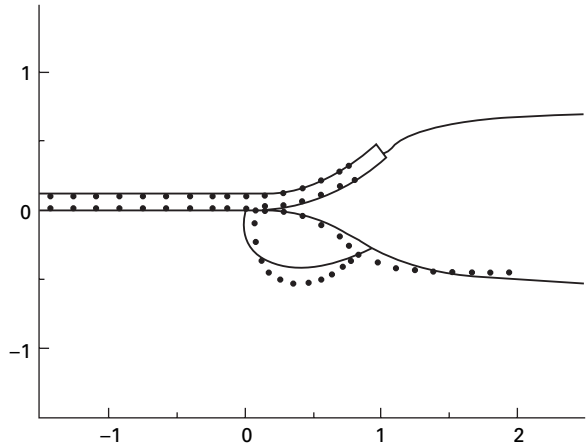
In the microelectronics context, dislocations can be observed to nucleate at various edges and corners. One can apply DDD simulations to this phenomenon using the following strategy. Given the geometry of interest and assuming that one knows the associated stress field, one can place dislocation test loops at various locations. If the loop is large enough, it will grow into a stable configuration. If it is too small, it will shrink and disappear. The unstable stationary state separating these regimes represents the transition configuration which must be surmounted to activate dislocation nucleation. By finding the size and location of the smallest transition configuration, one can estimate whether dislocation nucleation is likely to occur and where it will take place. By following the evolution of a nucleated loop, one can determine the final dislocation configuration produced by the nucleation event. Since the critical loop sizes involved in thermal nucleation are likely to be on the nanometer level, such estimates are only qualitative. Nevertheless, it has been demonstrated[127] that the 'test loop' strategy is of considerable utility (Fig. 2.22). Beyond this heuristic approach, the atomistic study of dislocation nucleation has recently drawn great interest. Such studies have only been done for dislocations nucleated by nanoindenters, and Fig. 2.22 raises the intriguing prospect of using standard processing techniques to provide a simpler environment in which to study nucleation in a controlled way.



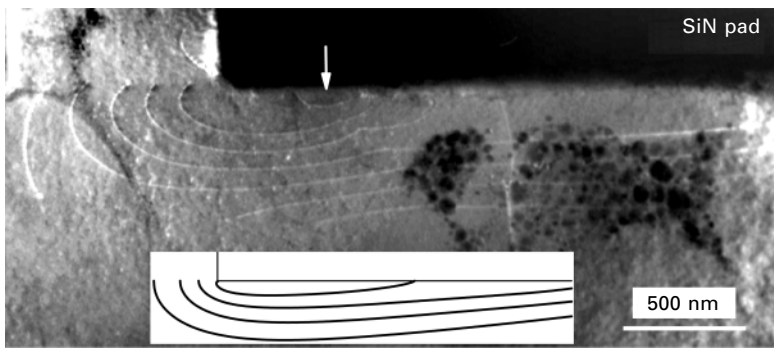
2.20 State of the art DDD simulation of the evolution of a dislocation in a complicated silicon structure. The stress calculation is done via the Finite Element Method. Only the surface tiles are shown. The structure represents one-fourth of an idealized dynamic random access memory (DRAM) structure.

2.5 Extensions of current discrete dislocation dynamics trends

There are many other DDD topics that are of current concern, or remain to be explored. These can perhaps be divided into three main categories. The first is the extension of simulations to include important physical effects which are either well understood theoretically, or for which good phenomenological models exist. Cross-slip[129] is already included in most



2.21 Wedge-shaped 'bird's-beak' structure formed by the oxidation of a silicon wafer partially covered by a silicon nitride mask. Very large stresses are generated as the oxidation process penetrates under the mask and pushes the nitride mask upwards, causing dislocations to be nucleated. The solid lines show the bird's beak shape predicted by the process-modeling code TSUPREM4[124], as well as the equilibrium dislocation loop calculated via DDD simulation[125]. The simulation used the geometry and stress-fields produced by the process-modeling calculation as input. The dots are obtained by scanning a TEM image of an actual structure grown with the process conditions assumed in the calculations[126].



2.22 (a) Dislocation pileup produced by a 0.5 μm thick 10 μm $\times$ 10 μm Si_3N_4 pad on a silicon substrate. The pad exerts an intrinsic stress of order 1 GPa on the Si substrate, generating a non-uniform field in the substrate large enough to nucleate dislocations at the pad edges and corners. The pad (dark area) is shown in plan view. The inset shows the pileup structure predicted by a DDD simulation allowing repeated loop nucleation at the most favorable point indicated by the loop-test procedure. The stress field was computed using the techniques described in Ref.[128], for the pad dimensions and intrinsic stress appropriate to the figure. The estimated critical loop radius was of order 2 nm.

codes, but it is not difficult to list other important mechanisms that need similar attention. Thus dislocation dissociation is important in the context of modeling strong dislocation interactions more completely[55,130,131], it is a vital mechanism in determining the properties of superalloys[132], and it is central to understanding the evolution of stacking faults in strained semiconductor layers[133–135]. Dislocation climb[136] is of importance in a number of contexts, especially those involving the plastic response of crystals damaged by radiation or by ion implantation. Dislocation pinning by local obstacles[137–139] and dislocation freezing by the gettering of impurities[140,141] are also important in these contexts.

The second area represents opportunities to improve the absolute accuracy of the simulations. This includes the formally well-defined tasks of including elastic anisotropy, accurate long-range fields, and image corrections. As we have discussed, the third of these has been well studied for free surfaces, but the case of elastically mismatched interfaces, important in small structures and in metal films, has been largely ignored. As we have emphasized above, all of these effects are often on the scale of the current core error. Thus, a more important task at present may be to match DDD simulations done with simple configurations to atomistic calculations to see if the effective core radius can be determined. Also needed are better numbers for mobility, cross-slip probabilities, and dislocation–obstacle interactions, presumably derived from direct measurement or from atomistic calculations[138,142–144]. An interesting commentary on the issue of connecting microscopic calculations to DDD simulations has been given by Bulatov[145].

Lastly, there are important phenomena that we do not yet know how to include in a DDD simulation at all. Primary among these is dislocation nucleation, which is currently the subject of intense investigation at the fundamental level[146–148]. DDD simulations may be able to tell us where dislocations are most easily nucleated and what they do after they appear, but at present they can predict the behavior of a stressed material only in terms of the evolution of a known initial dislocation structure. How a single crystal of pure copper subject to a nanoindenter, or a strained SiGe layer taken to annealing temperatures, come to be filled with dislocations is still a mystery. One factor known to promote dislocation nucleation is crystal damage, from radiation, ion implantation, or mechanical surface damage. In the first two cases, it is common to form small prismatic loops[149,150], which apparently can unfault[151] to nucleate glide dislocations. The processes by which crystal damage, which tends primarily to form sessile defects, can nucleate glide behavior are poorly understood and need further atomistic study before they can be included in DDD simulations. The presence of interfaces or free surfaces also leads to intriguing unresolved issues. It is not known for instance what the structure of a dislocation lying in a crystal/oxide interface should be. The interaction of dislocations with grain boundaries is not well understood.

Finally it is known that the dislocation properties themselves can change at an interface, and indeed just the addition of an impurity layer to a free surface can drastically change the plastic behavior of both semiconductor and metal layers[152,153].

Among the various scales dealt with by multiscale modeling, dislocation dynamics has an especially heterogeneous character dictated by the great variety of different things a dislocation can do, each of which must be included in a phenomenological way in any simulation. Since DDD simulation is also the youngest field in the multiscale modeling canon, it is not surprising that activity continues at a fever pitch. As the laundry list of topics presented in this section shows, an enormous amount of work remains to be done before the field of DDD simulation can be said to have reached its full potential.

2.6 Acknowledgments

The author wishes to thank L. Kubin, W. Cai, H. Mughrabi, O. Kraft, N. Ghoniem, and J. Benedict for their assistance and comments.

2.7 References

1. J.P. Hirth and J. Lothe, *Theory of Dislocations*, New York: Wiley, 1982.
2. T. Mura, *Micromechanics of Defects in Solids*, Dordrecht: Kluwer, 1982.
3. D. Hull and D.J. Bacon, *Introduction to Dislocations*, Oxford: Butterworth-Heinemann, 1984.
4. V.L. Indenbohm and J. Lothe eds *Elastic Strain Fields and Dislocation Mobility*, Amsterdam: North-Holland, 1992.
5. V.V. Bulatov and W. Cai, *Computer Simulation of Dislocations*, Oxford: Oxford University Press, 2006.
6. L.P. Kubin and B. Devincre, *Mesoscale Simulations of Dislocations and Plasticity*, Oxford: Oxford University Press, forthcoming.
7. J. Lothe, *Phil. Mag*. **15**, 353 (1967).
8. L.M. Brown, *Phil. Mag*. **15**, 363 (1967).
9. M.O. Peach and J.S. Koehler, *Phys. Rev*. **80**, 436 (1950).
10. J. Lothe, in *Elastic Strain Fields and Dislocation Mobility*, eds V.L. Indenbohm and J. Lothe, Amsterdam: North-Holland, 1992, 175.
11. D.J. Bacon, *Phys. Stat. Sol*. **23**, 527 (1967).
12. A.J.E. Foreman, *Phil. Mag*. **15**, 1011 (1967).
13. A.J.E. Foreman, P.B. Hirsch and F. Humphreys, in *Fundamental Aspects of Dislocation Theory*, eds J.A. Simmons, R. de Wit and R. Bullough, NBS pub. no. 317, Washington DC: National Bureau of Standards, 1970, 1083.
14. A.A. Shtolberg, *Phys. Stat. Sol*. **43**, 523 (1971).
15. D.J. Bacon, U.F. Kocks and R.O. Scattergood, *Phil. Mag*. **28**, 1241 (1973).
16. R.O. Scattergood and D.J. Bacon, *Acta Met*. **30**, 1665 (1982).
17. M.S. Duesbery and K. Sadananda, *Phil. Mag. A* **63**, 535 (1991).
18. M.S. Duesbery, N.P. Louat, and K. Sadananda, *Phil. Mag. A* **65**, 311 (1992).

19. M.S. Duesbery, N.P. Louat and K. Sadananda, *Acta Metall. Mater*. **40**, 149 (1992).
20. L.P. Kubin, G.R. Canova, M. Condat, B. Devincre, V. Pontikis and Y. Brechet, *Solid State Phenom*. **23/24**, 455 (1992).
21. G.R. Canova, Y.J.M. Brechet and L.P. Kubin, *Solid State Phenom*. **34**, 101 (1993).
22. B. Devincre and L.P. Kubin, *Modell. Simul. Mater. Sci. Eng*. **2**, 559–570 (1994).
23. H.M. Zbib, M. Rhee and J.P. Hirth, in *Advances in Engineering Plasticity and its Applications*, eds T. Abe and T. Tsuta, New York: Elsevier, 1996, 15.
24. H.M. Zbib, M. Rhee and J.P. Hirth, *Int. J. Mech. Sci*. **40**, 113–127 (1998).
25. M. Rhee, H.M. Zbib, J.P. Hirth, H. Huang and T. Diaz de la Rubia, *Modell. Simul. Mater. Sci. Eng*. **6**, 467–492 (1998).
26. K.W. Schwarz and J. Tersoff, *Appl. Phys. Lett*. **69**, 1220 (1996).
27. K.W. Schwarz, *Phys. Rev. Lett*. **78**, 4785 (1997).
28. K.W. Schwarz and F.K. LeGoues, *Phys. Rev. Lett*. **79**, 1877 (1997).
29. K.W. Schwarz, *J. Appl. Phys*. **85**, 108 (1999).
30. J.C. Ramirez, A.F. Bower and L.B. Freund, Brown University Report (1990).
31. R.V. Kukta, Ph.D. Thesis, Brown University (1998).
32. X.J. Xin, R.H. Wagoner and G.S. Daehn, *Scripta Materialia* **39**, 397 (1998).
33. D. Weygand, L.H. Friedman, E. Van der Giessen, and A. Needleman, *Modell. Simul. Mater. Sci. Eng*. **10**, 437 (2002).
34. D. Weygand and P. Gumbsch, *Mater. Sci. Eng. A* **400–401**, 158–161 (2005).
35. N.M. Ghoniem and L.Z. Sun, *Phys. Rev. B* **60**, 128 (1999).
36. N.M. Ghoniem, S.-H. Tong and L.Z. Sun, *Phys. Rev. B* **61**, 913 (2000).
37. V.V. Bulatov, W. Cai, J. Fier, M. Hiratani, T. Pierce, M. Tang, M. Rhee, K. Yates and A. Arsenlis, in *Supercomputing 2004*, Washington DC: IEEE Computer Society, 2004, 19.
38. W. Cai, V.V. Bulatov, T.G. Pierce, M. Hiratani, M. Rhee, M. Bartelt and M. Tang, in *Solid Mechanics and its Applications*, eds H. Kitagawa and Y. Shibutani, Dordrecht: Kluwer, 2004, Vol. 115, 1–11.
39. Z. Wang, N. Ghoniem, S. Swaminarayan and R. LeSar, *J. Comp. Phys*. **219**, 608 (2006).
40. R.J. Amodeo and N.M. Ghoniem, *Phys. Rev. B* **41**, 6958 (1990).
41. H.Y. Wang and R. LeSar, *Phil. Mag. A* **71**, 149 (1995).
42. A.A. Benzerga, Y. Brechet, A. Needleman and E. Van der Giessen, *Modell. Simul. Mater. Sci. Eng*. **12**, 159 (2004), and references contained therein.
43. D. Raabe, *Z. Metallkd*. **6**, 493–497 (1996).
44. M. Rhee, J.S. Stölken, V.V. Bulatov, T. Diaz de la Rubia, H.M. Zbib and J. Hirth, *Mater. Sci. Eng*. **A309–310**, 288 (2001).
45. X. Han and N.M. Ghoniem, *Phil. Mag*. **85**, 1205 (2005).
46. H. Huang, N. Ghoniem, T. Diaz de la Rubia, M. Rhee, H. Zbib and J.P. Hirth, *ASME J. Eng. Mater. Technol*. **121**, 143 (1999).
47. A. Walcker, D. Weygand and O. Kraft, *Mat. Sci. Eng. A* **400–401** 397 (2005).
48. W. Cai, V.V. Bulatov, J. Chang, J. Li and S. Yip, in *Dislocations in Solids*, eds F.R.N. Nabarro and J.P. Hirth, Amsterdam: North-Holland, 2004, Vol. 12, 1.
49. T.J.R. Hughes, in *Computational Methods for Transient Analysis*, eds T. Belytschko and T.J.R. Hughes, Amsterdam; North-Holland, 1983, 67.
50. W.H. Press, S.A. Teukolsky, W.F. Vetterling and B.P. Flannery, *Numerical Recipes*, Cambridge: Cambridge University Press, 1992.
51. S.D. Gavazza and D.M. Barnett, *J. Mech. Phys. Solids* **24**, 171 (1976).
52. W. Cai, A. Arsenlis, C.R. Weinberger and V.V. Bulatov, *J. Mech. Phys. Solids* **54**, 561 (2006).

53. L.M. Brown, *Phil. Mag*. **10**, 441 (1964).
54. D. Kuhlmann-Wilsdorf, *Mat. Res. Innovat*. **1**, 265 (1998).
55. V.B. Shenoy, R.V. Kukta and R. Phillips, *Phys. Rev. Lett*. **84**, 1491 (2000).
56. L.K. Wickham, K.W. Schwarz and J.S. Stölken, *Phys. Rev. Lett*. **22**, 4574 (1999).
57. R. Madec, B. Devincre and L.P. Kubin, *Comp. Mat. Sci*. **23**, 219 (2002).
58. R. Madec and L.P. Kubin, in *IUTAM Symposium on Mesoscopic Dynamics of Fracture Process and Materials Strength*, eds H. Kitagawa and Y. Shibutani, Dordrecht: Kluwer, 2004, 69.
59. K.W. Schwarz, *Modell. Simul. Mater. Sci. Eng*. **11**, 609 (2003).
60. Z. Wang, N. Ghoniem and R. LeSar, *Phys. Rev. B* **69**, 174102 (2004).
61. L. Greengard and V. Rokhlin, *J. Comp. Phys*. **73**, 325 (1987).
62. I. Groma and B. Bakó, *Phys. Rev. Lett*. **84**, 1487 (2000).
63. M. Zaiser, M.-Carmen Miguel and I. Groma, *Phys. Rev. B* **64**, 224102 (2001).
64. G. Saada, *Acta Metall*. **8**, 841 (1960).
65. D. Gómez-García, B. Devincre and L.P. Kubin, *Phys. Rev. Lett*. **96**, 125503 (2006).
66. E.H. Yoffe, *Phil. Mag*. **6**, 1147 (1961).
67. S.J. Shaibani and P.M. Hazzledine, *Phil. Mag*. **44**, 657 (1981).
68. Yu. A. Belov, in *Elastic Strain Fields and Dislocation Mobility*, eds V.L. Indenbohm and J. Lothe, Amsterdam: North-Holland, 1992, 39.
69. E.H. Tan and L.Z. Sun, *MRS Proc*. **795**, 47–52 (2004).
70. R. Martinez and N.M. Ghoniem, *J. Comp. Meth. Eng. Sci*. **3(2)**, 229 (2002).
71. M. Tang, G. Xu, W. Cai and V.V. Bulatov, *MRS Proc.* **795**, 29 (2004).
72. X.J. Xin, K. Kollu and R.H. Wagoner, *Mater. Sci. Eng. A* **309**, 520 (2001).
73. T.A. Khraishi and H.M. Zbib, *J. Eng. Mater. Technol*. **124**, 342 (2002).
74. L. Yan, T.A. Khraishi, Y.L. Shen and M.F. Horstemeyer, *Modell. Simul. Mater. Sci. Eng*. **12**, S289 (2004).
75. J. Boussinesq, *Application des potentiels à l'étude de l'équilibre et du mouvement des solides élastiques*, Paris: Gauthier-Villars, 1885, 45.
76. V. Cerruti *Roma, Acc. Lincei, Mem. fis. mat.* (1882).
77. S.P. Timoshenko and J.N. Goodier, *Theory of Elasticity*, New York: McGraw-Hill, 1970, 398.
78. M.C. Fivel, T.J. Gosling and G.R. Canova, *Modell. Simul. Mater. Sci. Eng*. **4**, 581 (1996).
79. X.H. Liu and K.W. Schwarz, *Modell. Simul. Mater. Sci. Eng*. **13**, 1233 (2005).
80. J. Lothe, in *Fundamental Aspects of Dislocation Theory*, eds J.A. Simmons, R. de Wit and R. Bullough, NBS Pub. no. 317, Washington DC, National Bureau of Standards, 1970, 11.
81. M. Tang, W. Cai, G. Xu and V.V. Bulatov, *Modell. Simul. Mater. Sci. Eng*. **14**, 1139 (2006).
82. Z. Wang, R.J. McCabe, N.M. Ghoniem, R. LeSar, A. Misra and T.E. Mitchell, *Acta Mater*. **52**, 1535 (2004).
83. X.H. Liu, F.M. Ross and K.W. Schwarz, *Phys. Rev. Lett*. **85**, 4088 (2000).
84. K.W. Schwarz, *Phys. Rev. B*, **38**, 2398 (1988).
85. R. Madec, B. Devincre and L.P. Kubin in *IUTAM Symposium on Mesoscopic Dynamics of Fracture Process and Materials Strength*, eds H. Kitagawa and Y. Shibutani, Dordrecht, Kluwer, 2004, 35–44.
86. K.W. Schwarz (unpublished).
87. W. Cai, V.V. Bulatov, J. Chang, J. Li and S. Yip, *Phil. Mag*. **83**, 539–567 (2003).
88. R.W.K. Honeycombe, *The Plastic Deformation of Metals*, London: Edward Arnold, 1984.

89. J.G. Sevillano, in *Plastic Deformation and Fracture of Materials (Materials Science and Technology Vol. 6*), eds H. Mughrabi Weinheim: VCH, 1993, 40.
90. H. Mughrabi, T. Ungár, W. Kienle and M. Wilkens, *Phil. Mag*. **53**, 793 (1986).
91. V.V. Bulatov, M. Tang and W. Cai (unpublished).
92. V.V. Bulatov, L.L. Hsiung, M. Tang, A. Arsenlis, M.C. Bartelt, W. Cai, J.N. Florando, M. Hiratani, M. Rhee, G. Hommes, T.G. Pierce and T. Diaz de la Rubia, *Nature* **440**, 1174 (2006).
93. P.J. Jackson, *Prog. Mat. Sci*. **29**, 139 (1985).
94. R. Madec, B. Devincre, L. Kubin, T. Hoc and D. Rodney, *Science* **301**, 1879 (2003).
95. B. Devincre, T. Hoc and L.P. Kubin, *Mater. Sci. Eng. A* **400–401**, 182 (2005).
96. E.A. Stach, K.W. Schwarz, R. Hull, F.M. Ross and R.M. Tromp, *Phys. Rev. Lett*. **84**, 947 (2000).
97. K.W. Schwarz, *Phys. Rev. Lett*. **93**, 145503 (2003).
98. See the collection of articles in *Mechanical Properties in Small Dimensions*, eds R.P. Vinci and S.P. Baker, *MRS Bull*. **27**, 1 (2002).
99. L.B. Freund and S. Suresh, *Thin Film Materials*, Cambridge: Cambridge University Press, 2003.
100. R.R. Keller, J.M. Phelps and D.T. Read, *Mater. Sci. Eng. A* **288**, 42 (1996).
101. F.C. Frank and J.H. van der Merwe, *Proc. Roy. Soc*. (London) **A198**, 216 (1949).
102. J.W. Matthews and A.E. Blakeslee, *J. Crystal Growth* **27**, 118 (1974).
103. L.B. Freund, *J. Appl. Mech*. **54**, 553 (1987).
104. L.B. Freund, *J. Appl. Phys*. **68**, 2073 (1990).
105. V.T. Gillard, W.D. Nix and L.B. Freund, *J. Appl. Phys*. **76**, 7280 (1994).
106. K.W. Schwarz, J. Cai and P.M. Mooney, *Appl. Phys. Lett*. **85**, 2238 (2004).
107. D. Weiss, H. Gao and E. Arzt, *Acta Mater*. **49**, 2395 (2001).
108. O. Kraft, L.B. Freund, R. Phillips and E. Arzt, *MRS Bull.* **27**, 30–37 (2002).
109. M. De Koning, R. Miller, V.V. Bulatov and F.F. Abraham, *Phil. Mag. A*, **82**, 2511 (2002).
110. M. de Koning, R.J. Kurtz, V.V. Bulatov, C.H. Henager, R.G. Hoagland, W. Cai and M. Nomura, *J. Nucl. Mat*. **323**, 281–289 (2003).
111. A. Hartmaier, M.J. Buehler and H. Gao, *Defect and Diffusion Forum* **224–225**, 107 (2003).
112. A. Hartmaier, M.J. Buehler and H. Gao, *Mater. Sci. Eng A* **400–401**, 260 (2005).
113. G. Dehm, T.J. Balk, B. von Blanckenhagen, P. Gumbsch and E. Arzt, *Zeitschrift fur Metallkunde (Germany)* **93**, 383 (2002).
114. B. von Blanckenhagen, P. Gumbsch and E. Arzt, *Phil Mag. Lett*. **83**, 1 (2003).
115. B. von Blanckenhagen, E. Arzt and P. Gumbsch, *Acta Materialia* **52**, 773 (2004).
116. P. Pant, K.W. Schwarz and S.P. Baker, *Acta Materialia* **51**, 3243 (2003).
117. R.S. Fertig, P. Pant, K.W. Schwarz and S.P. Baker (unpublished).
118. R. Venkatraman, P.W. Davies, P.A. Flinn, D.B. Fraser, J.C. Bravman and W.D. Nix, *J. Electron. Mater*. **19**, 1231 (1990).
119. B.W. Lagow, I.M. Robertson, M. Jouiad, D.H. Lassila, T.C. Lee and H.K. Birnbaum, *Mater. Sci. Eng. A* **309–310**, 445 (2001).
120. I.M. Robertson, A. Beaudoin, K. Al-Fadhalah, L. Chun-Ming, J. Robach, B.D. Wirth, A. Arsenlis, D. Ahn and P. Sofronis, *Mater. Sci. Eng. A* **400–401**, 245 (2005).
121. H.D. Espinosa, S. Berbenni, M. Panico and K.W. Schwarz, *Proc. Natl. Acad. Sci*. **102**, 16933 (2005).

122. P.M. Fahey, S.R. Mader, S.R. Stiffler, R.L. Mohler, J.D. Mis and J.A. Slinkman, *IBM J. Res. Dev*. **36**, 158 (1992).
123. *Advanced Semiconductor Manufacturing Conference and Workshop, 2005 IEEE/SEMI* (IEEE Product No. CH37654).
124. *TSUPREM4 User's Manual Version 1999.4* (Avant! Corporation, Fremont, CA, 1999).
125. D. Chidambarrao, X.H. Liu and K.W. Schwarz, *J. Appl. Phys*. **92**, 6278 (2002).
126. J. Vanhellemont, S. Amelinckx and C. Claeys, *J. Appl. Phys*. **61**, 2176 (1987).
127. M. Kammler, D. Chidambarrao, K.W. Schwarz, C.T. Black and F.M. Ross, *Appl. Phys. Lett*. **87**, 133116 (2005).
128. K.W. Schwarz and D. Chidambarrao, *J. Appl. Phys*. **85**, 7198 (1999).
129. W. Püschl, *Prog. Mater. Sci*. **47**, 415 (2002).
130. C. Shin, M. Fivel, D. Rodney, R. Phillips, V.B. Shenoy and R. Dupuy, *J. Phys. IV (France)* **11**, 19 (2001).
131. D. Weygand, in *IUTAM Symposium on Mesoscopic Dynamics of Fracture Process and Materials Strength*, eds H. Kitagawa and Y. Shibutani, Dordrecht: Kluwer, 2004, 23.
132. S.I. Rao, T.A. Parthasarathy, D.M. Dimiduk and P.M. Hazzledine, *Phil. Mag*. **84**, 3195–3215 (2004).
133. P.M.J. Marée, J.C. Barbour, J.F. Van der Veen, C.W.T. Bulle-Lieuwma and M.P.A. Viegers, *J. Appl. Phys*. **62**, 4413 (1987).
134. S.W. Bedell, K. Fogel, D.K. Sadana, H. Chen and A. Domenicucci, *Appl. Phys. Lett*. **85**, 2493 (2004).
135. Y. Kimura, N. Sugii, S. Kimura, K. Inui and W. Hirasawa, *Appl. Phys. Lett*. **88**, 031912 (2006).
136. D. Raabe, *Phil. Mag. A* **77**, 751–760 (1998).
137. M. Hiratani, H.M. Zbib and M.A. Khaleel, *Int. J. Plast*. **9**, 1271 (2003).
138. D.J. Bacon and Yu. N. Osetsky, *Mater. Sci. Eng. A* **400–401**, 353 (2005).
139. C.S. Shin, M.C. Fivel, M. Vwerdier and C. Robertson, *Mater. Sci. Eng. A*, **166** (2005).
140. S.M. Hu and W.J. Patrick, *J. Appl. Phys* **46**, 1869 (1975).
141. I.V. Peidous, K.L. Loiko, D.A. Simpson, T. La and W.R. Frensley, *Mat. Res. Soc. Proc*. **673**, P3.9.1 (2001).
142. T. Vegge, T. Rasmussen, T. Leffers, O.B. Pedersen and K.W. Jacobsen, *Phys. Rev. Lett*. **85**, 3866 (2001).
143. T. Vegge, T. Rasmussen, T. Leffers, O.B. Pedersen and K.W. Jacobsen, *Phil. Mag. Lett*. **81**, 137 (2001).
144. W. Cai and V.V. Bulatov, *Mater. Sci. Eng A* **387–389**, 277 (2004).
145. V.V. Bulatov, *J. Comp.-Aided Mat. Design* **9**, 133 (2002).
146. K.J. Van Vliet, J. Li, T. Zhu, S. Yip and S. Suresh, *Phys. Rev. B* **67**, 104105 (2003).
147. G. Xu and C. Zhang, *J. Mech. Phys. Sol*. **51**, 1371 (2003).
148. G. Xu in *Dislocations in Solids, Vol. 12*, eds F.R.N. Nabarro and J.P. Hirth, Amsterdam: Elsevier, 2004, 83.
149. M.W. Thompson, *Defects and Radiation Damage in Metals*, Cambridge: Cambridge University Press, 1969.
150. K.S. Jones, S. Prussin and E.R. Weber, *Appl. Phys. A* **45**, 1 (1988).
151. A. Kubota and W.G. Wolfer, *Mater. Sci. Eng. A* **400–401**, 362 (2005).
152. E.A. Stach, R. Hull, R.M. Tromp, M.C. Reuter, M. Copel, F.K. LeGoues and J.C. Bean, *J. Appl. Phys*. **83**, 1931 (1998).
153. S.P. Baker, R.-M. Keller-Flaig and J.B. Shu, *Acta Mater*. **51**, 3019–3036 (2003).

3

Phase-field modelling of material microstructure

L.-Q. CHEN, Penn State University, USA

3.1 Introduction

When speaking about structures within the context of processing/structures/properties relationships in materials science and engineering, this almost always refers to the microstructures of materials. Different processing conditions lead to different microstructures and thus to variations in properties. A microstructure may contain a wide variety of structural features such as phases of different compositions and/or crystal structures, grains of different orientations, domains of different structural variants, domains of different electrical or magnetic polarizations, as well as structural defects such as interphase boundaries, grain boundaries, domain walls, cracks, surfaces, and dislocations. The length scales of these structural features range through anstrons (crack tips), nanometers (interfacial width, dislocation core, nuclei, small domains and grains), and microns (grains and domains).

Microstructures evolve during materials processing or in service at high temperatures as a result of phase transformations or particle or domain coarsening. The common processing variables are temperature and composition. Microstructures can also be modified by external fields such as an applied stress or electrical or magnetic field. The time scale for microstructure evolution in materials typically spans from seconds to days or even months. One of the main goals for materials design is to capture the optimum microstructures having the most desirable properties during processing and to freeze their evolution during applications.

As one can imagine, finding an optimum microstructure requires exhaustive and costly experimentations in the processing parameter space of temperature, composition, and time. As a result, there has been increasing interest in the recent past in utilizing computer simulations and modeling to reduce the number of time-consuming experiments. Due to the complexity of microstructures and their evolution, as well as the time and spatial scales that are involved, however, one cannot rely on a single available computational method to directly predict the microstructure evolution of a system starting

with arbitrarily specified initial state, temperature, composition, and time. The discussions in this chapter will be limited to computer simulations at the microstructure level, which requires the knowledge of structures and properties of individual structural features in a microstructure, much of which can, in principle, be obtained using atomistic level first-principles calculations discussed in other chapters in this book. In particular, this chapter discusses one of the microstructure evolution models, the phase-field approach for modeling microstructure evolution processes.

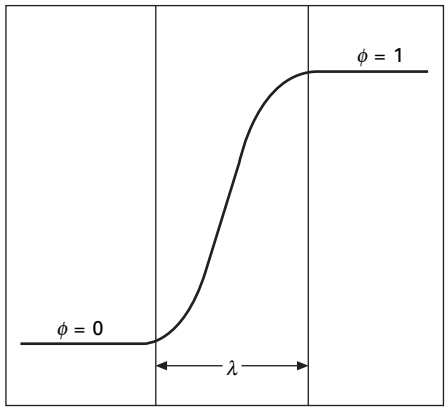
The phase-field approach has received tremendous attention since the early 1990s due to its flexibility to model a wide variety of materials processes and microstructure evolution. A phase-field model requires information on the structures and properties of individual structural features in a microstructure as input and predicts the microstructure evolution based on fundamental thermodynamic and kinetic principles. This chapter is not intended to be a review of the phase-field method; it serves rather as a brief explanation of the method. A number of reviews on the method have already appeared in the last few years[1–3]. Because of the personal interest of the author, the discussions are focused on solid-state processes and microstructure evolution.

3.2 Model description

3.2.1 Representation of a microstructure

In order to model the formation and evolution of a microstructure, it is necessary to be able to digitize a microstructure on a computer. In a phase-field model, a set of continuous fields is employed to represent a microstructure. These are typically uniform inside a phase or domain sufficiently far away from the interfaces or walls. The field variables have the same uniform values in the same phase or the same types of domains within a given microstructure. Different values of the field variables, for example 0 and 1, distinguish different phases or domains. Across the interfaces between different phases or domains, the field variables vary continuously from one uniform value corresponding to one type of phase or domain to another uniform value corresponding to another phase or domain. Therefore, the interfaces in a phase-field model are diffuse and possess a certain thickness. The variation of a field variable across an interface is schematically shown in Fig. 3.1.

Field variables can be either conserved or non-conserved, depending on whether they satisfy the local conservation law, $\partial\phi/\partial t = -\nabla \cdot \boldsymbol{J}$ where ϕ is a field variable and $\boldsymbol{J}$ is the corresponding flux. For example, composition and temperature fields are both conserved while long-range order parameter fields describing ordered domain structures are non-conserved. It is easy to understand that the artificial phase field in solidification modeling of a single-component liquid is non-conserved since its value can go from 0 to 1 for the whole



3.1 A schematic diagram showing the distribution of a field variable across an interface. $\phi = 0$ and $\phi = 1$ represent two phases separated by a diffuse interface with a thickness roughly λ.

system. Conserved and non-conserved variables require different kinetic laws for their evolution as discussed below.

Field variables can be physical or artificial. Physical fields refer to well-defined order parameters which can be experimentally measured. The interfacial width described by a physical field is also expected to reflect the actual interfacial width in a microstructure. In phenomenological theories of phase transformations, order parameters are used to characterize the nature and the critical temperatures of phase transformations which produce the microstructures. A well-known example is the long-range order parameter for order–disorder transformations. The corresponding order parameter field can be employed to describe the antiphase domain structures which result from ordering. Another example is a composition field which describes a two-phase microstructure with differences in compositions between the two phases. A composition field is also sufficient to describe the morphological evolution during phase separation through either nucleation and growth or spinodal decomposition or during precipitate coarsening in a binary alloy. Other examples of physical order parameters include electric polarization for a ferroelectric phase transition and magnetization for a ferromagnetic phase transition. On the other hand, artificial fields are introduced for the sole purpose of avoiding tracking the interfaces during a microstructure evolution. Essentially all phase-field models of solidification employ an artificial field called the 'phase field'. The interfacial width described by artificial fields has no directional relationship to the physical width of a real interface. The thermodynamic and kinetic coefficients in the phase-field equations are chosen to match the corresponding parameters in the conventional sharp-interface equations through sharp- or thin-interface analyses[4–8].

Field variables can be a scalar or a vector, and they can be a single component or multicomponent. For example, a long-range order parameter for order–disorder transition and composition for phase separation are scalars while polarization and magnetization are polar and axial vectors, respectively. An ordered phase such as the B2 phase based on a body-centered-cubic (bcc) lattice can be characterized by a single scalar long-range order parameter while the physical characterization of an $L1_2$ ordered phase on a face-centered-cubic (fcc) lattice requires a scalar order parameter with three components[9–12]. Many examples of microstructures require more than one type of order parameters. For example, precipitation of an ordered intermetallic phase in a disordered matrix, a fundamental process in many technologically important alloy systems such as Al-alloys for automotive applications and Ni-base superalloys for aerospace applications, involves both ordering and compositional clustering, and thus the characterization of the resulting precipitate microstructures requires both composition and order parameter fields. Another example is a composite microstructure of ferroelectric and ferromagnetic crystals which requires two types of field variables, the electric polarization and magnetization. Some examples of field variables for a number of common phase transformations and microstructures are listed in Table 3.1.

3.2.2 Thermodynamics of microstructures

Within the diffuse-interface context[13], the total free energy of an inhomogeneous microstructure is written as a functional of all the field variables that characterize the phase transitions, domain structures, and defects. In general, it contains four types of energetic contributions, i.e.

Table 3.1 Examples of field variables

Microstructure	Field variables
A isostructural two-phase system either from nucleation and growth or from spinodal decomposition in a binary system	Composition $c(r)$
Order–disorder transitions and antiphase domain structures	Long-range order parameter $\eta(r)$
Precipitation of ordered intermetallic phases in a disordered matrix in a binary alloy	$c(r)$ and $\eta\ (r)$
Ferroelectric transitions and domain structures	$P_i(r)$
Ferromagnetic transformations and magnetic domain structures	$M_i(r)$
Martensitic transformations and martensite microstructures	ε_{ij}°

$$F = \int_V \lfloor f_{\text{local}} + f_{\text{gra}} + f_{\text{appl}} \rfloor \, \mathrm{d}V + \int_{V,V'} [f_{\text{non-local}}] \, \mathrm{d}V\mathrm{d}V' \qquad [3.1]$$

where f_{local} is the local bulk chemical free energy density that is a function of one or more of the order parameters: c_i (composition of component i), η_{I} (long-range order parameters), p_i (polarization component i), m_i (magnetization component i), ε_{ij} (strain component ij), and ϕ_i (order parameters or fields describing the distribution of grains, dislocations, etc). f_{gra} is the gradient energy density, i.e. the energy penalty for the inhomogeneities in the order parameter fields. It is non-zero only at and around interfaces, and therefore its introduction automatically includes the domain-wall energy contribution. f_{appl} represents the coupling potential energy between applied fields such as applied stress, electric field, or magnetic field and the corresponding order parameters such as strain, polarization, and magnetization. The last term in equation [3.1] includes contributions from any one or more of the long-range interactions such as elastic, electrostatic, and magnetostatic interactions. It is the competition among the different contributions to the total free energy that is responsible for the formation of the many fascinating microstructure patterns observed during various phase transformations. Due to the importance of various energetic contributions in the phase-field modeling, each individual term in equation [3.1] is briefly discussed below.

Local bulk chemical free energy density

The actual bulk chemical free energy as a function of temperature, strain, and composition and/or other thermodynamic parameters is typically not available for most systems. Free energies cannot be directly measured experimentally. They are very difficult to compute from first principles, although there have been recent attempts to obtain free energies of pure and binary systems by combining first-principles calculations, lattice dynamics, and statistical thermodynamics[14]. Therefore, most of the exiting phase-field simulations have employed double-well or multi-well functions with regard to the order parameters or the Landau-type of free energy models with the coefficients fitted to experimentally-measured phase transition temperatures, single-crystal properties, etc. The main difference among different phase-field models lies in the construction of f as a function of field variables for different phase transformations and microstructure evolution processes.

To illustrate some examples of free energy functions, let us start with a simple and familiar system of spinodal decomposition and formation of a two-phase binary system described by a composition field. The most commonly employed model is a simple double-well potential:

$$f = -\frac{A}{2} c^2 + \frac{B}{4} c^4 \qquad [3.2]$$

where A and B are constants. At equilibrium, equation [3.2] describes two phases with compositions $+\sqrt{A/B}$ and $-\sqrt{A/B}$, respectively. The spinodal compositions are given by $+\sqrt{A/3B}$ and $-\sqrt{A/3B}$. Another possibility is a regular solution model:

$$f(c) = \varphi c(1-c) + RT\,[c\,\ln(c) + (1-c)\,\ln(1-c)] \qquad [3.3]$$

where R is the ideal gas constant, T is the absolute temperature, and φ is a material constant related to heat of mixing. For an ideal solution, φ is equal to zero. If φ is negative, a single homogeneous phase is stable. Otherwise, a homogeneous state is not stable with respect to phase separation to two phases through either nucleation and growth or spinodal decomposition.

Similarly, for an order–disorder transition described by a long-range order parameter, one could use the following simple model free energy:

$$f = \frac{1}{2}\,\alpha_o(T - T_c)\,\eta^2 + \frac{1}{4}\,\beta\eta^4 \qquad [3.4]$$

where α_o and β are phenomenological coefficients and T_c is the instability temperature at which the high-temperature phase becomes unstable with respect to ordering. The equilibrium value for the order parameter as a function of temperature is given by

$$\eta_e = \pm\sqrt{\alpha_o(T - T_c)/\beta} \qquad [3.5]$$

To illustrate the role of strain in phase transformations in a homogeneous system, let us consider a simple model system with two degenerate states for the product phase and characterized by a single-order parameter, η (Fig. 3.1). The transformation is assumed to be second order and the strain is purely dilatational, i.e. the transformation involves only volume changes ($\delta_{ij}\varepsilon_{ij}$ where i and j are Cartesian indices, 1, 2, and 3). Assuming that the coupling between the strain (ε_{ij}) and the order parameter (η) is linear-quadratic [15], the thermodynamics of the system can be described by

$$f = \frac{1}{2}\,\alpha_o(T - T_c)\,\eta^2 + \frac{1}{4}\,\beta\eta^4 - \gamma\eta^2\delta_{ij}\varepsilon_{ij} + \frac{1}{2}\,\lambda_{ijkl}\varepsilon_{ij}\varepsilon_{kl} \qquad [3.6]$$

where γ represents the degree of coupling between the order parameter and strain and λ_{ijkl} is the elastic modulus tensor. The phenomenological coefficients α_o and β can be obtained by fitting experimentally measured thermodynamic properties under a clamped boundary condition ($\varepsilon_{ij} = 0$). The first two terms represent the bulk free energy as a function of order parameter at $\varepsilon_{ij} = 0$. The third and fourth terms describe the strain contribution to the total free energy with δ_{ij} the Kronecker-delta function defined as

$$\delta_{ij} = \begin{cases} 1 \text{ if } i = j \\ 0 \text{ if } i \neq j \end{cases}$$

Equation [3.6] can also be viewed as an expansion to fourth order with respect to the order parameter and to second order in strain. Including only terms to second order in strain is equivalent to assuming linear elasticity. In the equation, the Einstein summation convention is employed, i.e. repeated indices imply summation.

The transformation strain, also called the eigenstrain or stress-free strain, is given by

$$\varepsilon_{ij}^{o} = s_{ijkl}\delta_{kl}\gamma\eta^2 \qquad [3.7]$$

which can be obtained by minimizing the free energy with respect to strain, ε_{ij}. In equation [3.7], s_{ijkl} is the compliance tensor. It displays a quadratic dependence of strain on the order parameter, a result of linear-quadratic coupling in the free energy expression.

One can also formulate the local free energy as a function of order parameter at a given stress state:

$$f = \frac{1}{2}\alpha_o(T - T_c)\eta^2 + \frac{1}{4}\beta\eta^4 - \gamma\eta^2\delta_{ij}\varepsilon_{ij} + \frac{1}{2}\lambda_{ijkl}\varepsilon_{ij}\varepsilon_{kl} - \sigma_{ij}\varepsilon_{ij} \qquad [3.8]$$

where σ_{ij} is the stress tensor. Eliminating ε_{ij} from equation [3.8], we have the free energy as a function of stress:

$$g = \frac{1}{2}\alpha_o(T - T_c)\eta^2 + \frac{1}{4}\beta'\eta^4 - \frac{1}{2}s_{ijkl}\varepsilon_{ij}\varepsilon_{kl} - \sigma_{ij}\varepsilon_{ij}^{o} \qquad [3.9]$$

where $\beta' = \beta - 2\gamma^2\delta_{ij}s_{ijkl}\delta_{kl}$. The above discussion shows that it is important to specify the mechanical boundary conditions for determining the coefficients in the free energy expression.

For real alloy systems in binary and multicomponent systems, the free energies as a function of composition can, in principle, be obtained using the CALPHAD approach (Calculations of Phase Diagrams)[16,17]. The empirical CALPHAD approach employs regular-solution-type of free energy models and optimizes the thermodynamic parameters using existing experimental data and information from first-principles calculations for pure, binary, and sometimes ternary systems[18–20]. However, it should be cautioned that the non-equilibrium segments of the CALPHAD free energies are not necessarily unique or even correct. Moreover, in many cases, order parameters other than composition are involved in a given process, and hence additional effort is often required to incorporate the thermodynamic database in a phase-field model. There have been numerous efforts to use thermodynamic databases in phase-field simulations[12,21–27]. For example, two types of approaches have been proposed for the case of modeling the morphological evolution and coarsening kinetics of γ' precipitates in Ni-base alloys.

One approach is to construct the thermodynamic database in terms of the site-occupation probabilities of different species on the four sublattices of a

fcc lattice using the CALPHAD approach and convert it to a free energy in terms of physical long-range order parameters and composition that can be directly used in a phase-field simulation[12]. The advantage of such an approach lies in the physical nature of the long-range order parameters that correspond to the *fcc–L*1_2 ordering, and the interfacial width reflects the actual physical width between the precipitates and matrix. The main disadvantage of this approach is its difficulty modeling large three-dimensional microstructures due to the relatively small physical width of the interfaces and the need to resolve the composition and order parameter profiles across such a small interfacial width in a numerical simulation. Furthermore, the approach is limited to systems in which the physical order parameters can be clearly defined based on site occupation probabilities.

The other approach is to borrow the ideas from phase-field models for solidifications using artificial order parameters. Since thermodynamic databases are represented using regular-solution-type of free energy models as a function of composition for each individual phase, artificial phase-fields are required to connect the free energies for all the individual phases to obtain a single free energy function. For the particular example of precipitation of γ' in Ni-base alloys, one can obtain two free energy functions as a function of composition, one for the precipitates and one for the matrix, i.e. $f^{\gamma'}(c_{\gamma'}, T)$ and $f^{\alpha}(c_{\alpha}, T)$[26]. To link the two free energy functions as well as to distinguish the four possible types of antiphase domains for the γ' precipitates, four artificial fields, $\eta_i(x, t)(i = 1, 2, 3, 4)$, are required, in addition to the physical field, composition c. Only one of the order parameters is equal to 1 within a γ' precipitate, and they are all zero in the matrix. They vary smoothly from 1 to 0 across the interface between γ' and α. The single free energy function in terms of composition and artificial fields can then be written as

$$f(c, \eta_1, \eta_2, \eta_3, T) = (1 - h(\eta_1, \eta_2, \eta_3, \eta_4))f^{\alpha}(c_{\alpha}, T) + h(\eta_1, \eta_2, \eta_3, \eta_4)f^{\gamma'}(c_{\gamma'}, T) + g(\eta_1, \eta_2, \eta_3, \eta_4) \quad [3.10]$$

where c_{α} and $c_{\gamma'}$ are the mole fraction of *Al* atoms in α and γ' phases, respectively, T is the absolute temperature, $g(\eta_1, \eta_2, \eta_3)$ is a double- or multi-well potential which is zero at 0 and 1 for the artificial order parameters. The function $h(\eta_1, \eta_2, \eta_3, \eta_4)$ is a monotonously changing function from 0 to 1. It is required to have the following properties:

$$h(0) = 0, h(1) = 1, \mathrm{d}h/\mathrm{d}\eta_i|_{\eta_i=0} = \mathrm{d}h/\mathrm{d}\eta_i|_{\eta_i=1} = 0 \quad [3.11]$$

These properties ensure that the equilibrium values, 0 and 1 for the phase fields, are not affected by the chemical free energies $f^{\gamma'}(c_{\gamma'}, T)$ and $f^{\alpha}(c_{\alpha}, T)$. An example which satisfies the conditions is

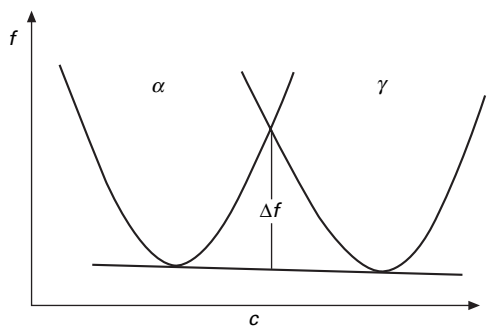
$$h(\eta_1, \eta_2, \eta_3, \eta_4) = 3(\eta_1^2 + \eta_2^2 + \eta_3^2 + \eta_4^2) - 2(\eta_1^3 + \eta_2^3 + \eta_3^3 + \eta_4^3) \quad [3.12]$$

According to Steinbach *et al.*[28] and Kim *et al.*[6], one may regard the interfacial region to be a mixture of the two phases with compositions c_α and $c_{\gamma'}$ and with the same chemical potential, i.e. c_α and $c_{\gamma'}$ satisfy the following set of equations for a binary system

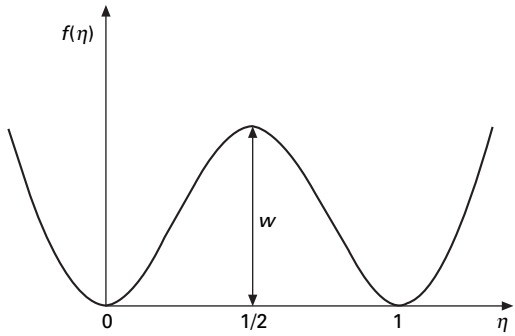
$$c = [1 - h(\eta_1, \eta_2, \eta_3, \eta_4)]\, c_\alpha + h(\eta_1, \eta_2, \eta_3, \eta_4)\, c_{\theta'} \qquad [3.13]$$

$$\frac{\partial f^\alpha(c_\alpha)}{\partial c_\alpha} = \frac{\partial f^{\theta'}(c_{\theta'})}{\partial c_{\theta'}} \qquad [3.14]$$

The main advantage of this approach is the fact that for an interface at equilibrium, the contribution of the actual chemical free energy, $f^{\gamma'}(c_{\gamma'}, T)$ and $f^\alpha(c_\alpha, T)$, i.e. Δf (see Fig. 3.2), to the total interfacial energy, is eliminated at equilibrium because of conditions [3.13] and [3.14]. The interfacial energy and interfacial width are entirely determined by the double- or multiwell potential depth (Fig. 3.3) and the gradient energy coefficient in the artificial phase field. As a result, a larger interfacial width may be employed to fit the same interfacial energy (see the next section for the qualitative relation between interfacial width and the double-well potential depth, equation [3.16]), thus increasing the length scale of a phase-field simulation even with the usual numerical methods using uniform grids. It is also shown that it is reasonably straightforward to extend the model multicomponent systems[29]. However, the implementation of this model requires the numerical solution to the above coupled equations [3.13] and [3.14] for c_α and $c_{\gamma'}$ for each set of c and artificial fields at each time step, and this process can be computationally very expensive. Furthermore, the depth of the double-well potential (w) cannot be made too small compared to the actual chemical driving force (described by $f^{\gamma'}(c_{\gamma'}, T)$ and $f^\alpha(c_\alpha, T)$), i.e. Δf in Fig. 3.2, for phase transformations during a phase-field simulation, and thus the interfacial width that one can use is also limited. Otherwise numerically stability may develop in a phase-field simulation, leading to incorrect path for the



3.2 Illustration of schematic free energy curves of solid and liquid. Δf is the maximum excess chemical free energy across an interface.



3.3 Illustration of a double-well potential with potential depth of *w*.

microstructure evolution. Finally, such a formulation is physically less appealing since the order parameters are artificial, and artificially increased interfacial width may have consequences on the kinetics of microstructure evolution, e.g. the kinetics of coarsening.

Gradient energy and interfacial energy

As mentioned above, in the diffuse-interface description, the free energy of an inhomogeneous system, such as a microstructure, also depends on the gradient energies. The magnitude of gradient energy is specified by the gradient energy coefficient which characterizes the energy penalty due to the field inhomogeneity at the interfaces. Examples of interfaces include solid–liquid interphase boundaries during solidification, grain boundaries, matrix–precipitate interfaces during precipitation reactions and precipitate coarsening, and domain walls in ordered phases, ferroelectric crystals, and ferromagnetic systems. For a given free energy model and a given set of gradient energy coefficients, the specific interfacial energy (interfacial energy per unit area) can be calculated for an equilibrium interface. It is important to realize that the integral of the gradient energy term only counts part of the interfacial energy. The total interfacial energy for a flat interface should be calculated from[13]

$$\gamma = \frac{\Delta F}{S} = \int \left[f_{\text{local}} - f_{\text{equil}} + f_{\text{gra}} \right] \mathrm{d}^3 r = \int \left[\Delta f + f_{\text{gra}} \right] \mathrm{d}^3 r \quad [3.15]$$

where S is the interfacial area and f_{equil} is the bulk equilibrium free energy density as a function of composition represented by the common tangent line for a binary system, or it is the homogeneous bulk free energy density if the flat interface is a ferroic domain wall or a grain boundary. Analytical expressions for the interfacial energy in terms of free energy parameters and the gradient energy coefficients are only available for very simple cases where an analytical

solution for the equilibrium profile of field variable across the interface can be derived. For example, for a simple double-well potential with a single phase-field variable or a simple composition field, the equilibrium profile is described by a hyperbolic tangent function.

For more general cases, the interfacial energy has to be computed numerically. However, in general, the interfacial energy (σ), interfacial width (λ), the total well depth of a free energy model (Δf), and the gradient coefficient (α) obey the following set of relationships:

$$\sigma \propto \sqrt{\alpha \Delta f},\ \lambda \propto \sqrt{\alpha / \Delta f},\ \sigma \sim \lambda \Delta f \qquad [3.16]$$

The last relation in [3.16] is particularly useful in estimating the interfacial width if one has the knowledge on the bulk thermodynamics of a system and the interfacial energy.

In crystalline solids, interfacial energies are generally anisotropic. The interfacial energy anisotropy is usually a function of temperature with the degree of anisotropy larger at low temperatures. The type and degree of interfacial energy anisotropy affect the particle shapes or interface orientations during microstructure evolution. A number of approaches have been proposed to describe the interfacial energy anisotropy in phase-field models. One straightforward and most often employed approach to introduce interfacial energy anisotropy is to make the square-root of the gradient coefficient have the same directional dependence as the interfacial energy (according to relation [3.13]). Another approach is to add higher-order gradient energy terms[30–33]. For phase-field models with only physical order parameters, the interfacial energy anisotropy can be introduced naturally and physically through the proper introduction and coupling of gradient terms which take into account the underlying crystalline symmetry. This is particularly true for modeling the anistropic domain wall energies for ordered intermetallic phases[9,11], ferroelectric domain structures, and ferromagnetic domain structures.

Non-local interactions

The non-local contributions refer to the long-range interactions such as elastic interactions, electrostatic interactions, and magnetostatic interactions. While the bulk chemical free energy depends only on the volume fraction of each phase or domain, the energies associated with these long-range interactions are both functions of the volume fraction and morphologies of the coexisting phases or domains[34–36]. Therefore, they very often play dominant roles in the formation of microstructure patterns in solid states.

These long-range interactions can be obtained by solving the corresponding mechanical and electrostatic, and magnetostatic equilibrium equations for a given microstructure. For example, for the long-range elastic interactions,

the following mechanical equilibrium equation has to be solved under given mechanical boundary conditions:

$$\frac{\partial \sigma_{ij}}{\partial r_j} = 0 \text{ with } \sigma_{ij}(r) = \lambda_{ijkl}(r)\left\lfloor \varepsilon_{kl}(r) - \varepsilon^{o}_{kl}(c, \eta, \ldots)\right\rfloor \qquad [3.17]$$

where σ_{ij} is the local elastic stress, r_j is the jth component of the position vector, $\boldsymbol{r}$, $\lambda_{ijkl}(r)$ is the elastic stiffness tensor which varies with space, $\varepsilon_{jk}(r)$ is the total strain state at a given position in a microstructure, and ε^{o}_{kl} is the local stress-free strain or transformation strain or eigenstrain which is also a function of position through its dependence on field variables. The resulting elastic energy is a function of phase-field variables and thus the microstructure[36]. Various levels of approximations and different approaches have been proposed to solve the elasticity equation [3.17] with arbitrary distribution of eigenstrains, i.e. microstructure. For the case of homogeneous approximation and periodic boundary conditions, it was shown by Khachaturyan and Shatalov[37] that an analytical solution for the displacements, strains, and thus the strain energy could be obtained in the Fourier space. Therefore, in the case of homogeneous approximation, the elastic energy computation does not incur any significant computation. For systems with small elastic homogeneity, first-order approximations may be employed[38,39]. For large elastic inhomogeneities, first-order approximations are not sufficient and it is numerically more expensive to compute the elastic energy contributions. However, recently a number of approaches have been proposed for obtaining elastic solutions in systems with large elastic inhomogeneity[40–43]. Obtaining elasticity solutions for microstructures or domain structures in thin films is usually more difficult than for a bulk system with periodic boundary conditions. However, for simplified cases, analytical or efficient numerical methods can be designed. For example, in a film with a rigid or pre-described substrate strain, Li *et al.* showed that elastic solutions can be obtained analytically or at least semi-analytically[44,45] by assuming the same elastic constants between the film and substrate. The elastic interactions arising from surface stresses in epilayers have been obtained by a similar approach using Fourier transforms[46]. Recently, it has been shown that the interactions between precipitates and structural defects such as dislocations can be described using the same approach as discussed above[47–50]. This eigenstrain concept, widely used in micromechanics, for describing defects has been successfully applied to modeling the formation of Cottrell atmosphere and nucleation of a coherent precipitate around an edge dislocation[47,48] as well as dislocation motion[49,50].

For a system involving long-range electrostatic interactions, the local electric displacement satisfies the following electrostatic equilibrium equation with appropriate boundary conditions:

$$\frac{\partial D_i}{\partial r_i} = 4\pi\rho_{\mathrm{f}} \quad \text{or} \quad \frac{\partial E_i}{\partial r_i} = 4\pi\rho_{\mathrm{t}} = 4\pi(\rho_f - \nabla \cdot \boldsymbol{p}) \qquad [3.18]$$

where D_i and E_i are the ith components of the electric displacement field and the electric field, respectively, ρ_{f} is the local free charge and ρ_{t} is the local total charge which includes the bound charge associated with electric polarization, $\boldsymbol{p}$. Similar to elastic interactions, these equations can be analytically solved using Fourier transforms for periodic boundary conditions. For magnetostatic interactions, the structure of the equilibrium equation is almost exactly the same as the electrostatic equilibrium and can also be solved using Fourier transforms[36].

3.2.3 Evolution equations

In all phase-field models, the temporal and spatial evolution of the field variables follows a set of kinetic equations. All conserved fields, c_i, evolve with time according to the Cahn–Hilliard equation[51], or simply the diffusion equation in the case that no gradient energy is introduced for the conserved variable, whereas the non-conserved fields, η_{p} (including polarization), are governed by the Allen–Cahn equation[52], i.e.

$$\frac{\partial c_i(\mathbf{r}, t)}{\partial t} = \nabla M_{ij} \nabla \frac{\delta F}{\delta c_j(\mathbf{r}, t)} \qquad [3.19]$$

$$\frac{\partial \eta_{\mathrm{p}}(\mathbf{r}, t)}{\partial t} = -L_{pq} \frac{\delta F}{\delta \eta_q(\mathbf{r}, t)} \qquad [3.20]$$

where M_{ij} and L_{pq} are related to atom or interface mobility. F is the total free energy of a system which is a functional of all the relevant conserved and non-conserved fields given by equation [3.1]. The evolution of magnetization is assumed to follow the Landau–Lifshits–Gilbert equation (see e.g.[53]):

$$(1 + g^2) \frac{\partial \boldsymbol{m}}{\partial t} = -\gamma_{\mathrm{o}} \boldsymbol{m} \times \boldsymbol{H}_{\mathrm{eff}} - \frac{\gamma_{\mathrm{o}} g}{m_s} \boldsymbol{m} \times (\boldsymbol{m} \times \boldsymbol{H}_{\mathrm{eff}}) \qquad [3.21]$$

where $\boldsymbol{m}$ is magnetization vector, m_{s} is the saturation magnetization, γ_{o} is the gyromagnetic ratio, g is the damping constant, and $\boldsymbol{H}_{\mathrm{eff}}$ is the effective magnetic field.

In order to relate the phase-field parameters to the experimentally measurable thermodynamic and kinetic properties, one has to examine the phase-field equations in the sharp- and/or thin-interface limit. This is particularly true for phase-field models with artificial field variables for which the corresponding kinetic parameters are not directly related to the measurable physical properties. A sharp-interface analysis[4] matches the phase-field parameters at the limit of

zero interfacial thickness to experimentally measured thermodynamic and kinetic properties while a thin-interface analysis[5,6,8] allows the variation of the phase-field variable over a certain thickness for the interface. It has been shown by Karma and Rappel[5] that a phase-field simulation using the thin-interface asymptotics permits one to use a larger interface width and thus larger grid size.

3.2.4 Input parameters

Phase-field modeling requires the input parameters for determining the thermodynamics of a microstructure and its kinetics of evolution. Therefore, the parameters that are needed include those that enter or are required to determine the bulk chemical free energy as a function of order parameters as discussed in the last section. The elastic constants and the lattice parameter dependence as a function of order parameters are necessary to determine the elastic energy contribution to the thermodynamics of a microstructure. For ferroelectric and ferromagnetic crystals, the lattice parameter dependence on polarization and magnetization is described by the electrostrictive or magnetostrictive coefficients. The interfacial or domain-wall energies together with the bulk free energy density are used to determine the gradient energy coefficients. For a number of special cases such as diffusional ordering and phase separation on a fixed crystalline lattice, it is possible to calculate the gradient energy coefficients using interatomic interaction energies. Furthermore, diffusional mobility of different atomic species involved in microstructure evolution process as well as the mobility of interface or domain-wall motion determines how fast the field variables and thus the microstructures evolve as a function of time. It is also possible to use the CALPHAD approach to determine the databases for diffusional mobilities in binary or multicomponent systems using diffusivity information from experiments and first-principles calculations of relatively simpler systems. The typical input parameters needed for a phase-field simulation are summarized in Table 3.2.

3.2.5 Numerical solutions

The evolution profiles of the field variables, and thus the microstructure evolution, are obtained by numerically solving the systems of evolution equations subject to appropriate initial and boundary conditions. Most of the phase-field simulations employ the second-order finite-difference discretization in space using uniform grids and the forward Euler method for time stepping to solve the phase-field equations for simplicity. It is well known that in such an explicit scheme, the time step has to be small to keep the numerical solutions stable. Dramatic savings in computation time and improvement in numerical accuracy can be achieved by using more advanced numerical

Table 3.2 Required input parameters in a phase-field simulation

Thermodynamics kinetics	Input parameters
Bulk free energy density function	All the parameters required to construct bulk free energy density function such as transition temperatures, susceptibilities, equilibrium compositions at given temperature, heat capacities, etc.
Interfacial energies	Local free energy density and gradient energy coefficients
Elastic energy	Elastic constants, compositional dependence of lattice parameters, electrostrictive coefficients, magnetostrictive coefficients
Kinetic coefficients	Diffusional mobilities, interface, grain boundary, or domain-wall mobilities

approaches such as the semi-implicit Fourier Spectral method[54,55]. Spectral discretization is particularly convenient and attractive for systems with long-range interactions. In addition, to effectively resolve the interfacial dynamics, particularly for a system involving very different characteristic spatial scales, e.g. very large domain sizes with very few interfaces, adaptive mesh schemes are desirable. The usual practice of spatial adaptation is accomplished either through local mesh refinement and coarsening[56] or though mesh moving or transformation[57]. Very recently, based on the real-space moving mesh methods proposed in[58–60], preliminary versions of adaptive schemes are being developed that allow one to maintain the applicability of the spectral codes[61]. This is achieved by working with both a computational space with uniform grids and a real space with adaptive grids. Based on test simulations on isolated domains in a matrix, it is possible to achieve an order of magnitude improvement in efficiency due to the smaller number of grid points required to achieve the same accuracy despite the significant overhead involved in evolving the moving mesh.

3.3 Advantages and disadvantages

The phase-field method offers a number of advantages. First of all, with the phase-field approach, one is able to model the evolution of arbitrary morphologies and complex microstructures without explicitly tracking the positions of interfaces. This is particularly true for modeling three-dimensional microstructures for which conventional front-tracking approaches would have been very difficult. Secondly, it can be applied to essentially all types of microstructure problems related to vastly different materials processes by appropriately choosing either physical or artificial field variables. For example,

it has been successfully applied to solidification, solid-state phase transformations, coarsening and growth, and many others (see review[1–3]). Third, the phase-field model can describe different processes such as phase transformations (driven by bulk free energy reduction) and particle coarsening (driven by interfacial energy reduction) within the same formulation. It is relatively straightforward to incorporate the effect of coherency and applied stresses as well as electrical and magnetic fields. Finally, it is possible to link phase-field models with thermodynamic and kinetic databases for obtaining the materials parameters[12,21–27] or to derive the free energy models from microscopic models[62–64]. One of the main disadvantages of the phase-field approach is the fact that the method is still computationally very intensive, particularly for three-dimensional systems. Therefore, it is essential that efficient and accurate numerical algorithms should be developed and implemented. Furthermore, it relies on more fundamental calculations such as first-principles calculations or experimental data for the input parameters. Finally, the physical size that a phase-field simulation can handle is, in many cases, limited by the usually small physical width of real interfaces in microstructures as compared to phase and domain sizes. Although the thin-interface analysis is able to significantly increase the artificial interfacial width that one can use in a phase-field simulation, and thus the physical size of a simulation, there is still a limit on the width that one can use in phase-field simulations as discussed in the last section. Another strategy is to employ the adaptive mesh approach to increase the length scale for a phase-field simulation.

3.4 Recent developments and future opportunities

Phase-field models have been applied to many different materials processes. Since the review articles were published in 2002[2], there have been a number of new developments which are worth mentioning. One of the new directions is the increasing effort in developing multiscale models to predict microstructure evolution starting from first principles. For example, it was recently shown that it is possible to obtain all the necessary thermodynamic information for the input to a phase-field model from first-principles calculations combined with cluster expansions; it includes the bulk free energies of matrix and precipitate phases, the interfacial energy and its anisotropy, and the lattice mismatch[14]. It has also been shown that it is feasible to obtain the fundamental properties of solid–liquid interfaces such as interfacial energies and mobilities as well as their anisotropies[65] for input to phase-field simulations of solidifications[66]. Recently, Shen and Wang showed that the g-surface can be obtained from first-principles calculations and can then be used in phase-field simulations of dislocation motion[67]. For modeling the evolution of relatively complex microstructures, such information-

passing from one level to another seems to be the most realistic approach for multiscale modeling.

For applications to more complicated systems such as multicomponent alloys, it is not possible to obtain all the necessary structural, thermodynamic, and kinetic parameters directly from first-principles atomistic calculations. As a result, phase-field models are increasingly relying on existing or future thermodynamic, kinetic, and crystallographic databases obtained from empirical modeling. As outlined above, it is possible to directly construct the free energy function of a phase-field model from existing databases using the CALPHAD method[12, 21–27]. The compositional dependence of atomic mobilities from databases can also be incorporated (see for example[25,26]). However, in order to take into account the effect of elastic energy in solid-state processes, additional databases, such as the crystallographic lattice parameters and elastic constants, have to be constructed. An attempt to construct lattice parameter databases is already underway[68]. With independently assessed reliable databases, it will be possible to predict the microstructure evolution in complex multicomponent alloys using the phase-field method.

Another new direction is the phase-field crystal model proposed by Elder *et al.*[69,70]. It describes the crystal structures and microstructures on atomic or subatomic length and diffusive time scales. The main advantage of this approach is its ability to model both plastic and elastic deformations with a time scale larger than atomistic molecular dynamics simulations. It has been employed to model epitaxial growth, material hardness, grain growth, reconstructive phase transitions, and crack propagation. The main disadvantage will be the small length scale and the difficulty in constructing the local free energy density for a real material.

There is also increasing interest in exploring problems involving non-periodic low-dimensional systems such as thin films and surfaces[45,46,71–83] and interactions between phase and defect microstructures such as dislocations[47,84–86], etc.

Three-dimensional phase-field simulations are very computationally intensive, and thus they require innovative numerical algorithms. Furthermore, many of the practical microstructures have compositional or structural domains which are much larger than the width of the interfaces. Despite the effort to increase the artificial interfacial width that one can employ in a phase-field model, for example, using the thin-interface analysis, in many cases, numerical methods using uniform grid sizes will not be sufficient to perform three-dimensional phase-field simulations with the desirable system size. Therefore, it is important that efficient adaptive numerical algorithms be developed. However, it is particularly challenging to develop adaptive algorithms for systems in which the interfaces are abundant in a microstructure and long-range interactions such as elastic interactions are involved. A recent model combining a moving mesh method with the spectral method seems to be

promising for systems involving long-range interactions, but the amount of interfaces is low[61].

There have been a number of existing studies on nucleation and growth using the phase-field method[63,87–90]. Clearly, significant effort is still required to establish a robust, physical, and quantitative approach to induce both homogeneous nucleation in the bulk and heterogeneous nucleation around defects within the phase-field approach.

It is possible to combine the phase-field modeling of microstructure evolution and effective property calculation of a microstructure to obtain the temporal evolution of properties[43] or to simply use the microstructure evolution obtained from a phase-field simulation in a constitutive model to predict the materials behavior, e.g. to use a strengthening model to predict the effect of microstructure on alloy strengthening. It is expected that there will be increasing effort linking microstructure evolution to property prediction.

3.5 Acknowledgments

The author is grateful for the financial supports from NSF under the grant numbers DMR-0122638, DMR-0205232 and DMR-0507146.

3.6 References

1. A. Karma, Phase field methods, in *Encyclopedia of Materials Science and Technology*, eds K.H.J. Buschow *et al*., Oxford: Elsevier, 2001, 6873–6886.
2. L. Q. Chen, Phase-field models for microstructure evolution, *Annu. Rev. Mater. Res*. **32**, 113–140, (2002).
3. W. J. Boettinger, J. A. Warren, C. Beckermann and A. Karma, Phase-field simulation of solidification, *Annu. Rev. Mater. Res*. **32**, 163–194 (2002).
4. G. Caginalp and W. Xie, Phase-field and sharp-interface alloy models, *Phys. Rev. E* **48**(3), 1897–1909 (1993).
5. A. Karma and W. J. Rappel, Phase-field method for computationally efficient modeling of solidification with arbitrary interface kinetics, *Phys. Rev. E.* **53**(4), R3107–3020 (1996).
6. S. G. Kim, W. T. Kim, and T. Suzuki, Phase-field model for binary alloys, *Phys. Rev. E* **60**(6), 7186–7197 (1999).
7. K. R. Elder, M. Grant, N. Provatas and J. M. Kosterlitz, Sharp interface limits of phase-field models – art. no. 021604. *Phys. Rev. E* **64**(2), Art. No. 021604 (2001).
8. A. Karma, *Phase-field formulation for quantitative modeling of alloy solidification. Phys. Rev. Letts*, **87**(11), Art. No. 11570 (2001).
9. R. J. Braun, J. W. Cahn, G. B. Mcfadden and A. A. Wheeler, anisotropy of interfaces in an ordered alloy: a multiple-order-parameter model, *Philos. Trans. R. Soc. London, Ser.* A **355**(1730), 1787–1833 (1997).
10. D. Y. Li and L. Q. Chen, Shape evolution and splitting of coherent particles under applied stresses, *Acta Mater.* **47**(1), 247–257 (1998).
11. Y. Wang, D. Banerjee, C. C. Su and A. G. Khachaturyan, Field kinetic model and computer simulation of precipitation of Ll(2) ordered intermetallics from fcc solid solution, *Acta Mater.* **46**(9), 2983–3001 (1998).

12. J. Z. Zhu, Z. K. Liu, V. Vaithyanathan and L. Q. Chen, Linking phase-field model to calphad: application to precipitate shape evolution in Ni-base alloys, *Scripta Mater.* **46**(5), 401–406 (2002).
13. J. W. Cahn and J. E. Hilliard, Free energy of a nonuniform system. I. Interfacial free energy, *J. Chem. Phys.*, **28**, 258–267 (1958).
14. V. Vaithyanathan, C. Wolverton and L. Q. Chen, Multiscale modeling of precipitate microstructure evolution, *Phys. Rev. Lett.* **88**(12), Art. No. 125503 (2002).
15. E. K. H. Salje, *Phase transitions in ferroelastic and co-elastic crystals: an introduction for mineralogists, material scientists, and physicists*. Cambridge Topics in Mineral Physics and Chemistry, Cambridge, New York: Cambridge University Press, 1990.
16. L. Kaufman and H. Bernstein, *Computer Calculation of Phase Diagram*, New York: Academic Press, 1970.
17. N. Saunders and A. P. Miodownik, *Calphad (Calculation of Phase Diagrams): A Comprehensive Guide*, Oxford, New York: Pergamon, 1998.
18. L. Q. Chen, C. Wolverton, V. Vaithyanathan and Z. K. Liu, Modeling solid-state phase transformations and microstructure evolution, *MRS Bull.* **26**(3), 197–202 (2001).
19. C. Wolverton, X. Y. Yan, R. Vijayaraghavan and V. Ozolins, Incorporating first-principles energetics in computational thermodynamics approaches, *Acta Mater.*, **50**(9), 2187–2197 (2002).
20. Z.-K. Liu, L.-Q. Chen, P. Raghavan, Q. Du, J. O. Sofo, S. A. Langer and C. Wolverton, An integrated framework for multi-scale materials simulation and design, *J. Comput. Aided Mater. Des.*, **11**, 183–199 (2004).
21. B. Bottger, U. Grafe, D. Ma and S. G. Fries, Simulation of microsegregation and microstructural evolution in directionally solidified superalloys, *Mater. Sci. Technol.* **16**(11–12), 1425–1428 (2000).
22. H. Kobayashi, M. Ode, S. G. Kim, W. T. Kim and T. Suzuki, Phase-field model for solidification of ternary alloys coupled with thermodynamic database, *Scripta Mater.* **48**(6), 689–694 (2003).
23. R. S. Qin and E. R. Wallach, A phase-field model coupled with a thermodynamic database, *Acta Mater.* **51**(20), 6199–6210 (2003).
24. I. Loginova, J. Agren and G. Arnberg, On the formation of widmanstdtten ferrite in binary Fe-C – phase-field approach, *Acta Mater.*, **52**(13), 4055–4063 (2004).
25. K. Wu, Y. A. Chang and Y. Wang, Simulating interdiffusion microstructures in Ni-Al-Cr diffusion couples: a phase field approach coupled with Calphad database, *Scripta Mater.* **50**(8), 1145–1150 (2004).
26. J. Z. Zhu, T. Wang, S. H. Zhou, Z. K. Liu and L. Q. Chen, Quantitative interface models for simulating microstructure evolution, *Acta Mater.* **52**(4), 833–840 (2004).
27. T. Koyama and H. Onodera, Computer simulation of phase decomposition in Fe-Cu-Mn-Ni quaternary alloy based on the phase-field method, *Mater. Trans.* **46**(6), 1187–1192 (2005).
28. I. Steinbach, F. Pezzolla, B. Nestler, M. Seesselberg, R. Prieler, G. J. Schmitz and J. L. L. Rezende, A phase field concept for multiphase systems, *Physica D.* **94**(3), 135–147 (1996).
29. P. R. Cha, D. H. Yeon and J. K. Yoon, A phase field model for isothermal solidification of multicomponent alloys, *Acta Mater.* **49**(16), 3295–3307 (2001).
30. T. A. Abinandanan and F. Haider, An extended Cahn-Hilliard model for interfaces with cubic anisotropy, *Philos. Mag. A* **81**(10), 2457–2479 (2001).

31. J. S. Langer, Models of pattern formation in first-order phase transitions, in *Directions in Condensed Matter Physics*, eds G. Grinstein and G. Mazenko, Singapore. World Scientific, 165–186 (1986).
32. J. E. Taylor and J. W. Cahn, Diffuse interfaces with sharp corners and facets: phase field models with strongly anisotropic surfaces, *Physica D*, **112**(3–4), 381–411 (1998).
33. W. Dreyer and W. H. Muller, toward quantitative modeling of morphology changes in solids with phase field theories: atomistic arguments for the determination of higher gradient coefficients, *Bioceramics* **15**, 901–914 (2003).
34. C. M. Bishop, R. E. Garcia and W. C. Carter, Effect of charge separation on the stability of large wavelength fluctuations during spinodal decomposition, *Acta Mater.* **51**(6), 1517–1524 (2003).
35. L. Q. Chen, Y. Z. Wang and A. G. Khachaturyan, Transformation-induced elastic strain effect on the precipitation kinetics of ordered intermetallics, *Philos. Mag. Letts*, **64**(5), 241–251 (1991).
36. A. G. Khachaturyan, *Theory of Structural Transformations in Solids*, New York: John Wiley & Sons, (1983).
37. A. G. Khachaturyan and G. A. Shatalov, Elastic-interaction potential of defects in a crystal, *Sov. Phys. Solid State*, **11**, 118–123 (1969).
38. A. Onuki, Long-range interaction through elastic fields in phase separating solids, *J. Phys. Soc. Jpn.* **58**, 3069–3072 (1989).
39. C. Sagui, D. Orlikowski, A. Somoza and C. Roland, Three-dimensional simulations of Ostwald ripening with elastic effects, *Phys. Rev. E*. **58**, 569–577 (1998).
40. S. Y. Hu and L. Q. Chen, A phase-field model for evolving microstructures with strong elastic inhomogeneity, *Acta Mater.* **49**(11), 1879–1890 (2001).
41. P. H. Leo, J. S. Lowengrub and H. J. Hou, A diffuse interface model for microstructural evolution in elastically stressed solids, *Acta Mater.* **61**, 2113–2130 (1998).
42. Y. U. Wang, Y. M. M. Jin and A. G. Khachaturyan, Phase field microelasticity theory and modeling of elastically and structurally inhomogeneous solid, *J. Appl. Phys.* **92**(3), 1351–1360 (2002).
43. J. Z. Zhu, L. Q. Chen and J. Shen, Computer simulation of morphological evolution during phase separation and coarsening with strong inhomogeneous elasticity, *Modell. Simul. Mater. Sci. Eng.* **9**, 499–511 (2001).
44. Y. L. Li, S. Y. Hu, Z. K. Liu and L. Q. Chen, Phase-field model of domain structures in ferroelectric thin films, *Appl. Phys. Letts*, **78**(24), 3878–3880 (2001).
45. Y. L. Li, S. Y. Hu, Z. K. Liu and L. Q. Chen, Effect of substrate constraint on the stability and evolution of ferroelectric domain structures in thin films, *Acta Mater.* **50**(2), 395–411 (2002).
46. Z. Suo and W. Lu, Composition modulation and nanophase separation in a binary epilayer, *J. Mech. Phys. Solids* **48**(2), 211–232 (2000).
47. S. Y. Hu and L. Q. Chen, Solute segregation and coherent nucleation and growth near a dislocation – a phase-field model integrating defect and phase microstructures, *Acta Mater.* **49**(3), 463–472 (2001).
48. S. Y. Hu and L. Q. Chen, Diffuse-interface modeling of composition evolution in the presence of structural defects, *Comput. Mater. Sci.* **23**(1–4), 270–282 (2002).
49. Y. U. Wang, Y. M. Jin, A. M. Cuitino and A. G. Khachaturyan, Phase field microelasticity theory and modeling of multiple dislocation dynamics, *Appl. Phys. Letts.* **78**(16), 2324–2326 (2001).
50. Y. U. Wang, Y. M. Jin, A. M. Cuitino and A. G. Khachaturyan, Application of phase field microelasticity theory of phase transformations to dislocation dynamics: model

and three-dimensional simulations in a single crystal, *Philos. Mag. Letts.* **81**(6), 385–393 (2001).
51. J. W. Cahn, On spinodal decomposition, *Acta Metall.* **9**, 795–801 (1961).
52. S. M. Allen and J. W. Cahn, A microscopic theory of domain wall motion and its experimental verification in Fe-Al alloy domain growth kinetics, *Journal de Physique*. **C7**, C7–51 (1977).
53. J. Fidler and T. Schrefl, Micromagnetic modelling – the current state of the art, *J. Phys. D* **33**(15), R135–R156 (2000).
54. L. Q. Chen and J. Shen, Applications of semi-implicit fourier-spectral method to phase field equations, *Comput. Phys. Commun.* **108**(2-3), 147–158 (1998).
55. J. Z. Zhu, L. Q. Chen, J. Shen and V. Tikare, Coarsening kinetics from a variable-mobility Cahn-Hilliard equation: application of a semi-implicit fourier spectral method *Phys. Rev. E.* **60**(4), 3564–3572 (1999).
56. M. J. Berger and J. Oliger, Adaptive mesh refinement for hyperbolic partial differential equations, *J. Comput. Phys.* **53**, 484–512 (1984).
57. D. S. McRae and K. R. Laflin, Dynamic grid adaption and grid quality, *Handbook of Grid Generation*, eds. J.F. Thompson, B.K. Soni and N.P. Wetherill, Boca Raton, FL: CRC Press, 1999.
58. J. U. Brackbill and J. S. Saltzman, Adaptive zoning for singular problems in two-dimensions, *J. Comput. Phys.* **46**, 342–368 (1982).
59. W. Z. Huang and R. D. Russell, Moving mesh strategy based on a gradient flow equation for two-dimensional problems, *SIAM J. Sci. Comput.* **20**, 998–1015 (1996).
60. H. D. Ceniceros and T. Y. Hou, An efficient dynamically adaptive mesh for potentially singular solutions, *J. Comput. Phys.* **72**, 609–639 (2000).
61. W. M. Feng, P. Yu, S. Y. Hu, Z.-K. Liu, Q. Du and L. Q. Chen, Spectral adaptive moving mesh method for phase-field equations, *J. Comput. Phys.* **220**, 498–510 (2006).
62. M. Ohno and T. Mohri, Disorder-L1(0) transition investigated by phase field method with cvm local free energy, *Mater. Trans*. **42**(10), 2033–2041 (2001).
63. R. Poduri and L. Q. Chen, Non-classical nucleation theory of ordered intermetallic precipitates – application to the Al-Li alloy, *Acta Mater.* **44**(10), 4253–4259 (1996).
64. V. G. Vaks, Microscopic derivation of Ginzburg-Landau-type functionals for alloys and their application to studies of antiphase and interphase boundaries, *JETP Lett*, **76**(2), 93–98 (2002).
65. J. J. Hoyt, M. Asta, and A. Karma, Method for computing the anisotropy of the solid-liquid interfacial free energy, *Phys. Rev. Letts*, **86**(24), 5530–5533 (2001).
66. J. Bragard, A. Karma, Y. H. Lee and M. Plapp, Linking phase-field and atomistic simulations to model dendritic solidification in highly undercooled melts, *Interface Sci.* **10**(2–3), 121–136 (2002).
67. C. Shen and Y. Wang, Incorporation of gamma-surface to phase field model of dislocations: simulating dislocation dissociation in fcc crystals, *Acta Mater.* **52**(3), 683–691 (2004).
68. T. Wang, J. Z. Zhu, R. A. Mackay, L. Q. Chen and Z. K. Liu, Modeling of lattice parameter in the Ni-Al system, *Metall. Mater. Trans. A*. **35A**(8), 2313–2321 (2004).
69. K. R. Elder and M. Grant, Modeling elastic and plastic deformations in nonequilibrium processing using phase field crystals, *Phys. Rev. E*, **70**(5), Art. No. 051605 (2004).
70. K. R. Elder, M. Katakowski, M. Haataja, and M. Grant, Modeling elasticity in crystal growth, *Phys. Rev. Letts*. **88**(24), Art. No. 245701 (2002).

71. A. Karma and M. Plapp, Spiral surface growth without desorption, *Phys. Rev. Letts.* **81**(20), 4444–4447 (1998).
72. M. Mahadevan and R. M. Bradley, Phase field model of surface electromigration in single crystal metal thin films, *Physica D* **126**(3–4), 201–213 (1999).
73. S. M. Wise and W. C. Johnson, Numerical simulations of pattern-directed phase decomposition in a stressed, binary thin film, *J. Appl. Phys.* **94**(2), 889–898 (2003).
74. W. Lu and D. Kim, Dynamics of nanoscale self-assembly of ternary epilayers. *Microelectron. Eng.* **75**(1), 78–84 (2004).
75. D. C. Kim and W. Lu, Self-organized nanostructures in multi-phase epilayers, *Nanotechnology*, **15**(5), 667–674 (2004).
76. L. Wei and D. Kim, Simulation on nanoscale self-assembly of ternary-epilayers, *Comput. Mater. Sci.* **32**(1), 20–30 (2005).
77. A. Artemev, J. Slutsker and A. L. Roytburd, Phase field modeling of self-assembling nanostructures in constrained films, *Acta Mater.* **53**(12), 3425–3432 (2005).
78. H. C. Yu and W. Lu, Ordering of nanovoids in an anisotropic solid driven by surface misfit, *J. Comput. Theor. Nanosci.* **2**(2), 256–262 (2005).
79. Y. U. Wang, Y. M. M. Jin and A. G. Khachaturyan, Mesoscale modelling of mobile crystal defects – dislocations, cracks and surface roughening: phase field microelasticity approach, *Philos. Mag.* **85**(2–3), 261–277 (2005).
80. S. Choudhury, Y. L. Li and L. Q. Chen, A phase diagram for epitaxial PbZr1-xTiXo.3 thin films at the bulk morphotropic boundary composition, *J. Amer. Ceramic Soc.* **88**(6), 1669–1672 (2005).
81. D. H. Yeon, P. R. Cha and M. Grant, Phase field model of stress-induced surface instabilities: surface diffusion, *Acta Mater.* **54**(6), 1623–1630 (2006).
82. H. Emmerich and C. Eck, Microstructure morphology transitions at mesoscopic epitaxial surfaces, *Continuum Mech. Thermodyn.* **17**(5), 373–386 (2006).
83. T. Takaki, T. Hasebe and Y. Tomita, Two-dimensional phase-field simulation of self-assembled quantum dot formation, *J. Cryst. Growth* **287**(2), 495–499 (2006).
84. S. Y. Hu, J. Choi, Y. L. Li and L. Q. Chen, Dynamic drag of solute atmosphere on moving edge dislocations – phase-field simulation, *J. Appl. Phys.* **96**(1), 229–236 (2004).
85. S. Y. Hu and L. Q. Chen, Spinodal decomposition in a film with periodically distributed interfacial dislocations, *Acta Mater.* **52**(10), 3069–3074 (2004).
86. S. Y. Hu, Y. L. Li and L. Q. Chen, Effect of interfacial dislocations on ferroelectric phase stability and domain morphology in a thin film – a phase-field model, *J. Appl. Phys*. **94**(4), 2542–2547 (2003).
87. A. Roy, J. M. Rickman, J. D. Gunton and K. R. Elder, Simulation study of nucleation in a phase-field model with nonlocal interactions, *Physi. Rev. E* **57**(3), 2610–2617 (1998).
88. J. P. Simmons, C. Shen and Y. Wang, Phase field modeling of simultaneous nucleation and growth by explicitly incorporating nucleation events, *Scripta Material.* **43**(10), 935–942 (2000).
89. L. Granasy, T. Borzsonyi and T. Pusztai, Crystal nucleation and growth in binary phase-field theory, *J. Cryst Growth* **237**, 1813–1817 (2002).
90. M. Castro, Phase-field approach to heterogeneous nucleation, *Phys. Rev. B* **67**(3), Art. No.035412 (2003).

4

Mesoscale modelling of grain growth and microstructure in polycrystalline materials

D. MOLDOVAN, Louisiana State University, USA;
D. WOLF, Argonne National Laboratory, USA

4.1 Introduction

The key challenge in multiscale computer simulation of polycrystalline materials consists in the development of the concepts and algorithms that link the different length and time scales controlling the evolution of complex systems. As a minimum, there are three levels that must be considered. First, the electronic and atomic levels, the 'microscale', which encompasses the nature of chemical bonding and all relevant atomic-level processes and mechanisms. Second, the level broadly defined as the 'mesoscale' focuses on the critical role played by material microstructure including, for example, the grain-boundary microstructure in polycrystals, porosity, precipitates, any dislocation substructures and cracks. Finally, the results of the mesoscale simulations provide the basis for a continuum description of the material, and thus enable macroscale simulations, for example, in the form of a finite-element or phase-field approach.

Here we demonstrate for the simple model case of grain growth how a rigorous computational and theoretical framework can be formulated which permits physical insights gained from atomic-level simulations to be transferred into the mesoscale and, ultimately, linked to the continuum level. This transfer requires quantification of the results of the atomic-level simulations by formulation of a theoretical model. This then enables examination of the statistical mechanics of the process in a realistic system, i.e. without the well-known time and length-scale restrictions inherent to the atomic-level simulations. These mesoscale simulations then provide the foundation for the formulation of a continuum theoretical framework, with input parameters extracted from the mesoscale simulations.

Grain growth is the process by which the average grain size in a polycrystal increases in time. The significance of this process comes from the profound influences of the average grain size and the grain-size distribution on a wide range of properties of polycrystalline materials. Grain size is a key parameter in sintered ceramics as well as metal and alloy microstructures and usually

has to be controlled during thermo-mechanical processing in order to ensure optimization of mechanical properties. The conventional picture of grain growth, derived from extensive studies of coarse-grained polycrystals, is that the process is driven by the reduction of the total area of the grain boundaries (GBs) in the material (for recent reviews, see Atkinson, 1988; Humphreys and Hatherly, 1995). The underlying mechanism involves curvature-driven GB migration, i.e. the motion of the GBs towards the center of their curvature. This curvature arises from the fact that the Herring relation (Herring, 1951) for the dihedral angles between the GBs joined at the triple junctions cannot be satisfied unless the GBs are curved. For a given grain diameter, d, this curvature is of the order of $1/d$. Nanocrystalline materials are therefore particularly unstable against grain growth, as evidenced by the observation of grain coarsening even at relatively low temperatures (Gleiter, 1989). A better understanding of the thermodynamic stability of their grain microstructure is obviously essential if they are to be useful for any applications at even moderate temperatures.

Quantitative experimental investigations of grain growth, particularly non-destructive ones, are difficult not only for a nm grain size but even for coarse-grained materials. Also, the classical theoretical treatments of grain growth have encountered considerable difficulties in fully incorporating the dynamical topological features of evolving microstructures during grain coarsening (Burke and Turnbull, 1952; Feltham, 1957; Smith, 1964; Hillert, 1965; Louat, 1974). This has led to extensive computer simulations of curvature-driven grain growth mostly in two-dimensional (2D) model systems. Largely inspired by the investigation of isotropic systems, such as soap froth (Smith, 1964), these simulations have used a variety of approaches, such as GB-dynamics models (Kermode and Weaire, 1990; Weaire and Lei, 1990), vertex models (Fullman, 1952; Thompson *et al.*, 1987; Kawasaki *et al.*, 1989; Weygand *et al.*, 1998), Potts models (Anderson *et al.*, 1984, Grest *et al.*, 1985; Anderson and Grest, 1989), Voronoi tessellation (Kurtz and Carpay, 1980; Kumar *et al.*, 1992), mean-field models (Rivier, 1983; Fradkov *et al.*, 1985; Beenakker, 1988) and continuum diffuse-interface models describing the energy stored in the GBs by a Ginzburg–Landau equation (Chen and Yang, 1994; Chen, L.-Q., 1995; Fan and Chen, 1995).

All these simulation models share the common feature of being mesoscopic in nature, i.e. the objects treated dynamically are the GBs and grain junctions rather than the atoms of which they are composed. Mesoscale simulations therefore require GB properties as input parameters, such as the energy and mobility of each GB for a given misorientation between neighboring grains and information on the energy and mobility of the triple junctions (Galina *et al.*, 1987; Gottstein *et al.*, 1999). While they are therefore computationally highly efficient, mesoscale simulations are limited by the fact that they require *at the outset* a full knowledge of the physical laws governing the

evolution of the microstructure. They therefore provide neither a test of the actual validity of the assumed grain-growth mechanism nor insights on the possible activation of other mechanisms that may have been overlooked. For example, the triple junctions are usually assumed to be much more mobile than the GBs, ensuring that at any instant the Herring relation (Herring, 1951) is satisfied at every triple junction. While recent experimental evidence has cast doubt on the general validity of this assumption (Galina *et al.*, 1987), within the framework of mesoscale simulations alone this important issue cannot be addressed.

By contrast with mesoscale simulations, the only physical input required for atomic-level simulations is the interatomic force law specifying how the atoms interact with each other. Given the dramatic increases in computer power during recent years, atomistic simulations of polycrystalline materials, albeit with a nm grain size, have now become feasible for the first time (Phillpot *et al.*, 1995; D'Agostino and Van Swygenhoven, 1996, Keblinski *et al.*, 1997, 1998; Schiotz *et al.*, 1998, 1999; Van Swygenhoven *et al.*, 1998, 2000; Haslam *et al.*, 2001, 2002, 2003, 2004). In a recent molecular-dynamics (MD) simulation study of grain growth in a nanocrystalline-palladium model microstructure, Haslam *et al.* (2001) showed that in addition to the conventional growth mechanism by curvature-driven GB-migration, grain rotations play an equally important role, by eliminating the common GB between neighboring grains while leading to grain coalescence. These insights can be incorporated into mesoscale simulations in which, instead of the atoms, the objects that evolve in space and time are the GBs, grain junctions and grain orientations, with a timescale controlled by that associated with GB migration, and the grain rotations and a length scale given by the grain size. These mesoscale simulations, incorporating physical insight and input GB parameters obtained by MD simulation, enable the investigation of the topology and long-time grain-growth behavior in a physically realistic manner.

Here we demonstrate for the simple model case of grain growth how a rigorous hierarchical computational and theoretical framework can be formulated that permits physical insights gained from atomic-level simulations to be transferred into the mesoscale, and ultimately be linked to the continuum level. The atomic level–mesoscale linkage requires quantification for transfer to the mesoscale of the results of the atomic-level simulations by formulation of a theoretical model. The mesoscale simulations then enable examination of the statistical mechanics of the process in a realistic system, i.e. without the well-known time and length-scale restrictions inherent to the atomic-level simulations. Input parameters extracted from these simulations subsequently provide the basis for the formulation of a continuum theoretical framework.

The chapter is organized as follows. In addition to describing the MD simulation approach, Section 4.2 summarizes the most important results of

the MD simulations of grain growth, including the underlying atomic-level mechanisms thus elucidated. The section also reviews how MD simulation has been used to determine GB energies and mobilities. Section 4.3 offers a detailed description of our mesoscale simulation approach that, in Section 4.4, is validated by a comparison against the evolution of the same system studied in the MD simulations. We then determine the topology and kinetics of grain growth in a 2D polycrystal containing thousands of grains and in the presence of both GB migration and grain rotations (Section 4.5). In this section we also illustrate how some of these results can be captured analytically within an extended von Neumann–Mullins theory of grain growth. Finally, the discussion in Section 4.6 offers a perspective on how the atomistic and mesoscale simulations complement each other in elucidating grain growth, and how the effects of external stress might be incorporated within a general, hierarchical multiscale simulation framework.

4.2 Molecular dynamics simulation of grain growth

As described in detail in the paper by Haslam *et al.* (2001), a textured, columnar microstructure is ideal for this study because, while providing a fully 3D treatment of the underlying GB physics, its evolution during grain growth is only a 2D problem (in the x–y plane) and hence computationally highly efficient. Only a few lattice planes (their number being determined by the truncation radius of the interatomic potential) need to be considered in the periodically repeated texture (z) direction, and the evolution of such a microstructure is easily visualized and analyzed. Furthermore, such a microstructure can be considered as a physical realization of the 2D microstructures on which virtually all of the theory of grain growth has been developed and for which most mesoscopic simulations have been performed, including those described in this chapter (see Sections 4.3 and 4.5). Experimentally, such 2D microstructures can, at least in principle, be realized in the form of unsupported, textured thin films.

4.2.1 Simulation approach

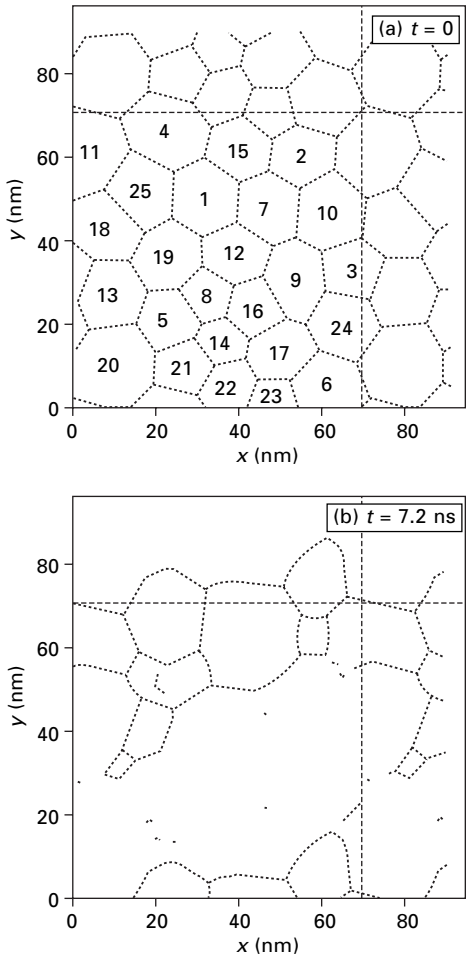
As described in detail in the paper by Haslam *et al.* (2001), an initial microstructure containing 25 periodically repeated grains with an average grain diameter of $d = 14$ nm was chosen for this study. An embedded atom method (EAM) potential parameterized for Pd (Foiles and Adams, 1986) was used throughout. This particular choice was motivated by the earlier extensive MD simulations of nanocrystalline Pd by the Argonne group (Keblinski *et al.*, 1999(a, b); Yamakov *et al.*, 2000). This potential was fitted to give the correct zero-temperature lattice parameter ($a_0 = 3.89$ Å), cohesive

energy, elastic constants and vacancy-formation energy. Based on free-energy simulations, the melting point for this potential was estimated to be about 1400 K (Foiles and Adams, 1986), which is considerably lower than the experimental value of 1825 K. However, in the previous simulations of GB diffusion in Pd (Keblinski *et al.*, 1999a), no signs of GB-induced melting at 1400 K were detected for even rather long simulation times, suggesting that the melting point may be significantly higher, probably closer to $T_m \approx 1500$ K.

The initial microstructure shown in Fig. 4.1(a) was generated using a 2D Voronoi construction (Voronoi, 1908; Chen D., 1995) to define the shapes and the relative sizes of polygons (Weaire and Kermode, 1984). A 3D periodic polycrystal is then formed by filling each of these polygons with a 3D perfect-crystal fcc lattice with a chosen orientation (for details, see Haslam *et al.*, 2001). As the common texture (or tilt) axis in the z direction Haslam *et al.*, chose the <001> direction; all GBs in the system are therefore <001> tilt boundaries. This direction was chosen because the energies of <001> tilt GBs do not exhibit cusps for certain 'special' misorientations (Wolf, 1990; Wolf and Merkle, 1992) and thus avoid any issues related to the distinct behaviors of 'special' and 'random' high-angle GBs during grain growth. To introduce GBs into the system, each grain must be assigned a tilt angle about <001> that differs from those of its neighbors. In principle, these angles could be chosen randomly; this would yield a microstructure inevitably containing some GBs with very small misorientation angles, i.e. with a spacing between the GB dislocations that can become comparable to, or even larger than, the length of the GB segment itself. So as to avoid this problem, Haslam *et al.* limited themselves to larger misorientation angles, by generating only misorientations such that each GB segment contains at least three or four dislocation cores; this has the additional advantage that the GBs are more clearly visible in our simulations (as continuous lines of miscoordinated atoms; see Fig. 4.1(a)). In practice, the random set of grain orientations was refined using a Monte-Carlo procedure to avoid such low-angle GBs; for the resulting set of grain orientations, the minimum misorientation angle between any two grains was 14.9° (Haslam *et al.*, 2001). It is important to note that, since the GB plane is fixed by the Voronoi construction, the tilt GBs thus obtained are generally asymmetric.

4.2.2 Grain-growth mechanisms

The comparison of Fig. 4.1(b) with 4.1(a), showing the positions of only the miscoordinated atoms in one of the 6(001) planes in the simulation cell at the end and beginning of the simulation, illustrates the overall evolution of this microstructure at T = 1400 K over a period of 7.2 ns, corresponding to approximately 1.4 million MD time steps. We note that most of the



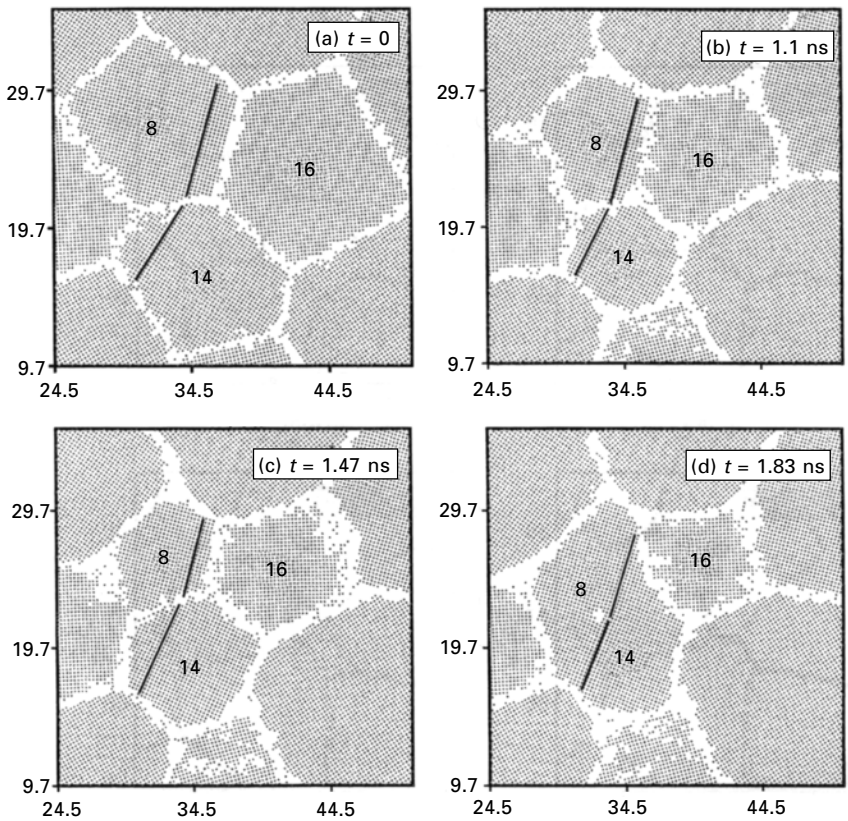
4.1 (a) The 25 grains in the initial microstructure with average grain diameter of 14 nm are clearly delineated by the miscoordinated atoms. (b) After simulation for 7.2 ns (≈ 1.4 million time steps) at T = 1400K, considerable grain growth has taken place and only 11 grains remain. The simulation-cell borders are delineated by the dashed lines; however, the plots are extended into the periodically repeated regions to show more clearly the initial size and shape of each grain.

miscoordinated atoms in the grain interiors are associated with dislocation cores; the remainder indicates the presence of vacancies and interstitials.

As seen from Figs 4.1(a) and (b), during the 7.2 ns of the MD simulation the system evolved from an initial microstructure containing 25 grains with relatively narrow distributions in the grain sizes and shapes, to one consisting of only 11 grains with far broader distributions in both. The final microstructure

is clearly dominated by one highly elongated grain that is far larger than any of the grains in the initial microstructure; this dominant grain has the same orientation as the initial grain number 9 (see Fig. 4.1(a)).

As described in detail in the paper of Haslam *et al.* (2001), this grain is the result of grain-coalescence events following coupled grain rotation and GB migration involving, among others, grains 5, 8, 9, 14 and 16. The four snapshots in Fig. 4.2 illustrate the time evolution of this process from the viewpoint of the perfectly coordinated atoms in the grain interiors, showing



4.2 Atomic-level mechanism of the first rotation-coalescence event from the viewpoint of the atoms in the grain interiors (i.e. all the miscoordinated atoms, located in the white regions, were omitted). The solid lines indicate <110> directions. (a) Initial configuration of grains 8, 14 and 16, with $\theta_{8\text{–}14} \approx 18°$; (b) after $t = 1.11$ ns, $\theta_{8\text{–}14} \approx 11°$; (c) after $t = 1.47$ ns, $\theta_{8\text{–}14} \approx 9°$; (d) after $t = 1.83$ *ns*, $\theta_{8\text{–}14} \approx 4°$. Notice the decreasing areas of grains 8 and 16, and the initially decreasing and later increasing area of grain 14; these indicate that the rotation-coalescence and GB-migration mechanisms of grain growth are intimately coupled.

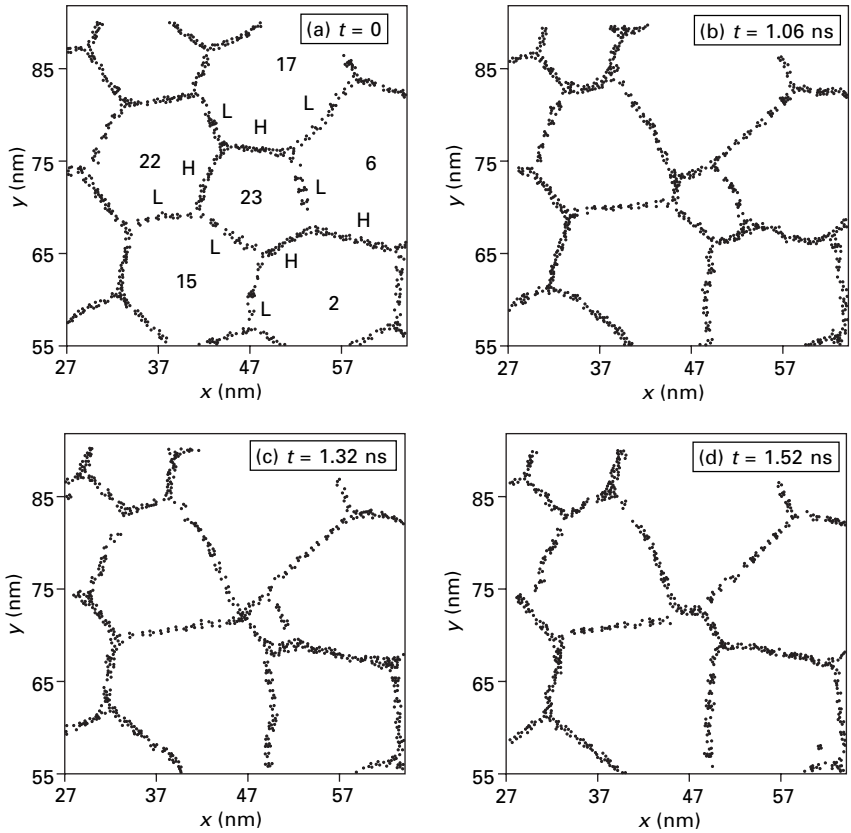
the rotation of grain 14 towards its coalescence with grain 8 and resulting in the elongated grain 8–14. In the initial configuration in Fig. 4.2(a) the two grains are misoriented by ≈18°. According to Fig. 4.2(b), during the first 1.1 ns grains 8, 14 and 16 shrink due to GB migration; also, as indicated by the solid lines, grain 14 has already undergone some significant rotation towards the orientation of grain 8. The interplay between migration and rotation is further illustrated in Fig. 4.2(c) in which it can be seen that grain 14, while continuing to rotate, has actually grown again, by the mechanism of GB migration; it is also interesting to note that the inclination of the GB between grains 8 and 14 has changed during this process. Finally, the coalescence of grains 8 and 14 is evident in Fig. 4.2(d) with the two grains now being closely aligned, with the low-angle GB between them having been reduced to a single dislocation (the core of which shows up as the white spot in the center of the new grain 8–14).

A complementary view of these events discussed by Haslam *et al.* (2001) from the viewpoint of the miscoordinated atoms situated mostly in the GBs demonstrates that the rotation–coalescence mechanism may be viewed as an 'unzipping' of the GB between the two grains, i.e. the gradual increase in the spacing between the GB dislocations until the GB has disappeared, leaving behind the new, dislocation-free perfect-crystal grain 8–14. This coalescence event becomes the trigger for subsequent grain growth in the vicinity of the large, elongated grain.

The above observations demonstrate that grain growth can be triggered by the coalescence of two neighboring grains, resulting not only in the elimination of the GB between them but also the elimination of two triple junctions on two of their neighboring grains. This, in turn, leaves behind a highly curved GB on each of these grains, resulting in their destabilization and subsequent disappearance by a relaxation process that involves a region of considerable size in the vicinity of these two grains (for details, see Haslam *et al.*, 2001).

In a different region of this model system, the simulations yielded an example of grain growth dominated entirely by the conventional mechanism of curvature-driven GB migration. This is illustrated in Fig. 4.3, which shows how grain 23 is being 'eaten up' by its neighbors. The snapshots in Figs 4.3(a)–(d) reveal that after only 1.06 ns, grain 23 has already shrunk considerably (Fig. 4.3(b)) and, after 1.52 ns, it has completely disappeared (Fig. 4.3(d)).

Insight into the underlying driving forces may be gained from an analysis of the energetics of the GBs in the vicinity of grain 23. It turns out that the GBs delineating grain 23 have either $\theta_{ij} \geq 30°$ or $\theta_{ij} \leq 20°$. To distinguish these, in Fig. 4.3(a) the GBs are labeled either as 'H' (indicating 'high-angle') or 'L' (indicating 'low-angle'). The structures of the low-angle GBs consist of identifiable dislocation cores, although these cores are rather closely



4.3 Four snapshots illustrating grain growth by curvature-driven GB migration. (a) $t = 0$; (b) after $t = 1.06$ ns; (c) after $t = 1.32$ ns; (d) after $t = 1.52$ ns. In (a), the GBs surrounding the disappearing grain 23 are labeled 'H' (indicating a 'high-angle' GB, defined here by $\theta_{ij} \geq 30°$) and 'L' (indicating a 'low-angle' GB, defined by $\theta_{ij} \leq 20°$).

spaced due to the low-angle cutoff of 14.9° imposed when generating the initial microstructure; by contrast, in the high-angle GBs the miscoordinated atoms are distributed continuously all along the interface.

The comparison of Figs 4.3(a) and (b) reveals that during the shrinkage of grain 23, the high-angle, high-energy GBs (whose continuously disordered structure consists of completely overlapping dislocation cores) shorten and move significantly while, simultaneously, the inclination of the GB 23–17 changes, presumably in order to lower the energy of this high-angle GB. Remarkably, the lower-angle, lower-energy GBs (i.e. those exhibiting a discrete dislocation structure) migrate rather little; also, the lengths of the dislocation GBs 23–6 and 23–15 remain almost unchanged while the dislocation GB between grains 17 and 22 increases its length significantly (i.e. without

having to migrate). Finally, in Fig. 4.3(c) the high-angle GB between grains 22 and 23 has disappeared completely, leaving behind an unstable quadruple junction; the latter subsequently decomposes into two triple junctions while grain 23 disappears and a new GB between grains 15 and 17 is formed.

This analysis demonstrates a major role of the high-angle, high-energy GBs during the disappearance of grain 23 by curvature-driven GB migration; this role seems to be the result of not only their higher energy but also their higher mobility, consistent with theoretical arguments summarized by Humphreys (1998). Their role as 'drainable' energy reservoirs in the polycrystal becomes clear from Figs 4.3(a) and (b): on balance, the shrinkage of grain 23 eliminates a fraction of the high-energy GB area altogether while converting another fraction into a low-energy, dislocation GB. We therefore conclude that grain growth by curvature-driven GB migration is dominated by the high-energy GBs in the system: the large amount of energy released by their elimination or conversion into lower-angle GBs, coupled with their high mobility, represent the relaxation mechanism by which the polycrystal lowers its energy towards that of the single crystal.

4.2.3 Topological discontinuities during grain growth

In mesoscale simulations of microstructural evolution it is well known that processes and events which discontinuously change the topology of the evolving microstructure represent major computational hurdles towards making the simulations physically realistic. Such discontinuities are usually associated with GB segments that become too small to be considered as independent elements of the microstructure with well-defined physical properties. To overcome such topological bottlenecks in the simulation, physical assumptions and specific mechanisms must usually be introduced which transform the microstructure into a state from which it can continue to evolve in a physically meaningful manner.

Two such topological instabilities are well-known to occur during grain growth (Kermode and Weaire, 1990). The first is known as a T2 event by which a small, three-sided grain disappears. The second is a neighbor-switching event (now commonly referred to as a T1 event) during which two grains which were initially neighbors of each other separate along the common GB and move apart while, simultaneously, two other grains which were initially not neighbors move towards each other to form a new, common GB. It has been suggested that neighbor switching can occur not only during grain growth but also during high-temperature plastic flow (Ashby and Verrall, 1973; Zelin and Mukherjee, 1996).

Several T1 and T2 events were observed in the MD simulations of Haslam *et al*. (2001). An example of a T1 event can be seen in Figs 4.3(c) and (d), illustrating the mechanism by which grains 22 and 23 lose each other as

neighbors while grains 15 and 17 move towards each other to now become neighbors. The saddle-point configuration for the switch involves the four-grain junction clearly seen in Fig. 4.3(c); however, immediately after the switch the grain 23 disappears. Another example of a T1 process is discussed in more detail in the paper of Haslam *et al.*, as is a T2 event associated with the instability and subsequent disappearance of three-sided grains.

4.2.4 Extracting grain-boundary properties from atomistic simulations

As illustrated in the sections that follow, the transfer of insights gained from the MD simulations into a mesoscale representation of the system requires the analysis and quantification of both the driving and retarding forces controlling both GB migration and grain rotations. In turn this requires determination of the free energies and their derivatives with respect to the misorientation angle of all the GBs surrounding each grain. In addition one needs to determine the GB mobilities as well.

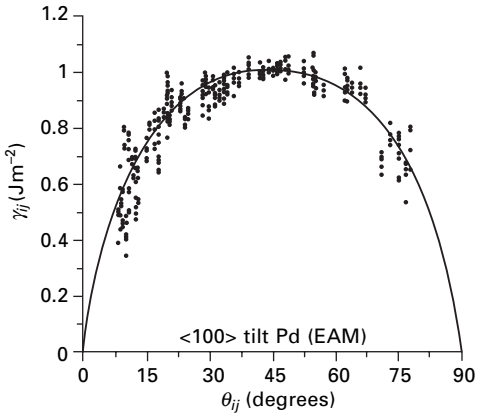
From the microscopic point of view the GBs are 2D regions separating two perfectly coordinated crystalline regions (the grains), their most obvious feature being the loss of perfect-crystal nearest-neighbor coordination of the GB atoms. In covalent materials such as Si, this loss gives rise to the formation of dangling bonds. Generally, the existence of the GBs in the system can be characterized by the Gibbsian excess energy per unit area, $\gamma = (E_N(GB) - E_N(id))/A$. This energy is readily determined by evaluating the energy of a system of N atoms containing the GB, $E_N(GB)$, and subtracting from it the energy of an N-atom perfect-crystal reference system, $E_N(id)$, that does not contain the interface of area A.

The GB energies shown in Fig. 4.4 were obtained for the 25-grain <100> columnar model system by Haslam *et al.* (2001) (see Section 4.2 and Fig. 4.1). The scatter of the data in Fig. 4.4 is due to the fact that the asymmetric tilt GBs in the system include a distribution in GB inclinations for any given misorientation. The solid line in Fig. 4.4 represents a least-square fit to a well-tested, empirical extension of the Read–Shockley equation (Read and Shockley, 1950), to high misorientation angles, θ_{ij}, between two grains i and j (Wolf, 1989):

$$\gamma(\theta_{ij}) = \gamma_{\max} \sin(2\theta_{ij}) \{1 - r \ln [\sin(2\theta_{ij})]\} \quad [4.1]$$

where $\gamma_{\max}$ is the plateau value and r is a measure of how steeply the GB-energy rises at small angles. The value of the parameters $\gamma_{\max}$ and r fitted for <001> tilt GBs in palladium are $\gamma_{\max} = 1.01\ \mathrm{Jm}^{-2}$ and $r = 0.693$ (Haslam *et al.*, 2001).

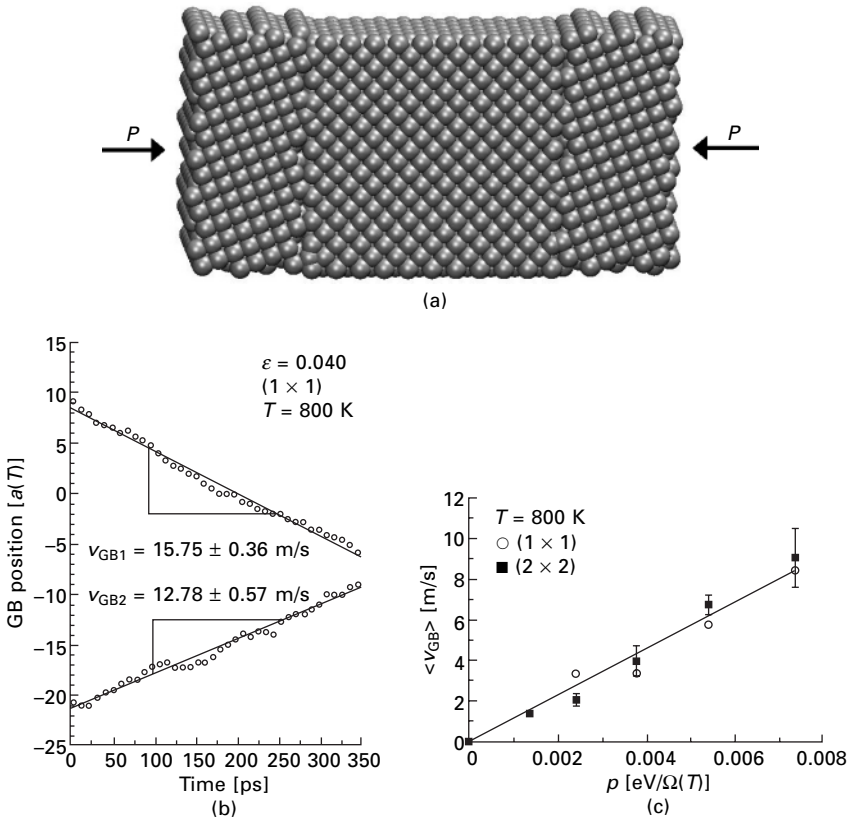
The GB mobility, which characterizes the viscous motion of a GB in response to a driving force acting on the boundary, is another key input



4.4 Variation of the energy γ_{ij}, of a GB between two grains *i* and *j*, as a function of the misorientation angle θ_{ij} for <001> tilt GBs in Pd obtained from the simulation of 24-grain polycrystal containing 24 hexagonal grains with a uniform grain shape and a grain size of ~ 14 nm. The solid line represents a least-squares fit of the data to equation (4.1). The larger scatter of the data points for any given misorientation is due to variation of the GB energy with the inclination of the GB plane (see Haslam *et al.,* 2001).

parameter for the mesoscale simulations. Two distinct MD approaches have been developed to study GB motion in pure materials. In the first one, Schonfelder *et al.* (1997) applied external strain (in the elastic regime) on a bicrystal to drive GB migration (see Fig. 4.5(a)). In this simulation the driving force for GB migration is given by the difference between the stored elastic energies in the two grains at fixed elastic strain. Indeed, as shown in Fig. 4.5(b), this simulation showed that for a given strain the two GBs present in the 3D periodic simulation cell move approximately with constant velocity towards each other, thereby enlarging the energetically favored grain at the expense of the grain with the higher elastic-energy density. These simulations also showed good linearity in the velocity–driving force relationship (see Fig. 4.5(c)), the slope of which is proportional to the GB mobility. The same simulation method was also used recently by Zhang *et al.* (2004). A second type of MD simulation method was developed by Upmanyu *et al.* (1998, 1999) and relies on capillarity (i.e. GB curvature) rather than elasticity to drive GB migration.

Very much like the GB energy, the GB mobility also depends on the GB misorientation and the crystallographic orientation of the GB plane. Unfortunately, however, current knowledge and understanding of this property is very much limited by the lack of good experimental results and paucity of simulation results. Based on the existing MD simulations and experimental data it is known that high-angle GBs have very high mobility that, like the GB energy, is independent of the GB misorientation (Humphreys and Hatherly,



4.5 (a) Schematic representation of a bicrystal simulation-cell used in a typical MD simulation study of elastically driven GB migration. Due to the 3D periodic border conditions imposed on the simulation cell the system will contain two identical GBs. (b) MD bicrystal simulation result of a Cu model system at T = 800 K and 0.04 elastic strain showing the average positions of the two GBs versus time (taken from Schonfelder *et al.,* 1997). Despite the stochastic nature of the GB migration process the displacement–time behavior of each GB is well represented by a linear relationship with slopes giving the drift velocities. (c) GB velocity averaged over the two GBs versus the elastic driving force at T = 800 K.

1995). Also, there is credible evidence (Humphreys, 1998) that the mobility of <001> tilt GBs with θ_{ij} in the range 10–20° increases with misorientation and saturates at about 20°. Finally, there is little evidence of any dependence of the mobility on θ_{ij} for angles smaller than typically about 10°; moreover, for $\theta_{ij} \approx 5°$ the mobility is thought to be about 10–100 times lower than for high angles. An empirical expression for the variation of the GB mobility with the misorientation capturing this dependence was given by Humphreys (1998):

$$m(\theta_{ij}) = m_{\max}\left\{1 - \exp\left[-B\left(\frac{\theta_{ij}}{\theta_0}\right)^n\right]\right\} \quad [4.2]$$

where $m_{\max}$ is the mobility of the high-angle GBs; B, θ_0 and n are materials parameters. For our mesoscale simulations an estimate for palladium gives $m_{\max} = 6.5 \times 10^{-8}$ m^4 J^{-1} s^{-1}, whereas the values used for the other parameters are those given by Humphreys (1998), i.e. $B = 5$, $n = 4$ and $\theta_0 = 20°$.

4.3 Mesoscale simulation methodology

The physical insights into the atomic-level mechanisms involved in grain growth gained from MD simulations can be transferred to the microstructural length and time scales, enabling mesoscale simulations of grain growth in which there are no atoms but merely GBs and grain junctions, with energies and mobilities that depend solely on the misorientations between the grains and the GB-plane inclinations. The inherent time scale of such simulations is therefore given by that associated with the processes of grain rotation and GB migration rather than by the vibrations of the atoms, and the inherent length scale is the grain size rather than the interatomic separations.

By contrast with Newton's law of motion coupling the acceleration of the atoms to the forces acting on them, the force laws governing GB migration and grain rotation are inherently viscous rather than conservative. A constant driving force for migration, P (in our case represented by the product of GB energy and GB curvature), results in a constant drift velocity, v, rather than an acceleration, according to (Humphreys and Hatherly, 1995)

$$v = mP \quad [4.3]$$

where m is the GB mobility.

The viscous force law in equation [4.3] is predicted by reaction-rate theory assuming that the mobility is independent of both the driving force and the details of the mechanism of GB migration. As described below, a similar law can be formulated for the angular velocity of some rotating grain, ω, with respect to an axis through its center of mass.

We finally note that, since little is known about the role of the GB-plane inclination in either the GB energy or GB mobility, throughout our mesoscale simulations we have ignored these possibly important two GB degrees of freedom. Moreover, we have assumed throughout that the triple junctions can move effortlessly in response to a driving force, thus ignoring any effects that a finite triple-junction mobility might have on the grain growth.

4.3.1 Theory of diffusion-accommodated grain rotation

The MD simulations revealed that for a grain size in the range of a few nanometers the microstructure evolves via two concurrent growth mechanisms,

(i) the conventional grain growth–grain shrinkage mechanisms mediated by GB migration and (ii) grain growth by the grain rotation–grain coalescence mechanism. In order to incorporate the latter into a mesoscopic representation, a phenomenological theory is required that analytically and quantitatively captures the MD observations.

The theoretical basis for incorporation of grain rotation into mesoscale simulations is provided by the theory of diffusion-accommodated grain rotation developed previously (Moldovan *et al.*, 2001). Since, as we shall demonstrate, the correct treatment of these grain rotations is essential to the successful linking of the atomic scale and the mesoscale, we briefly review the pertinent feature of this theory.

Considering a columnar microstructure of column height unity, the cumulative torque acting on some grain i with respect to its center of mass is given by (Harris *et al.*, 1998; Moldovan *et al.*, 2001)

$$\tau = \sum_j L_j \frac{\mathrm{d}\gamma_j}{\mathrm{d}\theta_j} \qquad [4.4]$$

where L_j denotes the length of the GB between grain i and some other grain j, with energy γ_j. θ_j is the misorientation angle across the GB. As observed in the MD simulations, analogous to GB migration grain rotation is a viscous process; i.e. the angular velocity of the grain, ω, with respect to an axis through its center of mass is given by

$$\omega = M\tau \qquad [4.5]$$

where M is the 'rotational mobility' of the grain. As an extension of the theory of diffusion-accommodated GB sliding by Raj and Ashby (1973), our theory of GB-diffusion or lattice-diffusion accommodated grain rotation yields a general expression for the grain-size dependence of M (Moldovan *et al.*, 2001)

$$M(R) = C/R^p \qquad [4.6]$$

where $p = 5$ for GB-diffusion accommodation and $p = 4$ for accommodation via lattice diffusion through the grain interiors. C is a physical parameter that depends on material properties, temperature T, the grain shape and the accommodation mechanism (Moldovan *et al.*, 2001). For example, for GB-diffusion accommodated grain rotation and a regular-hexagonal grain shape, $C = 95\Omega D_{\mathrm{GB}}\delta/\mathrm{k_B}T$, where Ω is the atomic volume, δ the diffusion width of the GBs along the periphery of the grain, D_{GB} the GB self-diffusion coefficient and k_B Boltzmann's constant.

4.3.2 Mesoscale representation of the system

The mesoscale representation of the microstructure consists of a network of interconnected GB segments that delimit the grains forming the microstructure.

Such an initial structure is usually generated via a Voronoi construction, with 2D periodic border conditions applied to the simulation cell (Weygand *et al.*, 1998). Analogous to the meshing in finite-element simulations, the GB network is discretized by two sets of mesh points (or nodes; Weygand *et al.*, 1998), one coinciding with the triple points and the other discretizing the GBs. Each node is linked to neighboring nodes by straight segments. For example, a triple-point node is connected to three neighboring nodes while a GB node is connected to only two other nodes (either or both of which may be associated with triple points). The introduction of the GB mesh allows GB curvature to be modeled in terms of a series of straight segments. This also permits the equilibrium condition at the triple points, given by Herring (1951) relation, to be fulfilled.

Each grain is assigned an initial orientation with respect to a fixed coordinate system in the plane of the simulation cell; therefore, each GB is characterized by the misorientation angle between two neighboring grains. Even in a columnar 2D microstructure the full GB texture covers a large parameter space, consisting of both the relative misorientations of any two neighboring grains and the inclination of the GB plane with respect to a reference direction. Regardless of the actual texture of the system, the development of a mesoscale representation requires prior knowledge of the misorientation dependence of both GB energies and mobilities. So as to enable a direct comparison with the MD system, we assume all GBs to be <001> tilt boundaries. In practice, to account for the four-fold rotation symmetry of the <001> texture, initial grain orientation angles are assigned randomly within the 0–90° interval. As further discussed in Sections 4.4 and 4.5, the dependence of the GB energy and GB mobility on the GB misorientation functional forms given by equations [4.1] and [4.2] and Fig. 4.4 are assumed.

In our 2D system, grains can rotate about the fixed texture axis in either the clockwise or counterclockwise direction. Any such rotation leads to a decrease in the total excess energy of the GBs delimiting the grain. According to equation [4.4] the GBs with the largest contribution to the cumulative torque are those with the smallest misorientation angle, as they have the largest value of the gradient $d\gamma(\theta)/d\theta$. The rate of grain rotation is given by equation [4.5], with a rotational mobility, M, that depends strongly on the average grain size (see equation [4.6]).

4.3.3 Variational functional approach and implementation

As a rigorous, formal basis for our mesoscale simulations we adopted the theoretical approach of Needleman and Rice (1980) based on a variational principle for dissipative systems. (For a broader discussion of this principle within the framework of 'thermomechanics', see Ziegler, 1977.) This principle was originally formulated for GB and surface diffusion in the context of void

growth (Needleman and Rice, 1980) and later adapted for the simulation of curvature-driven grain growth by Cocks and Gill (1996) and Gill and Cocks (1996). Their modification describes the rate of power dissipation due to the competition between the reduction in the GB energy and the viscous drag during GB migration. Based on the Cocks and Gill formulation of the functional, Cleri (2000) developed a stochastic formulation that enables a velocity Monte-Carlo (VMC) simulation approach using the variational functional as a transition-rate generating probability. Building on the work of Cocks and Gill (1996) and Cleri (2000), we have recently incorporated anisotropic GB properties (Moldovan *et al.*, 2002a) and grain rotations into this approach (Moldovan *et al.*, 2002b). In the following we briefly describe this comprehensive mesoscale methodology for the simulation of grain growth.

To incorporate grain rotations into the simulation model, additional terms have to be added to the functional describing curvature-driven GB migration (Cocks and Gill, 1996). These terms describe the competition between the reduction in GB energy due to the rotation and the related energy dissipation during the viscous rotation. In addition to the velocity field, $\{\boldsymbol{v}\}$, of all the GB and triple-point nodes, the dissipated-power functional, $\Pi(\{\boldsymbol{v}\},\{\omega\},\{\boldsymbol{r}\},\{\phi\})$, therefore includes also a set of grain-rotation rates, $\{\omega\}$. Here $\{\boldsymbol{r}\}$ represents the set of time-dependent coordinates of the GB and triple-point nodes, while $\{\phi\}$ is the set of grain orientations. This extended functional reads:

$$\begin{aligned}\Pi(\{\mathbf{v}\},\{\omega\},\{\mathbf{r}\},\{\phi\}) &= \sum_{i=1}^{N_{\mathrm{sgm}}} \Pi^{i}_{\mathrm{mig}} + \sum_{j=1}^{N_{\mathrm{g}}} \Pi^{j}_{\mathrm{rot}} \\ &= \sum_{i}^{N_{\mathrm{sgm}}} \left[\gamma_i(\mathbf{v}_{i1}\cdot\mathbf{s}_{i1} + \mathbf{v}_{i2}\cdot\mathbf{s}_{i2}) + \frac{l_i}{6m_i}[(v^n_{i1})^2 + (v^n_{i2})^2 + (v^n_{i1}v^n_{i2})] \right] \\ &\quad + \sum_{j=1}^{N_{\mathrm{g}}} \left[\tau_j\omega_j + \frac{\omega_j^2}{2M_j} \right] \end{aligned} \qquad [4.7]$$

where the first sum is taken over all the GB segments, N_{sgm}, of lengths l_i used to represent the discretized GB network; the second sum includes all N_{g} grains in the system (see also equations [4.3]–[4.5]). The subscripts 1 and 2 in the first two terms in equation [4.7] (describing the energy balance during curvature-driven GB migration) indicate that the values of the quantities are taken at the two end points of each segment; $\boldsymbol{s}_{i1}$ and $\boldsymbol{s}_{i2}$ are the tangent unit vectors at the ends of segment i, considered to point away from the end of the segment. $\boldsymbol{v}_{i1}$ and $\boldsymbol{v}_{i2}$ are the velocities of the nodes associated with segment i. v^n_{i1} and v^n_{i2} are the normal components of the velocities at the end points of segment i. γ_i and m_i, the GB energy and mobility of segment i, depend on the misorientation across the segment.

We use Cleri's velocity Monte-Carlo approach (Cleri, 2000) to find the

set of GB-node velocities, $\{\mathbf{v}\}$, that minimizes the variational functional at each time step. The grain angular velocities, $\{\omega\}$, are given by equation [4.5]. Since the GBs are discretized, the Herring relation for the triple-point angles (Herring, 1951) is satisfied in every time step, an assertion verified in our simulations. This results in the GBs being curved, providing the necessary driving force for GB curvature-driven grain growth.

In addition to the well-known T1 and T2 topological discontinuities associated with neighbor switching and the disappearance of three-sided grains (Weaire and Kermode, 1984; see also Section 4.2.3), grain coalescence gives rise to the topological discontinuities associated with the disappearance of a GB and the delimiting two triple junctions. In practice, this topological discontinuity is incorporated in our simulations by considering two neighboring grains as having coalesced when the misorientation of their common GB is less than a certain minimal angle; at that point, the appropriate topological rearrangements are carried out. This minimal angle can be rationalized by considering the dislocation spacing, d, in a low-angle GB given by $d = b/\sin\theta \sim b/\theta$ where b is the Burgers vector. As seen in the MD simulations, to be physically meaningful any GB segment of length L_{GB} should contain at least two dislocations (see Section 4.2.1). The condition that $d < L_{GB}/2$ therefore defines a minimum misorientation angle, $\theta_{min} \sim 2b/L_{GB}$. For example, for a grain size of $d = 100$ nm and a regular hexagonal grain shape, this yields a value of $\theta_{min} \sim 0.5°$; by comparison, for $d = 14$ nm, as is the case for the 25-grain Pd system considered in the MD simulations, $\theta_{min} \sim 2.7°$.

4.3.4 Dimensional analysis of characteristic length and timescales

In conventional grain-growth simulations (representing coarse-grained materials) in which growth occurs solely by the mechanism of curvature-driven GB migration, no absolute length scale is present. However, in nanocrystalline materials the simultaneous presence of both GB migration and grain rotation introduces a physical length scale into the system, in the form of a physical grain size, R_c. Next we estimate the value of this physical length scale by using dimensional analysis (for details, see Moldovan *et al.*, 2002b).

For this purpose let us consider a single representative term in equation [4.7]. In the presence of both GB migration and grain rotation, this term has the form, $\Pi = \Pi_{mig} + \Pi_{rot} = [\gamma v + (v^2/m)l] + [\tau\omega + (\omega^2/M)]$. To write Π in dimensionless form we express (i) the GB energy, γ, in units of γ_{max} (see equation [4.2]), (ii) the GB mobility, m, in units of m_{max} (see equation [4.3]), (iii) the length, l, of a GB segment in units of the initial average grain radius, R_0, and (iv) the time, t, in units of $t_0 = R_0^2/\gamma_{max}m_{max}$.

Having chosen these fundamental physical quantities as reference units,

the reduced-units representation of all the other physical quantities of interest are: $v_0 = \gamma_{max} m_{max}/R_o$ for the node velocity v, $\omega_0 = \gamma_{max} m_{max}/R_0^2$ for the angular velocity ω, and $\tau_0 = R_0 \gamma_{max}$ for the torque τ. This choice allows us to define the dimensionless variables $\gamma^* = \gamma/\gamma_{max}$, $m^* = m/m_{max}$, $R^* = R/R_0$, $v^* = v/v_0$, $\tau^* = \tau/\tau_0$ and $\omega^* = \omega/\omega_0$. Using these, the above sum of the two representative variational-functional terms can be written in dimensionless form as follows:

$$\Pi^* = \frac{\Pi}{\Pi_0} = \left[\gamma^* v^* + \frac{v^{*2}}{m^*} l^* \right] + \left[\tau^* \omega^* + \frac{1}{\eta} \frac{\omega^{*2}}{M^*} \right] \quad [4.8]$$

where $\Pi_0 = (\gamma_{max}^2 m_{max})/R_0$ and $M^* = (R_0/R)^p$ (see (Moldovan *et al.*, 2002c). It is important to notice that here we have introduced the dimensionless parameter $\eta = (C/m_{max} R_0^{p-3})$. The time evolution of the microstructure described by this functional therefore depends on the value of the parameter η, which can be written as follows:

$$\eta = \frac{C}{m_{max} R_0^{p-3}} = \left(\frac{R_c}{R_0} \right)^{p-3} \quad [4.9]$$

Here

$$R_c = \left(\frac{C}{m_{max}} \right)^{\frac{1}{p-3}} \quad [4.10]$$

is a characteristic *physical* length scale which depends on the GB characteristics such as the mobilities, diffusivities and temperature. For example, for a hexagonal grain shape $C = 95 \delta D_{GB} \Omega/kT$ (Moldovan *et al.*, 2001) and equation [4.10] yields:

$$R_c = \left(\frac{95 \delta D_{GB} \Omega}{m_{max} k_B T} \right)^{\frac{1}{p-3}} \quad [4.11]$$

with the same meaning of the physical quantities involved as defined in Section 4.3.1. In particular, we recall that $p = 5$ (4) for GB-diffusion (lattice-diffusion) accommodation of grain rotation.

According to equation [4.10], the physical length scale R_c is essentially the ratio of two materials constants, one characterizing the grain rotational mobility (see equations [4.5] and [4.6]) and the other the GB mobility. Interestingly, R_c does not depend on the topological or geometrical characteristics of the microstructure, such as the average grain size or the grain-size distribution. Therefore, given that the rates of GB-migration and grain-rotation-induced grain growth are proportional to these respective mobilities, R_c may be viewed as a reference grain size in a typical microstructure at which the rates of the two growth mechanisms contribute about equally to the growth.

According to equation [4.9], η may be viewed as the ratio of two lengths: the average initial grain size R_0, which depends on the initial microstructure, and the physical length scale R_c determined solely by the temperature and material properties, such as the GB mobility, m_{max}, and the materials constant C. Moreover from the characteristic term of the variational functional in dimensionless form we infer the existence of two limiting values of the parameter η. Whereas $\eta = \infty$ characterizes the growth solely due to the grain rotation-coalescence mechanism, $\eta = 0$ describes the growth due to GB migration alone. Thus, when $R_0 >> R_c$, η is small and the growth is GB-migration dominated; by contrast, when $R_0 << R_c$, η is large and the growth is grain-rotation dominated.

We finally mention that the parameter η can be viewed analogous to the Reynolds number in the analysis of the Navier–Stokes equations describing the incompressible flow of viscous fluids in hydrodynamics. The Reynolds number characterizes the relative importance of inertial and viscous forces; at a certain value of the Reynolds number the flow regime changes from laminar to turbulent flow. It is also interesting to recall that the Reynolds number is used to characterize the flow similarity; for example, the flows around two geometrically similar bodies will be identical (in dimensionless variables) if the Reynolds numbers for the two bodies are the same. Introduction of the parameter η has similar implications. In particular, a certain critical value of η marks the transition from GB-migration-dominated to rotation-dominated growth and can be used to study similarity of the growth.

4.4 Validation of mesoscale simulations

In the mesoscale methodology described above, the objects that evolve in time and space are the GBs and grains, rather than the atoms themselves. Instead of the vibration periods of the atoms and the interatomic distances, the characteristic mesoscale time and length scales are therefore governed by the relevant microstructural processes themselves; in the case of grain growth these are the GB mobilities, the grain-rotational mobilities and the grain size. Having thus discarded the atoms from the MD simulation and having replaced Newton's law of motion by a collective, viscous force law, a critical test of our mesoscale methodology is its ability to reproduce the microstructural evolution observed in the MD simulations. To validate our mesoscale approach, here we describe mesoscale simulations of grain growth for exactly the same 25-grain columnar microstructure with an initial grain size of 14 nm that was investigated by MD simulation (see Section 4.2).

As described above, in the simultaneous presence of both growth mechanisms the variational functional depends explicitly on the dimensionless parameter η. Therefore, for our mesoscale simulations to correspond formally to the conditions used in MD simulations, we have to determine the appropriate value of η corresponding to $R_0 = 14$ nm.

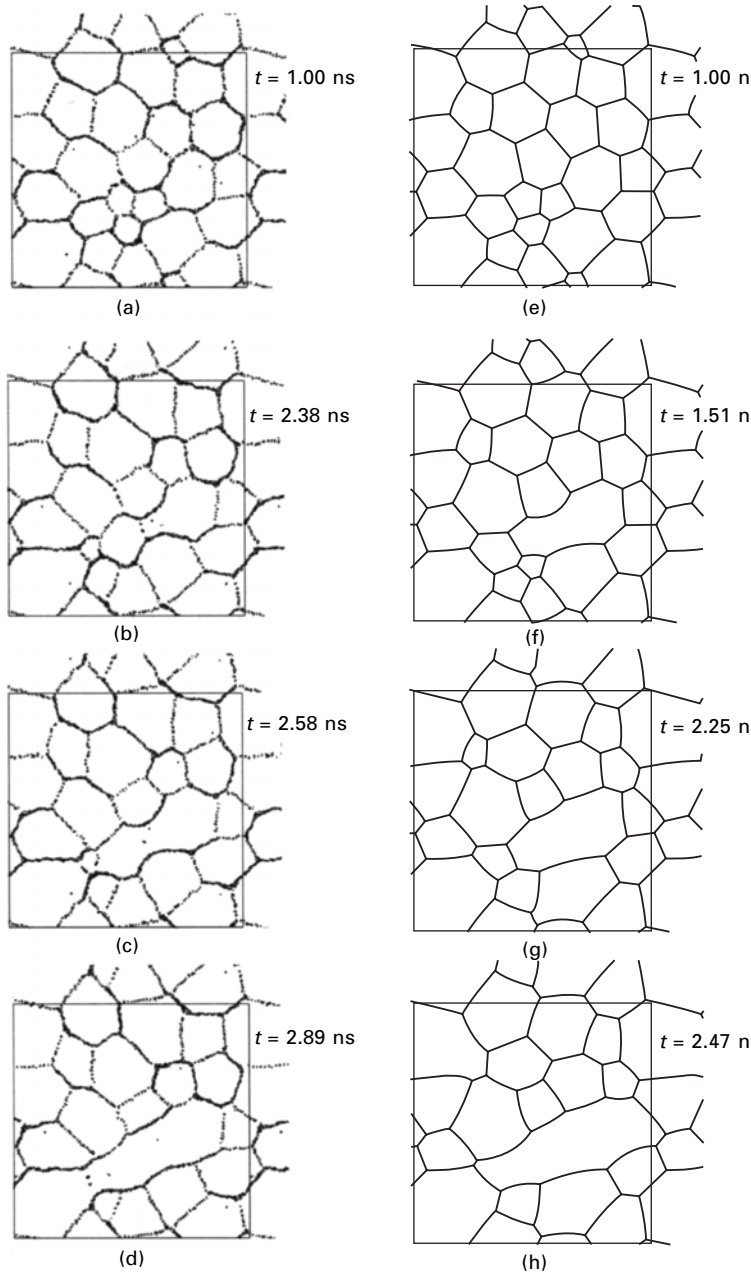
Table 4.1 Values of the parameters in equation [4.11] (in which we assume $p = 5$) used to estimate the characteristic grain size R_c for palladium at 1400 K. The temperature T, atomic volume $\Omega(T)$ at that temperature and initial grain size R_0 are fixed at the beginning of the simulation. The product $D_{GB}\delta$ of the GB diffusion constant and the diffusion width, the maximum GB mobility m_{max} and maximum GB energy γ_{max} were determined from independent simulations as described in Section 4.2.4 (see also Haslam *et al.*, 2001)

Parameter	Value
T	1400 K
$\Omega(T)$	15.5×10^{-30} m^3
R_0	14×10^{-9} m
$D_{GB}\delta$	6×10^{-19} m^3 s^{-1}
m_{max}	6.5×10^{-8} m^4 J^{-1} s^{-1}
γ_{max}	1.01 J m^{-2}

Inserting the values of the relevant physical parameters summarized in Table 4.1 into equation [4.11] and assuming $p = 5$, our estimate for palladium at 1400 K yields $R_c \approx 0.839$ nm and, hence, a value of $\eta \approx 0.0036$. This is the value used in our mesoscale simulations of the 25-grain columnar palladium polycrystal. In addition, the time must be converted from reduced units to physical units, $t_0 = R_0^2/\gamma_{max}m_{max}$ (see Section 4.3.4). For $R_0 = 14$ nm and the values from Table 4.1 for γ_{max} and m_{max}, we obtain $t_0 = 2.98$ ns.

Figures 4.6(a)–(h) compare the MD time evolution of the 25-grain system (Figs 4.6(a)–(d)) with the mesoscale model for $\eta = 0.0036$ (Figs 7.6(e)–(h)). The snapshots in Fig. 4.6(a) and (e) taken after 1.0 ns reveal excellent agreement between the two systems. Most importantly, both models predict a substantial decrease in the area of grain 23 (compare with the initial structure in Fig. 4.1(a)). Also, both models reveal a substantial change in the shapes and areas of grains 13, 17 and 24; moreover, the shape changes of grains 17 and 20 are well reproduced by the mesoscale model. Although, as the two systems evolve further, some small deviations start to develop between their evolution-paths, the mesoscale model predicts the disappearance of grain 23 by GB migration at about the same time as seen in the atomistic model. Moreover, it correctly describes the coalescence of grains 8 and 14 and of grains 16 and 9 and captures essentially the same time evolution of the system even after 3 ns.

This comparison demonstrates the ability of the mesoscale model to reproduce correctly the evolution of the MD system even in the presence of the grain rotations (Moldovan *et al.*, 2003a). However, great care is required in determining the appropriate values for the parameter η and the physical timescale characterizing the mesoscale model.



4.6 Four snapshots of the 25-grain system evolving by both MD (left-hand column) and mesoscale simulations for $\eta = 0.0036$ (right-hand column). This value of η corresponds to the physical parameters of the MD simulation shown in Fig. 4.1. Each snapshot gives the time in physical time units appropriate for the system at 1400 K.

4.5 Mesoscale simulation results

We are now ready to present our results of mesoscale simulations obtained for polycrystals containing thousands of grains and wide distributions in the grain misorientations, grain sizes and grain shapes. This allows us to investigate in a statistical manner the effects of the presence of both the GB migration and grain-rotation-coalescence mechanisms on the kinetic, topological and morphological aspects of grain growth. (For a full set of simulation results and a detailed discussion, see Moldovan *et al.*, 2002(a, b, c).)

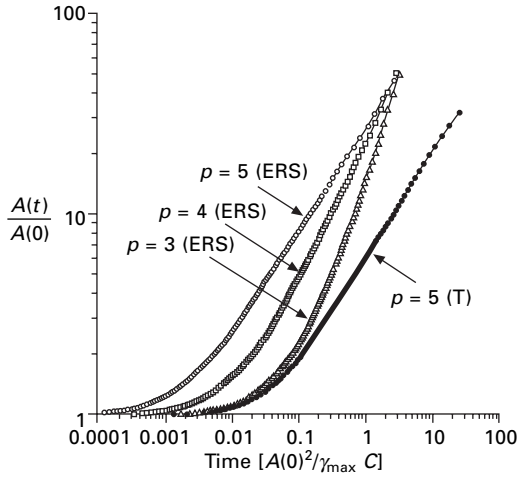
4.5.1 Grain growth by rotation-coalescence alone

We first focus on a (hypothetical) system evolving *exclusively* via the grain-rotation-coalescence mechanism. To explore the generality of this particular mechanism we consider two different functional forms for the misorientation dependence of the GB energy, $\gamma(\theta)$: (i) a realistic functional form for the <001> tilt GBs expressed as the extended Read–Shockley (ERS) formula in equation [4.1] and (ii) a simplified triangular (T) shape $\gamma(\theta)/\gamma_{max} = 4\theta/\pi$ for $0 < \theta < \pi/4$ and $\gamma(\theta)/\gamma_{max} = 2 - 4\theta/\pi$ for $\pi/4 < \theta < \pi/2$. Moreover, to improve the correlation between the value of the growth exponent, n (see below), and the grain-rotational-mobility parameter p in equation [4.6], we have performed simulations for $p = 3$, 4 and 5 employing both of these functional forms for $\gamma(\theta)$. While the cases for $p = 5$ and $p = 4$ find direct physical interpretation in terms of a diffusion-accommodation mechanism (see Section 4.3.1 and Moldovan *et al.*, 2001), the value $p = 3$ is primarily of mathematical interest.

We first investigate the growth kinetics. Figure 4.7 shows a log–log plot of the time variation of the average grain area, $A(t)$, normalized to the initial average grain area, $A(0)$, for the systems with various values of the parameter p and for $\gamma(\theta)$ given by both the ERS and T forms. We follow the time evolution of the total number of grains from initially 10 000 until only about 300 grains remain. After some transition period, which is different for each value of p, each system reaches a scaling regime in which the growth is characterized by a power law (i.e. $A(t) \sim t^n$) with a growth exponent n. Interestingly, the growth kinetics for the two rather different functional forms of $\gamma(\theta)$ are very similar. This suggests that the growth exponent assumes a universal value in a way that is similar to the growth exponent in isotropic GB systems (e.g. soap froth).

Moldovan *et al.*, (2002c) developed a theoretical model based on the mean-field approach that predicts that the exponent, n, which under certain conditions is universal, depends only on the parameter p in equation [4.6], i.e. on the specific accommodation mechanism, according to

$$n = \frac{2}{p - 1} \qquad [4.12]$$



4.7 Time variation of the average grain area (normalized using the initial value) for various values of the parameter *p*. In these simulations both the extended Read–Shockley (ERS) and triangular (T) formulas are used for the GB energy dependence on the misorientation.

Table 4.2 Analytical and simulated values for the growth exponents n characterizing the time dependence of the average grain area, $A(t) \sim t^n$, for the system evolving in the presence of grain rotation only. Simulation results are reported for two functional types of the misorientation dependence of the grain boundary energy $\gamma(\theta)$

		n (simulation)	
p	n (Eq. 4.12)	$\gamma(\theta)$: extended Read–Shockley	$\gamma(\theta)$: triangular shape
5	0.50	0.50 ± 0.02	0.51 ± 0.02
4	0.66	0.66 ± 0.02	0.67 ± 0.02
3	1.0	0.99 ± 0.02	0.95 ± 0.02

Table 4.2 summarizes the simulation findings together with the theoretically predicted values of n. One can see that the growth exponents for both functional forms of $\gamma(\theta)$ are almost the same and in very good agreement with the theoretically predicted exponents given by equation [4.12].

These simulation results, corroborated by our analytical derivation (see Moldovan *et al.*, 2002c), thus suggests the existence of a new universal coarsening exponent, n, characterizing the growth in 2D domain structures by the mechanism of grain-rotation-induced grain coalescence. This coarsening exponent depends only on the parameter p characterizing the size dependence of the rotational mobility of the grains.

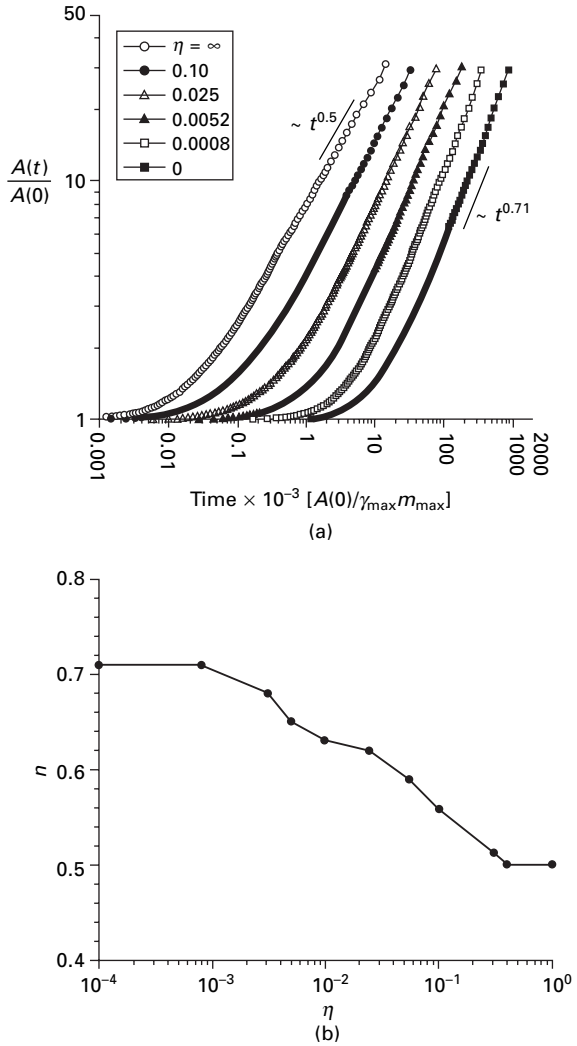
4.5.2 Grain growth by simultaneous GB migration and grain rotation

We now consider growth under the combined effects of both grain rotation and GB migration. As derived by dimensional analysis in Section 4.3.4, the parameter of the model that has to be 'tuned' in order to set the system in various growth regimes is η, given by equations [4.8] and [4.9].

We first analyze the growth law to determine the range of η values for which grain rotation and GB migration compete on an equal basis. We then perform a detailed analysis for one particular value of η ($\eta = 0.025$) in this crossover regime. Figures 4.8(a) and (b) reveal that the value of the growth exponent, n, in the growth law for the average grain area $A(t) \sim t^n$, depends strongly on the parameter η. Specifically (see Fig. 4.8(b)), we find that for $\eta < 0.001$, $n \approx 0.71$, i.e. the value for the system in the presence of GB migration only (Moldovan *et al.*, 2002a). By contrast, for $\eta > 0.3$ we find $n \sim 0.50$, i.e. the value for growth by grain-rotation-coalescence only (see also Fig. 4.7 and Table 4.2). In the range $0.001 < \eta < 0.3$, n decreases smoothly from $n = 0.71$ to $n = 0.5$, the middle of the crossover (in the linear-log representation of n vs. η, see Fig. 4.8(b)) corresponding to a value of $\eta \approx 0.025$.

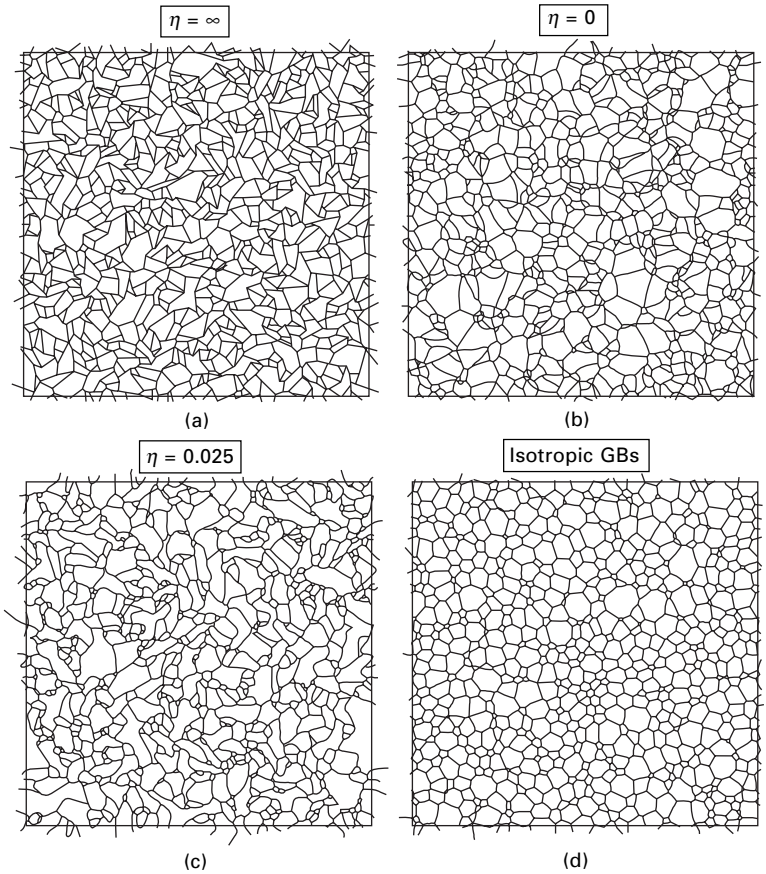
By rewriting equation [4.9] (with $p = 5$) as $R_0 = R_c/\eta^{1/2}$ one can adopt the alternate view of the crossover regime in terms of the initial average grain size R_0. We find that the GB-migration-dominated regime corresponds to large grain sizes, i.e, $R_0 > 31R_c(\eta < 0.001)$, while for $R_0 < 1.82R_c(\eta > 0.3)$ the system is in the grain-rotation-dominated regime. When considering these values one should bear in mind that, provided the growth process is followed over a long time such that significant changes of the average grain size occur, a continuous change from rotation-dominated to migration-dominated growth is obtained. In our simulations for growth in the scaling regime (evolution from 5000 to 500 grains) the average grain size increases by roughly a factor of three.

Figure 4.9 shows four snapshots of evolving microstructures with different values of η ranging between 0 and ∞; all snapshots contain about 700 grains evolved from initially 10 000 grains. It is interesting to see that in the presence of grain rotation only ($\eta = \infty$; see Fig. 4.9(a)), most of the grains become rather elongated. In addition, close inspection of this microstructure indicates the presence of a relatively small number of small grains. One can, therefore, anticipate a narrower grain-size distribution compared with the cases when GB migration is present. By contrast, for $\eta = 0$ (see Fig. 4.9(b)) there are no significantly elongated grains in the microstructure. In this sense the microstructure strongly resembles the isotropic microstructure shown in Fig. 4.9(d). However, there are some differences due mainly to the triple-point angles deviating from 120°, which is a consequence of the presence of both low-angle and high-angle GBs in the system.



4.8 (a) Time variation of the average grain area (normalized to the initial value) for a system characterized by various values of the parameter η. Notice that $\eta = \infty$ system evolves in the presence of grain-rotation coalescence only while $\eta = 0$ system evolves in the presence of curvature-driven GB migration only. (b) η dependence of the grain-growth exponent, n, for the system evolving in the presence of both grain rotation and GB migration.

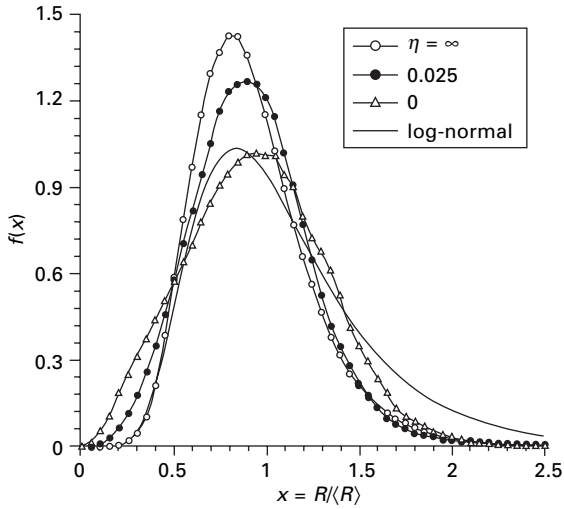
The morphology of the microstructure is quite different when both GB migration and grain rotation are present (see the system for $\eta = 0.025$ in Fig. 4.9(c)). This microstructure exhibits characteristics of both microstructures in Figs 4.9(a) and (b). For example, one can see the presence of some



4.9 Four typical snapshots of the microstructure when only about 700 (from initially 10 000) grains are left in a system evolving under various growth conditions. (a) Grain rotation only ($\eta = \infty$); (for this system the straight GB model was used in the simulations); (b) curvature-driven GB migration only ($\eta = 0$); (c) Simultaneous grain rotation and GB migration ($\eta = 0.025$); and (d) isotropic system, i.e. GB migration only in a system with isotropic GB properties.

elongated grains, the signature of grain rotation; also, the presence of some rather small grains compared to the average grain size reflects a broader distribution of the grain sizes which is a characteristic of GB migration only (see Fig. 4.9(b)). We also find that in the presence of GB migration (both with and without grain rotation), the dihedral angles between boundaries at the triple junctions quickly assume their equilibrium values and the Herring relation (Herring, 1951) is satisfied.

It is well-known that for curvature-driven grain growth, the average grain size $\langle R \rangle$ is a valid scaling parameter, which reduces the grain-size distribution



4.10 Grain-size distribution functions vs. reduced grain size, $x = R/\langle R \rangle$, for $\eta = \infty$, $\eta = 0.025$ and $\eta = 0$. For comparison, the log-normal distribution fitted to the distribution for GB migration only is shown.

function, $f(x) = f(R/\langle R \rangle)$, at different times to a common curve. Interestingly, our simulations reveal that this is still valid even for $\eta \neq 0$. Figure 4.10 compares $f(x)$ for $\eta = \infty$, $\eta = 0.025$ and $\eta = 0$ evaluated by averaging over a number of configurations while the system grew in the scaling regime. For comparison, we also show the log-normal distribution function fitted to the distribution in the presence of GB migration only ($\eta = 0$).

The shapes and peak positions are different for the three distribution functions. The distribution function for $\eta = \infty$ is narrower than for the other two systems and has a higher peak value. The peak position is around the reduced grain size $x_m \approx 0.8$, i.e. smaller than the corresponding peak values for the other systems. Notice also that for $\eta = \infty$ the distribution function drops more quickly to zero at small grain sizes, showing an apparent cutoff at $x_c \approx 0.2$. The presence of GB migration (for $\eta = 0$ and $\eta = 0.025$) leads to a widening of the grain-size distribution and a shift of the peak positions towards larger reduced grain sizes. Moreover, GB migration tends to eliminate the cutoff threshold x_c at small grain sizes; i.e. grains which are much smaller than the average start to be present more frequently in these systems. One can understand this by recalling that under GB migration small grains will shrink continuously until they disappear (by a T2 event or by a combination of T1 and T2 events). However, when only grain rotation takes place, there is no mechanism by which a grain can decrease its size. Rather, a grain may either disappear or may only increase its size by discrete coalescence events with its neighboring grains. Moreover, since the rotational mobility depends

so strongly on the average grain size (see equation [4.6]), small grains will be more likely to undergo coalescence events, and therefore disappear or grow faster, than larger ones.

Finally, we mention that there is no consensus on the precise form of the grain-size distribution of a microstructure evolving under curvature-driven grain growth. While experimental data suggest that a log-normal distribution function is appropriate, neither theoretical models nor simulation studies (see, e.g. Fig. 4.10) have reproduced such a distribution. Therefore, the deviations from the log-normal distribution are not surprising.

4.5.3 Extended von Neumann–Mullins relation

The mechanistic insights and GB parameters extracted from the atomic-level simulations, properly transferred into a mesoscopic conceptual and computational framework, have provided novel insights into not only the physics but also the statistical mechanics of grain growth. The ultimate goal of our multiscale simulations is, of course, to capture all these insights theoretically in the form of an analytic formulation of the process of grain growth. In the following we describe a simple theoretical framework capable of capturing the above simulation results.

For a system with isotropic GB properties evolving by curvature-driven GB migration, von Neumann (1952) derived the key equation

$$\frac{\mathrm{d}A_n}{\mathrm{d}t} = \frac{\pi \gamma m}{3}(n-6) \qquad [4.13]$$

relating the rate of area change of a grain to its number of sides, n, and to the product of the GB energy γ and mobility m. Equation [4.13], known as the von Neumann–Mullins (VNM) relation, allows a simplified description of a coarsening structure in an isotropic system in terms of only two variables: the area, A_n and the topological class, n, i.e. the number of sides of each grain. According to equation [4.13], the rate of area change for each grain is independent of the boundary shapes and of the rates of growth of the neighboring grains.

Focusing on the effect of GB anisotropy on grain growth, Moldovan *et al.* (2002a) recently investigated the adherence of curvature-driven grain growth to an *extended*, *averaged* VNM relation:

$$\frac{\mathrm{d}<A_n(t)>}{\mathrm{d}t} = <\gamma_n m_n>(\pi - <\theta_n>)\left[n - \frac{2\pi}{\pi - <\theta_n>}\right] \qquad [4.14]$$

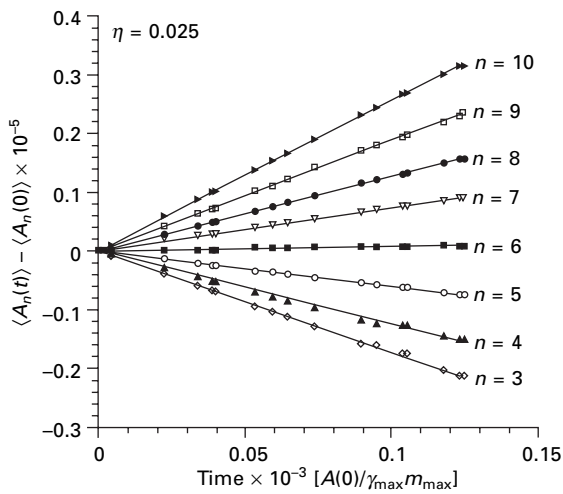
Here $<\gamma_n m_n>$ is the average value of γm for the GBs surrounding the grains of topological class n, and $<\theta_n>$ is the average value of the triple-point angles of the n-sided grains; <…> indicates an average over all grains with

n sides. One can see that equation [4.14] reduces to the VNM relation for isotropic systems by simply substituting $<\theta_n> = 2\pi/3$ and $<\gamma_n m_n> = \gamma m$.

We point out that equation [4.14] requires input from simulation, specifically the averaged quantities $<\gamma_n m_n>$ and $<\theta_n>$, whereas the VNM relation requires no simulation input. Our simulations reveal that, although the growth is considerably more complex in a system with anisotropic GB properties, the growth process can still be rationalized in terms of groups of grains belonging to the same topological class.

In the presence of both GB migration and grain rotation, two processes contribute to the change of the grain area. While GB migration leads to a continuous and smooth change of the area, growth due to the rotation–coalescence mechanism is sudden and discrete. Any rotation–coalescence event therefore changes the topological classes of the four grains involved. The validity of equation [4.14] is therefore not affected, except for the fact that the 'lifetime' of a given grain within its topological class is limited by any rotation–coalescence events in which the grain might participate.

Figure 4.11 shows the time evolution, over a relatively short period of time, of the *average* area, $<A_n(t)> - <A_n(0)>$, of n-sided grains in the scaling regime for the system with $\eta = 0.025$. One should notice that only grains that do not change their number of sides are considered when calculating $<A_n(t)>$. Although the rate of area change of individual grains in a topological class varies, the rate of change of the average area of the grains in each topological class is approximately constant. Figure 4.11 reveals that the average area of



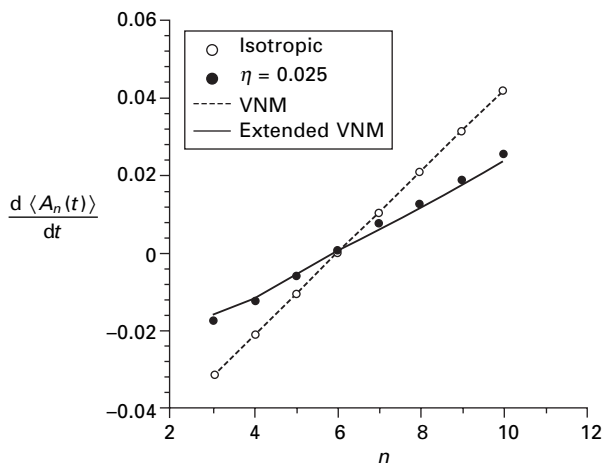
4.11 Time evolution of the average grain area, $<A_n(t)> - <A_n(0)>$, of n-sided grains for $\eta = 0.025$. Only grains that do not change their number of sides during the entire time interval of the averaging are used when calculating $<A_n(t)>$.

a six-sided grain does not change; by contrast, the average areas of grains with $n < 6$ decrease and those with $n > 6$ increase. Moreover the rates of average area change do not increase linearly with n, by contrast with the prediction from the VNM relation (see equation [4.13]).

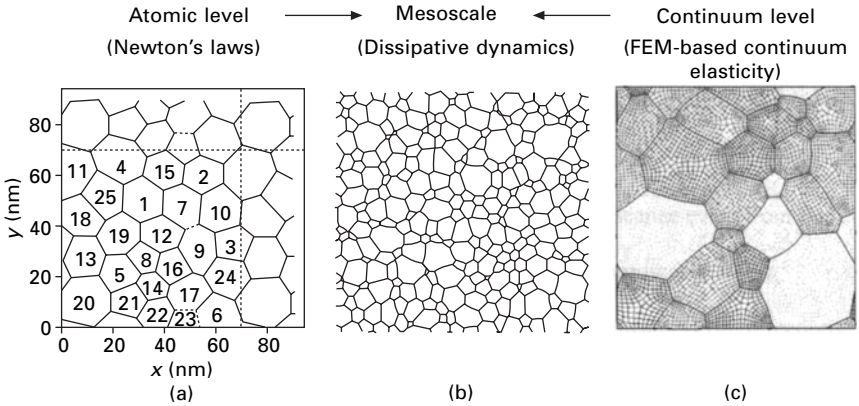
Figure 4.12 shows the simulation results for the rates of area change for grains belonging to various topological classes n, for both the isotropic system and for a system with $\eta = 0.025$. In addition the rates predicted from equation [4.13] and from the extended averaged VNM relation [4.14] are shown. The VNM relation is followed almost perfectly for the isotropic case. Although the average rate of grain-area change shows some deviation from linearity and is not at all represented by the VNM relation, for $\eta = 0.025$ the extended *averaged* VNM relation fits the simulation results well. This demonstrates that, although the degree of complexity increases in the simultaneous presence of anisotropic GB properties and grain rotation, the growth of grains which are not involved in any grain-rotation–coalescence events can still be rationalized in terms of groups of grains belonging to same topological class.

4.6 Summary and conclusions

For the simple model case of grain growth in a columnar polycrystalline microstructure, we have formulated a rigorous computational and theoretical framework which permits physical insights gained by MD simulation (see



4.12 Growth rates of the average areas of n-sided grains for the isotropic system and for the system with $\eta = 0.025$. The simulation results are well represented, respectively, by the von Neumann–Mullins relation (equation [4.13]; dotted line) and by the extended von Neumann–Mullins relation (equation [4.14]; solid line).



4.13 Multiscale simulation approach for polycrystalline materials illustrated for the case of grain growth in a polycrystalline fcc metal with a columnar microstructure. In (b) the objects of the mesoscale simulation that evolve in time and space are the discretized grain boundaries rather than the atoms. Instead of the vibration periods of the atoms and the interatomic distances, in (a), the related microstructural time and length scales are therefore given by the mobilities of the GBs and grains and by the grain size. The meshing of the grain interiors in (c) is necessary for determining the inhomogeneous stress distribution within the microstructure, while the atomic-level simulations are needed to establish the physical (i.e. mechanistic) basis for microstructural processes and to provide grain-boundary input parameters for the mesoscale simulation. The new paradigm for materials simulation is thus shifted from the continuum level down to the microstructural length and timescales, thus permitting atomic-level insights to be fully incorporated without losing the advantages of the finite-element approach.

Fig. 4.13(a)) to be transferred from the atomic level into the mesoscale (see Fig. 4.13(b)). This transfer requires quantification of the MD observations, for example, on the observed grain-growth mechanism(s) within a theoretical model. This model incorporates mesoscale physical properties, such as the GB energies and mobilities, extracted from the MD simulations as input parameters. This then enables examination of the statistical mechanics of the process in a realistic system, i.e. unencumbered by the well-known time and length-scale restrictions inherent to the MD simulations. Ideally, these mesoscale simulations enable the formulation of a continuum theoretical framework, with input parameters extracted from the mesoscale simulations, which quantitatively captures the insights into both the physics and statistical mechanics of the process gained from the simulations. In our simple model case, a generalized VNM relation was shown to capture analytically a key aspect of the mesoscale simulation results.

The centerpiece of our hierarchical multiscale approach consists of the

mesoscale methodology described in Section 4.3. It differs from the MD approach in two fundamental aspects. First, the mesoscale methodology naturally incorporates the rate-limiting microstructural time and length scales governed, respectively, by the GB and grain-rotational mobilities and by the grain size, rather than by atom-vibration periods and interatomic distances (compare Figs 4.13(a) and (b)). The well-known limitations of atomistic simulations pertaining to small system size and short inherent timescales are thus overcome.

Second, by contrast with Newton's law of motion coupling the acceleration of the atoms to the forces acting on them, the force law governing microstructural evolution is, by nature, viscous rather than conservative. Similar to the principle of maximum entropy production, a rigorous dynamical approach based on the functional of virtual-power dissipation can be formulated (Ziegler, 1977). This dissipative-dynamics approach uniquely defines the many-body path taken by the evolving system.

As an outlook we briefly describe how the effects of applied stress might be incorporated into the hierarchical multiscale approach outlined in this chapter. By meshing the grain interiors such that the grain-interior nodes link up with the already discretized GBs and grain junctions delimiting each grain (Fig. 4.13(c)), the inhomogeneous stress distribution arising from some externally applied stress can be computed using the finite-element approach. In an elastically anisotropic system, these stresses can provide an additional driving force for GB migration (Schoenfelder *et al.*, 1997) that acts against or in support of the capillarity driving force. At elevated temperatures, these stresses can also induce diffusion fluxes through the GBs and the grain interiors and, hence, give rise to diffusion creep and thus facilitate GB-diffusion accommodated processes, such as GB sliding and grain rotation (see Moldovan *et al.*, 2003b; Ding *et al.*, 2005).

Conceptually, this type of mesoscale approach may be viewed as a *dynamical* finite-element method (FEM) approach implemented, however, at the level of the rate-determining level of the materials microstructure. Operationally it increases the number of degrees of freedom in a conventional FEM simulation by the microstructural degrees of freedom associated with the GB network. The new paradigm for the simulation of engineering materials is thus shifted from the continuum level down to the critical microstructural length and timescales, thus permitting atomic-level insights to be fully incorporated without losing any of the strengths of the FEM-based continuum approach for the simulation of engineering structures.

We anticipate that over time this type of multiscale simulation methodology will have a major impact on the ability to simulate the properties of complex materials. The resulting novel class of simulations will provide a physical basis for realistic simulations of the complex interplay between diverse microstructural processes which limit the lifetime of engineering structures

subjected, for example, to materials fatigue and degradation. It will also provide a valuable tool for the rational design of materials with carefully tailored microstructures suitable for specific applications.

4.7 Acknowledgments

The work presented in this review chapter is part of a coordinated effort by the Argonne team during the period from 2000–2004. During this period we have enjoyed many stimulating discussions with our colleagues Andrew Haslam (now at Imperial College, London), Vesselin Yamakov (now at the National Insitute for Aerospace, Hampton, VA) and Simon Phillpot (now at the University of Florida). We are very grateful for their contributions that made this a most stimulating environment and team. D. Moldovan was supported by Louisiana Board of Regents (grant LEQSF(2003-06)-RD-A-13) and by NSF (grant CMS-0407785). D. Wolf was supported by the US Department of Energy, Office of Science under contract W-31-109-Eng-38.

4.8 References

Anderson M.P. and Grest G.S. (1989), Computer-simulation of normal grain-growth in 3 dimensions, *Phil. Mag. B* **59**, 293–329.

Anderson M.P., Srolovitz D.J., Grest G.S. and Sahni P.S. (1984), Computer-simulation of grain-growth. 1. Kinetics, *Acta Metall.* **32**, 783–791.

Ashby M.F. and Verrall R.A. (1973), Diffusion-accommodated flow and superplasticity, *Acta Metall.* **21**, 149–163.

Atkinson H.V. (1988), Theories of normal grain-growth in pure single phase systems, *Acta Metall.* **36**, 469–491.

Beenakker C.W.J. (1988), Numerical-simulation of a coarsening two-dimensional network, *Phys. Rev. A.* **37**, 1697–1702.

Burke J.E. and Turnbull D. (1952), Recrystallization and grain-growth, *Prog. Metal. Phys.* **3**, 220–244.

Chen D. (1995), Structural modeling of nanocrystalline materials, *Comput. Mater. Sci.* **3**, 327–333.

Chen L.-Q. (1995), A novel computer-simulation technique for modeling grain-growth, *Scripta Metall. et Mater.* **32**, 115–120.

Chen L.-Q. and Yang W. (1994), Computer-simulation of the domain dynamics of a quenched system with a large number of nonconserved order parameters – the grain-growth kinetics, *Phys. Rev. B* **50**, 15752–15756.

Cleri F. (2000), A stochastic grain-growth model based on a variational principle for dissipative systems, *Physica A* **282**, 339–354.

Cocks A.C.F. and Gill S.P.A. (1996), A variational approach to two dimensional grain-growth. 1. Theory, *Acta Mater.* **44**, 4765–4775.

D'Agostino G. and Van Swygenhoven H. (1996), Structural and mechanical properties of a simulated nickel nanophase, *Metastable Phases and Microstructures, MRS Symposia Proceedings No. 400*, Pittsburgh, 293–298.

Ding R., Moldovan D., Yamakov V. Wolf D. and Phillpot S.R. (2005), Effects of

microstructural inhomogeneity on dynamic grain growth during large-strain grain boundary diffusion-assisted plastic deformation, *Model. Simul. Mater. Sci. Eng*. **13**, 1129–1151.

Fan D. and Chen L.-Q. (1995), in the Proceedings of the American Ceramic Society Annual Meeting, Cincinnati, Ohio (1995).

Feltham P. (1957), Grain-growth in metals, *Acta Metall*. **5**, 97–105.

Foiles S.M. and Adams J.B. (1986), Thermodynamic properties of fcc transition metals as calculated with the embedded-atom method, *Phys. Rev. B* **40**, 5909–5915.

Fradkov V.E. Kravchenko A.S. and Shvindlerman L.S. (1985), Experimental investigation of normal grain-growth in terms of area and topological class, *Scripta Metall*. **19**, 1291–1296.

Fullman, R.L. (1952), in *Metal Interfaces*, Cleveland, ASM, OH: 179.

Galina A.V. Fradkov V.E. and Shvindlerman L.S. (1987), Influence of mobility and of triple grain junctions on grain boundary migration, *Phys. Met. Metall*. **63**, 165–168.

Gill S.P.A. and Cocks A.C.F. (1996), A variational approach to two dimensional grain growth. 2. Numerical results, *Acta Mater*. **44**, 4777–4789.

Gleiter H. (1989), Nanocrystalline materials, *Prog. Mater. Sci.* **33**, 223–315.

Gottstein G., Sursaeva V. and Shvindlerman L.S. (1999), The effect of triple junctions on grain boundary motion and grain microstructure evolution, *Interface Sci.* **7**, 273–283.

Grest G.S., Srolovitz D.J. and Anderson M.P. (1985), Computer-simulation of grain-growth. 4. Anisotropic grain-boundary energies, *Acta Metall*. **33**, 509–520.

Harris K.E., Singh V.V. and King A.H. (1998), Grain rotation in thin films of gold, *Acta Mater*. **46**, 2623–2633.

Haslam A.J., Phillpot S.R., Wolf D., Moldovan D. and Gleiter H. (2001), Mechanisms of grain-growth in nanocrystalline fcc metals by molecular-dynamics simulation, *Mater. Sci. Eng. A* **318**, 293–312.

Haslam A.J., Phillpot S.R., Moldovan D., Wolf D. and Gleiter H. (2002), Combined atomistic and mesoscale simulation of grain growth in nanocrystalline thin films, *Comput. Mater. Sci*. **23**, 15–32.

Haslam A.J., Moldovan D., Yamakov V., Wolf D., Phillpot S.R. and Gleiter H. (2003), Stress-enhanced grain-growth in a nanocrystalline material by molecular-dynamics simulation, *Acta Mater*. **51**, 2097–2112.

Haslam A.J., Yamakov V., Moldovan D., Wolf D., Phillpot S.R. and Gleiter H. (2004), Effects of grain growth on grain-boundary diffusion creep by molecular-dynamics simulation, *Acta Mater*. **52**, 1971–1987.

Herring C. (1951), in *The Physics of Powder Metallurgy* ed. W.E. Kingston, New York: McGraw-Hill, 143.

Hillert M. (1965), On the theory of normal and abnormal grain growth. Sur la theorie des croissances granulaires normale et anormale. Zur Theorie des normalen und des anomalen Kornwachstums, *Acta Mater*. **13**, 227–238.

Humphreys F.J. (1998), in *Grain Growth and Recrystallization in Polycrystalline Materials*, Proc. 3rd Int. Conf., eds H. Weiland, B.L. Adams and A.D. Rollett, Warrendale PA: TMS Publ., 13.

Humphreys F.J. and Hatherly M. (1995), *Recrystallization and Related Annealing Phenomena*, Oxford: Pergamon.

Kawasaki K., Nagai T. and Nakashima K. (1989), Vertex models for two-dimensional grain-growth, *Phil. Mag. B* **60**, 399–421.

Keblinski P., Phillpot S.R., Wolf D. and Gleiter H. (1997), Amorphous structure of grain boundaries and grain junctions in nanocrystalline silicon by molecular-dynamics simulation, *Acta Mater*. **45**, 987–998.

Keblinski P., Wolf D. and Gleiter H. (1998), Molecular-dynamics simulation of grain-boundary diffusion creep, *Interface Sci.* **6**, 205–212.

Keblinski P., Wolf D., Phillpot S.R. and Gleiter H. (1999a), Self-diffusion in high-angle fcc metal grain boundaries by molecular dynamics simulation, *Phil. Mag. A* **79**, 2735–2761.

Keblinski P., Wolf D., Phillpot S.R. and Gleiter H. (1999b), Structure of grain boundaries in nanocrystalline palladium by molecular dynamics simulation, *Scripta Mater*. **41**, 631–636.

Kermode J.P. and Weaire D. (1990), 2d-froth – A program for the investigation of 2-dimensional froths, *Comput. Phys. Commun*. **60**, 75–109.

Kumar S., Kurtz S.K., Banavar J.R. and Sharma, M.G. (1992), Properties of a three-dimensional Poisson–Voronoi tesselation: A Monte Carlo study, *J. Stat. Phys*. **67**, 523–551.

Kurtz S.K. and Carpay F.M.A. (1980), Microstructure and normal grain-growth in metals and ceramics. 1. Theory, *J. Appl. Phys*. **51**, 5725–5744.

Louat N.P. (1974), Theory of normal grain-growth, *Acta Metall*. **22**, 721–724.

Moldovan D., Wolf D. and Phillpot S.R. (2001), Theory of diffusion-accommodated grain rotation in columnar polycrystalline microstructures, *Acta Mater*. **49**, 3521–3532.

Moldovan D., Wolf D., Phillpot S.R. and Haslam A.J. (2002a), Mesoscopic simulation of two-dimensional grain growth with anisotropic grain-boundary properties, *Phil. Mag. A*. **82**, 1271–1297.

Moldovan D., Wolf D., Phillpot S.R. and Haslam A.J. (2002b), Role of grain rotation during grain growth in a columnar microstructure by mesoscale simulation, *Acta Mater*. **50**, 3397–3414.

Moldovan D., Yamakov V., Wolf D. and Phillpot S.R. (2002c), Scaling behavior of grain-rotation-induced grain growth, *Phys. Rev. Lett*. **89**, 206101–4.

Moldovan D., Wolf D. and Phillpot S.R. (2003a), Linking atomistic and mesoscale simulations of nanocrystalline materials: quantitative validation for the case of grain growth, *Phil. Mag*. **83**, 3643–3659.

Moldovan, D., Wolf D., Phillpot S.R., Mukherjee A.K. and Gleiter H. (2003b), Grain-boundary diffusion-controlled stress concentration in polycrystals, *Phil. Mag. Lett*. **83**, 29–38.

Needleman A. and Rice J.R. (1980), Plastic creep flow effects in the diffusive cavitation of grain-boundaries, *Acta Metall*. **28**, 1315–1332.

Phillpot S.R., Wolf D. and Gleiter H. (1995), Molecular-dynamics study of the synthesis and characterization of a fully dense, 3-dimensional nanocrystalline material, *J. Appl. Phys*. **78**, 847–861.

Raj R. and Ashby M.F. (1975), On grain boundary sliding and diffusional creep, *Metall. Trans.* **2**, 1113–1127.

Read W.T. and Shockley W. (1950), Dislocation models of crystal grain boundaries, *Phys. Rev*. **78**, 275–289.

Rivier N. (1983), On the structure of random tissues or froths, and their evolution, *Phil. Mag. B*. **47**, L45–L49.

Schiøtz J., Di Tolla F.D. and Jacobsen K.W. (1998), Softening of nanocrystalline metals at very small grain sizes, *Nature* **391**, 561–563.

Schiøtz J., Vegge T., Di Tolla F.D. and Jacobsen K.W. (1999), Atomic-scale simulations of the mechanical deformation of nanocrystalline metals, *Phys. Rev. B* **60**, 11971–11983.

Schoenfelder B., Wolf D., Phillpot S.R. and Furtkamp M. (1997), Molecular-dynamics method for the simulation of grain-boundary migration, *Interface Sci*. **5**, 245–262.

Smith C.S. (1964), *Mev. Mod. Phys.* **36**, 524–532.

Thompson C.V., Frost H.J. and Spaepen F. (1987), The relative rates of secondary and normal grain-growth, *Acta Metall*. **35**, 887–890.

Upmanyu M., Smith R.W. and Srolovitz D.J. (1998), Atomistic simulation of curvature driven grain boundary migration, *Interface Sci.* **6**, 41–58.

Upmanyu M., Srolovitz D.J., Shvindlerman L.S. and Gottstein G. (1999), Misorientation dependence of intrinsic grain boundary mobility: Simulation and experiment, *Acta Mater*. **47**, 3901–3914.

Van Swygenhoven H., Spaczer M. and Caro A. (1998), Role of low and high angle grain boundaries in the deformation mechanism of nanophase Ni: A molecular dynamics simulation study, *Nanostruct. Mater*. **10**, 819–828.

Van Swygenhoven H., Farkas D. and Caro A. (2000), Grain-boundary structures in polycrystalline metals at the nanoscale, *Phys. Rev. B* **62**, 831–838.

Von Neumann J. (1952), in *Metal Interfaces*, Cleveland, OH: ASM, 108.

Voronoi G. (1908), Nouvelles applications des parametres continus a la theorie des formes quadratiques, *Z. Reine Angew. Math.* **134**, 199–287.

Weaire D. and Kermode J.P. (1984), Computer-simulation of a two-dimensional soap froth. 2. Analysis of results, *Phil. Mag. B* **50**, 379–395.

Weaire D. and Lei H. (1990), A note on the statistics of the mature 2-dimensional soap froth, *Phil. Mag. Lett*. **62**, 427–430.

Weygand D., Brechet Y. and Lepinoux L. (1998), A vertex dynamics simulation of grain-growth in two dimensions, *Phil. Mag. B* **78**, 329–352.

Wolf D. (1989), A Read-Shockley model for high-angle grain-boundaries, *Scripta Metall*. **23**, 1713–1718.

Wolf D. (1990), Structure-energy correlation for grain-boundaries in fcc metals. 3. Symmetrical tilt boundaries, *Acta Metall*. **38**, 781–790.

Wolf D. and Merkle K. (1992), in *Materials interfaces, atomic-level structure and properties*, eds D. Wolf and S. Yip, London: Chapman and Hall, 87.

Yamakov V., Phillpot, S.R., Wolf, D. and Gleiter, H. (2000), in *Computer Simulation Studies in Condensed Matter Physics*, Vol. XIII eds D.P. Landau, S.P. Lewis and H.B. Schuettler, Berlin: Springer, 195.

Zelin M.G. and Mukherjee A.K. (1996), Geometrical aspects of superplastic flow, *Mater. Sci. Eng. A* **208**, 210–225.

Zhang H., Mendelev M.I. and Srolovitz D.J. (2004), Computer simulation of the elastically driven migration of a flat grain boundary, *Acta Mater*. **52**, 2569–2576.

Ziegler H. (1977), *An Introduction to Thermomechanics*, Amsterdam: North-Holland.

5

Finite element and homogenization modelling of materials

J. LLORCA, C. GONZÁLEZ and J. SEGURADO,
Polytechnic University of Madrid, Spain

5.1 Introduction

A basic principle of materials science is that the macrosocopic behavior of complex, heterogeneous materials depends mainly on the properties, volume fraction and spatial arrangement of the different phases and interfaces at length scales from nanometers to microns. This fact has led to the development of a number of simulation strategies which are grouped under the umbrella of continuum micromechanics, as the assumptions of continuum mechanics are valid in the range of length scales of interest. In contrast to other continuum mechanics approaches to materials, which develop constitutive equations for structural materials on the basis of fundamental considerations (deformation kinematics, thermodynamics, etc.), continuum micromechanics provides physically-based models in which the macroscopic behavior is linked to the details of the microstructure, and thus shows the route to optimize the behavior of materials for specific applications.

The basic strategy in continuum micromechanics is the simulation of the behavior of a representative volume element (RVE) of the microstructure of the material. The volume element should contain all the necessary information about the statistical description of the microstructure and, in addition, the RVE size should be large enough so that the average properties of this volume element are independent of its size and position within the material. Once the RVE has been defined (Section 5.2), the bridging of length scales (which stands as the central aim of continuum micromechanics) is carried out by obtaining the macroscopic (or effective) properties of the material from the analysis of the behavior of the RVE. This has been achieved traditionally through two different methods: homogenization techniques (Section 5.3) and the numerical simulation of the RVE behavior (Section 5.4). Combinations of the two techniques can be used to carry out multiscale modeling of materials and structures, and they are presented in Section 5.5 of this chapter. Future trends and expected developments in this area are briefly noted in the last section of this chapter.

Unless otherwise indicated, bold lowercase roman and greek letters stand, respectively, for first- and second-order tensors, and bold capital letters for fourth-order tensors. In addition, the different products are expressed as $(\mathbf{A}:\boldsymbol{\alpha})_{ij} = A_{ijkh}\alpha_{\mathrm{kh}}$, $\mathbf{A}::\mathbf{B} = A_{ijkh}B_{ijkh}$, and $(\boldsymbol{\alpha} \otimes \boldsymbol{\alpha})_{ijkh} = \alpha_{ij}\alpha_{kh}$. Finally, $\overline{\boldsymbol{\alpha}}$ represents the volumetric average of $\boldsymbol{\alpha}$.

5.2 Representative volume element

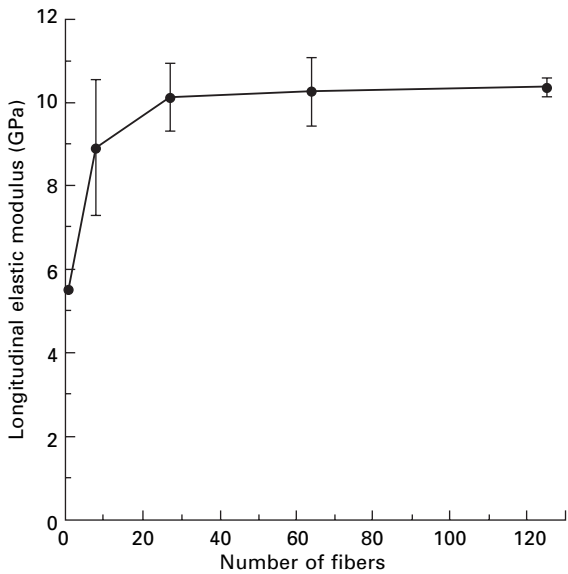
5.2.1 Concept of critical representative volume element

Continuum micromechanics relies on the simulation of the behavior of an RVE of the microstructure of the material. Evidently, the accuracy of the solution and the computer time required to solve the problem increase with the size of the RVE, and it is important to define clearly the meaning of RVE and also to provide analytical or numerical methods to determine its critical size.

Drugan and Willis (1996) identified two different ways to define the RVE. The first definition, based on the statistical nature of the material microstructure, defines the RVE as the smallest material volume statistically representative of the microstructure. The second definition is more pragmatic and defines the RVE just for a given material property or behavior. In this case, the RVE is the smallest volume of the heterogeneous material whose behavior represents accurately the actual macroscopic behavior. Accuracy is determined by the maximum allowed error between the exact value and that obtained from the RVE. This concept of critical RVE is more useful from the practical viewpoint and is normally used. Moreover, Drugan and Willis (1996) also introduced the concept of critical size of the RVE in the statistical sense. This guarantees the accuracy of the property obtained by averaging the results given by different RVEs of the same size, although the individual results may not be so precise. The critical size of the RVE in the statistical sense is always smaller, and very often it is cheaper, computationally, to simulate many small RVEs than a large one.

The critical RVE size to determine a material property with a prescribed accuracy depends on two parameters: the material microstructure and the phase properties. Analytical expressions to estimate that critical RVE size in the statistical sense have been rigorously developed only for specific cases. For instance, Monetto and Drugan (2004) determined the critical RVE size for the elastic behavior of an isotropic and homogeneous dispersion of ellipsoidal inclusions or voids embedded in a continuous matrix. They found that the critical size of the RVE was surprisingly small and that estimates of the elastic constants obtained on averaging the results provided by RVEs containing a few dozens of inclusions or voids were within 5 % of the exact value.

Analytical estimates of the critical RVE size are, however, exceptional and normally it has to be determined numerically for each given property in a heterogeneous material. The basic strategy is to compute the desired macroscopic property in RVEs with increasing size until the results converge to the actual value of an infinite size RVE (Gusev, 1997; Zohdi and Wriggers, 2001). The critical size is then determined as a function of the desired accuracy in the predictions. This methodology has been employed to determine the critical RVE size (in the statistical sense) for the elastic properties of a matrix containing a homogeneous dispersion of spherical inclusions or voids (Gusev, 1997), the thermo-elastic constants of aligned short-fiber composites (Hine *et al.*, 2002) or the tensile deformation of an elasto-plastic matrix reinforced with aligned ellipsoidal inclusions (Pierard *et al.*, 2007). The results obtained by Hine *et al.* (2002) for a random dispersion of aligned short fibers with an aspect ratio of 30 and a volume fraction of 0.15 are shown in Fig. 5.1, in which the average value of the longitudinal elastic modulus is plotted as a function of the number of fibers in the RVE. Three different realizations were generated and simulated for each number of fibers in the RVE, and the error bars in Fig. 5.1 show the 95 % confidence interval.



5.1 Estimations of the longitudinal elastic modulus of a short-fiber reinforced composite as a function of the number of fibers in the RVE. The solid circles stand for the average value and the error bars represent the 95 % confidence interval. The composite was made up of a homogeneous dispersion of aligned fibers with an aspect ratio of 30 embedded in a continuous matrix (see Hine *et al.*, 2002, for details).

These results show that RVEs with only 30 fibers were large enough to give predictions deviating only a few percent from the exact value provided that they were averaged over various realizations. Of course, bigger RVEs presented lower scatter, and the predictions obtained with 125 fibers in the RVE were practically superimposed. Thus, the optimum size of the RVE has to be determined in each case from the computational effort to generate and simulate many RVEs of small size or a few large ones. It is important to notice, however, that there is a minimum size of the RVE at which increasing the number of realizations does not improve the estimation of the property (Zohdi and Wriggers, 2001).

In general, the critical RVE size increases with the contrast in the phase properties, and larger RVEs are necessary to predict with the same accuracy the tensile deformation of a sphere-reinforced composite if the matrix behaves as an elasto-plastic solid than in the purely elastic case (González *et al.*, 2004). However, the critical factor controlling the RVE size is the development of percolation, which is very sensitive to the spatial arrangement of the phases. This behavior is typical of conductivity problems but may also arise in the simulation of the mechanical behavior, when plastic deformation or damage leads to the localization of the deformation in bands which propagate rapidly across the RVE. The macroscopic properties change suddenly once percolation has traveled through the RVE, and thus the effective response does not converge univocally to the exact value as the RVE size increases.

5.2.2 Microstructure generation and representation

As important as the size of the RVE to obtain the actual properties of the material is to ensure that the phase arrangement within the RVE reproduces faithfully the microstructure of the material. In general, the spatial distribution of the phases in the RVE can be obtained using different approaches. The first option is simply to reproduce in the RVE the phase arrangement provided by the microstructrural analysis of the material. This methodology was limited until very recently to materials with a two-dimensional microstructure, which were easily analyzed using quantitative metallography. The extension to materials with three-dimensional microstructure can be carried out by automated serial sectioning (Sidhu and Chawla, 2006) and X-ray computed microtomography (Youssef *et al.*, 2005). Automated serial sectioning works by taking optical micrographs after successive polishing cycles, from which a layer of material of controlled thickness is removed. The micrographs are analyzed using conventional image analysis software and segmented to black and white images, while retaining the true morphology of the phases in the microstructure. The segmented serial sections are stacked, and the three-dimensional microstructure is reconstructed. X-ray computed microtomography is based on X-ray radiography and provides detailed information at the micron

level in three dimensions of the spatial arrangement of the phases within the material. According to the Beer–Lambert law, each element in the recorded projection corresponds to a line integral of the attenuation coefficient along the beam path. The resulting image is a superimposed information of a volume in a plane. Three-dimensional images can be obtained by recording a large number of radiographs while the specimen is rotated between 0° and 180°. The spatial density distribution is then reconstructed by using a standard filtered back-projection algorithm.

These options to generate RVEs provide very realistic microstructures, but they are not suitable for all types of materials and often require very expensive experiments. A second approach to generate realistic RVEs uses the experimental data provided by two- and three-dimensional microstructural characterization techniques to compute two-point, three-point and higher-order correlation functions. These correlation functions are an important class of statistical descriptors that characterizate the spatial arrangement and heterogeneity of microstructural features (Tewari *et al*., 2004). This information can be used to reconstruct microstructures which, although not identical to the original material, show similar features as defined by the statistical correlation functions. Starting from an initial realization of the microstructure, the method proceeds to find a realization in which the calculated correlation functions best match the target functions. This is achieved by minimizing the sum of squared differences between the calculated and the target functions via stochastic optimization techniques (Torquato, 2001). The simulation of the properties of the generated microstructures is very useful to establish a direct link between microstructural statistical parameters which define the spatial arrangement of the phases and the macroscopic behavior. The disadvantage of the method is that the available reconstruction algorithms do not always provide a good representation of the original microstructure.

Finally, RVEs of model microstructures are often very helpful in numerical experiments in which the particular effect of one or more parameters is systematically analyzed. Recent examples of this methodology include the analysis of particle clustering on the tensile properties of sphere-reinforced metal–matrix composites (Segurado *et al*., 2003) or the influence of interface strength and toughness on the ductility of composites (Segurado and LLorca, 2005).

5.3 Homogenization techniques

Homogenization models were the workhorse to simulate the mechanical behavior of multiphase materials prior to the arrival of digital computers. They are still widely used because of the simplicity of their basic assumptions and their ability to describe accurately the overall or ‘effective’ behavior, particularly in the elastic regime. In addition, they often provide analytical

solutions for the constitutive equation of the materials, which can be used in structural analysis codes to assess the mechanical response of components. Detailed information on the homogenization tools available to simulate the behavior of heterogeneous materials can be found in Nemat-Nasser and Hori (1999) and Torquato (2001).

5.3.1 Basic equations

The central aim of the homogenization techniques is to describe the effective properties of the material from the mechanical properties and the microsocopic arrangement of the constituents. This bridging of the length scales is carried out by volume averaging and is denominated homogenization. Mathematically

$$\overline{\boldsymbol{\epsilon}} = \frac{1}{\Omega}\int_{\Omega} \boldsymbol{\epsilon}(\mathbf{x})\mathrm{d}\Omega \quad \text{and} \quad \overline{\boldsymbol{\sigma}} = \frac{1}{\Omega}\int_{\Omega} \boldsymbol{\sigma}\ (\mathbf{x})\mathrm{d}\Omega \qquad [5.1]$$

where $\boldsymbol{\epsilon}(\mathbf{x})$ and $\boldsymbol{\sigma}(\mathbf{x})$ are, respectively, the strain and the stress at a point in the material whose coordinates are given by the position vector $\mathbf{x}$, and $\overline{\boldsymbol{\epsilon}}$ and $\overline{\boldsymbol{\sigma}}$ stand for the macroscopic (or effective) strain and stress. The inverse operation, which provides the stresses and strains at a given point of the material from the effective properties, is denominated localization, and it is expressed as

$$\boldsymbol{\epsilon} = (\mathbf{x}) = \mathbf{A}(\mathbf{x}) : \overline{\boldsymbol{\epsilon}} \quad \text{and} \quad \boldsymbol{\sigma}(\mathbf{x}) = \mathbf{B}(\mathbf{x}) : \overline{\sigma} \qquad [5.2]$$

where $\mathbf{A}(\mathbf{x})$ and $\mathbf{B}(\mathbf{x})$ are the strain and stress concentration tensors. Evidently, the operations of homogenization and localization cannot be carried out exactly because of the complexity of the real materials, and it is necessary to make simplifications which lead to different homogenization models. The most important one is the mean-field approach, which approximates the stress and strain fields by phase-wise constant fields, $\boldsymbol{\epsilon}_i$ and $\boldsymbol{\sigma}_i$. Thus

$$\boldsymbol{\epsilon}_i = \mathbf{A}_i : \overline{\boldsymbol{\epsilon}} \quad \text{and} \quad \boldsymbol{\sigma}_i = \mathbf{B}_i : \overline{\boldsymbol{\sigma}} \qquad [5.3]$$

where $\mathbf{A}_i$ and $\mathbf{B}_i$ stand now for the strain and stress concentration tensors of the phase i. It follows from equation [5.1] directly that the stress and strain concentration tensors are related to each other by

$$\mathbf{I} = \sum_{i=1}^{N} f_i \mathbf{A}_i = \sum_{i=1}^{N} f_i \mathbf{B}_i \qquad [5.4]$$

where N is the number of phases, f_i the volume fraction of phase i, and $\mathbf{I}$ the identity tensor.

5.3.2 Linear behavior

If all the phases in the material are elastic, their constitutive equation can be expressed as

$$\boldsymbol{\sigma}_i = \mathbf{L}_i : \boldsymbol{\varepsilon}_i \qquad [5.5]$$

where $\mathbf{L}_i$ is the elastic stiffness tensor of phase i, and then the effective elastic stiffness tensor $\mathbf{L}$ (which relates the effective stress $\overline{\boldsymbol{\sigma}}$ with the effective strain $\overline{\boldsymbol{\epsilon}}$) can be computed as

$$\mathbf{L} = \sum_{i=1}^{N} f_i \mathbf{L}_i : \mathbf{A}_i \qquad [5.6]$$

The problem of obtaining the macroscopic properties of the heterogeneous material is thus reduced to determining the strain concentration tensors of the i phases in the material, which are related to each other through equation [5.5], and depend on the volume fraction, shape, spatial distribution and constitutive equation of each phase. Two obvious solutions are the Voigt and the Reuss models; the first one assumes that $\mathbf{A}_i = \mathbf{I},\ \forall i$ and leads to the well-known 'rule of mixtures' or isotrain approach, while $\mathbf{B}_i = \mathbf{I},\ \forall i$ in the second one, leads to the 'inverse rule of mixtures' or isostress model. However, these models (which lead to upper and lower bounds for the elastic properties) are normally very poor approximations, and better ones have been developed over the years.

The simplest one – within the framework of linear elasticity – is based on the pioneer work of Eshelby (1957), who analyzed the stress distribution in an elastic and isotropic ellipsoidal inclusion embedded in an elastic, isotropic and infinite matrix which is subjected to a remote strain $\overline{\boldsymbol{\epsilon}}$. Eshelby showed that the strain field within the inclusion was constant and given by (Eshelby, 1957; Hill, 1965)

$$\boldsymbol{\epsilon}_i = \mathbf{A}_i^{\text{dil}} : \overline{\boldsymbol{\epsilon}} \qquad \text{where} \qquad \mathbf{A}_i^{\text{dil}} = [\mathbf{I} + \mathbf{S}_i : \mathbf{L}_m^{-1} : (\mathbf{L}_i - \mathbf{L}_m)]^{-1} \qquad [5.7]$$

where $\mathbf{L}_m$ and $\mathbf{L}_i$ stand for stiffness tensors of the matrix and of the inclusion, respectively, and $\mathbf{S}_i$ is the Eshelby's tensor for the inclusion, whose components depend on the inclusion shape as well as on the matrix elastic constants.

The elastic stiffness tensor of the effective material, $\mathbf{L}$, can be obtained easily from equations [5.4], [5.6] and [5.7] as

$$\mathbf{L} = \mathbf{L}_{\text{m}} + f_i [(\mathbf{L}_{\text{i}} - \mathbf{L}_{\text{m}}) : \mathbf{A}_i^{\text{dil}}] \qquad [5.8]$$

where the superindex dil in the strain concentration tensor of the inclusion indicates that this expression is only exact when $f_i \to 0$, according to the hypotheses included in the work of Eshelby.

There are many variations of Eshelby's method to take into account the distortion in the matrix and inclusion stress fields induced by the presence of

neighbor particles. They can be used to compute the elastic properties of the heterogeneous material when the inclusion volume fraction is finite. One of the most popular ones is the Mori–Tanaka method (1973), which was reformulated by Benveniste (1987) in the context of the strain concentration tensors. Following Benveniste (1987), the inclusion strain concentration tensor, $\mathbf{A}_i^{\mathrm{MT}}$, can be obtained by interpolation between the strain concentration tensors for dilute conditions ($\mathbf{A}_i^{\mathrm{dil}}$) and when $f_i \to 1$ ($\mathbf{A}_i^{\mathrm{MT}} \to \mathbf{I}$). The simplest interpolation can be written as

$$\mathbf{A}_i^{\mathrm{MT}} = \mathbf{A}_i^{\mathrm{dil}} : [(1 - f_i)\mathbf{I} + f_i \mathbf{A}_i^{\mathrm{dil}}]^{-1} \qquad [5.9]$$

and the effective elastic properties in the case of a two-phase material can be computed as

$$\mathbf{L} = \mathbf{L}_m + f_i [(\mathbf{L}_i - \mathbf{L}_m) : \mathbf{A}_i^{\mathrm{MT}}] \qquad [5.10]$$

where f_{i} is the volume fraction of inclusions. The elastic constants provided by the Mori–Tanaka method provide an excellent solution in the case of a heterogeneous material with a matrix-inclusion topology in which the ellipsoidal inclusions (with the same orientation, aspect ratio and properties) are dispersed in a continuous matrix (Segurado and LLorca, 2002). For multiphase materials with matrix-inclusion topology (e.g the aspect ratio, orientation or properties of the inclusions are different), the best solution is provided by a two-step homogenization. Accordingly, the RVE is decomposed into grains, each one being viewed as a two-phase composite, whose properties are determined by the Mori–Tanaka method. The overall aggregate properties are computed by applying the Voigt or Reuss schemes over all the grains (Fig. 5.2) (Pierard *et al*., 2004).

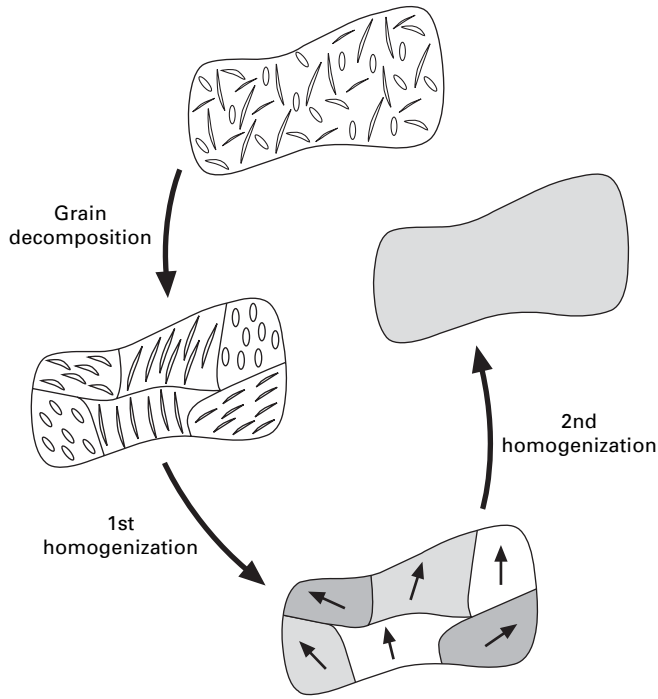
Another important mean-field approximation to the elastic properties of multiphase materials is the self-consistent method, which is particularly appropriate when the various phases form an interpenetrating network. This model was developed by Kröner (1958) to compute the effective elastic properties of polycrystalline solids. All the phases in the material are assumed to be embedded in an effective medium, whose properties are precisely the ones sought. The corresponding strain concentration tensor for each phase is obtained from Eshelby's dilute solution (equation [5.7]) substituting the matrix elastic constants by those of the effective medium, $\mathbf{L}$. Mathematically

$$\mathbf{A}_i^{\mathrm{SC}} = [\mathbf{I} + \mathbf{S}_i : \mathbf{L}^{-1} : (\mathbf{L}_i - \mathbf{L})]^{-1} \qquad [5.11]$$

and introducing equation [5.11] into equation [5.6] leads to

$$\mathbf{L} = \sum_{i=1}^{N} (f_i \mathbf{L}_i : [\mathbf{I} + \mathbf{S}_i : \mathbf{L}^{-1} : (\mathbf{L}_i - \mathbf{L})]^{-1}) \qquad [5.12]$$

which stands for a non-linear set of equations for the components of $\mathbf{L}$ which



5.2 Schematic of the two-step homogenization of multiphase composites with a matrix-inclusion topology.

can be solved numerically to obtain the elastic constants of the heterogeneous medium. Of course, Eshelby's tensor $\mathbf{S}_i$ depends on the shape and volume fraction of each phase as well as on the elastic constants of the effective medium given by $\mathbf{L}$.

Finally, it is possible to compute upper and lower bounds for the effective properties of the heterogeneous materials using the variational principles of the theory of elasticity and the information available on the microstructure. The simplest ones are the Voigt and Reuss bounds (Hill, 1963), which take into account only the volume fraction of each phase (*one-point* limits) and are normally of very little use. Hashin and Shtrikman (1963) derived much tighter bounds employing variational principles which involve polarization fields. These *two-point* bounds are the best possible ones for the effective elastic constants of two-phase materials which present a statistically isotropic microstructure given only the volume fraction information. They were improved more recently by *three-point* bounds (Torquato, 2001) which incorporate information about the phase arrangement through certain statistical correlation parameters.

5.3.3 Non-linear behavior and damage

All the techniques described above can be extended to study the mechanical behavior if one or more phases exhibit a non-linear behavior. Extension of these methods to deal with elasto-plastic phases requires a linearization procedure to re-write the local constitutive laws in such a way that homogenization schemes valid for linear elasticity might apply, and the effective stiffness tensor can be computed from those of each phase through the chosen linear approximation. The schemes for the analysis of elasto-plastic materials were reviewed recently by Ponte-Castañeda and Suquet (1998) and Chaboche *et al*. (2005); they can be divided into two main groups, depending on whether they are based on the use of the tangent or on the secant stiffness tensors of the phases.

Secant formulations deal with the plastic deformation within the context of non-linear elasticity, and the relationship between stress and strain in each phase is given by a secant stiffness tensor. They are only valid for monotonic and proportional loading, but they have been very popular because they are easy to implement. Predictions based on the tangent method were introduced in the seminal work of Hill (1965), who linearized the local constitutive laws written in rate-form and introduced an instantaneous elasto-plastic tangent modulus to compute the mechanical response through a step-by-step iterative procedure. This initial approach led, however, to too stiff predictions of the flow stress, and the origin of this error was traced to the anisotropic nature of the tangent stiffness tensor during plastic deformation. More precise estimations of the composite behavior were obtained by making use of isotropic or transversely isotropic projections of the tangent operator of the matrix instead of the anisotropic one in some steps of the computation of the macroscopic tangent operator (González and LLorca, 2000; Chaboche *et al*., 2005).

Although homogenization methods have been used exhaustively to compute the internal stresses in multiphase materials during deformation, their application to compute the mechanical behavior of materials, including progressive damage, has been very limited. This is because the mean-field approximations provide average magnitudes of the stress and strain fields in each phase, while damage is normally triggered by the extreme values. Moreover, damage leads to the rapid localization of the deformation, but this phenomenon cannot be captured easily by homogenization methods. Regardless of these limitations, homogenization models have been used to simulate the effect of interface decohesion and reinforcement fracture in composite materials. The standard strategy was to introduce progressively a new phase as the critical condition for damage was reached. Broken or decohered reinforcements were often represented by penny-shaped cracks or a transversally isotropic phase with zero stiffness perpendicular to the fracture plane, and various

algorithms were developed to redistribute the load shed by the damaged reinforcements into the matrix and the intact reinforcements (Estevez *et al*., 1999; González and Llorca, 2000).

5.4 Computational micromechanics

Computational micromechanics obtains the macroscopic or effective behavior of the heterogeneous material by solving the boundary value problem for an RVE of the microstructure by numerical methods. The simulations carried out within this framework provide the local values of the field variables, and thus can take into account accurately the nucleation and growth of damage and the subsequent localization of the strain upon deformation. These major advantages over the homogenization methods were hindered up to a few years ago by the power of digital computers, which limited the size of the RVE. In fact, most simulations were carried out in materials and loading conditions amenable of two-dimensional analyses (e.g. transverse loading of fiber-reinforced composites) or with simplistic RVEs formed by one or two microstructural features (particles, grains) and periodic boundary conditions (Brockenborough *et al*., 1991; LLorca *et al*., 1991). Despite these limitations, the simple computational micromechanics models developed in the 80s and 90s were very useful to clarify the relationship between some microstructural details (for instance, phase shape, volume fraction and aspect ratio) and the mechanisms of non-linear deformation and damage in heterogeneous materials.

5.4.1 Numerical strategies

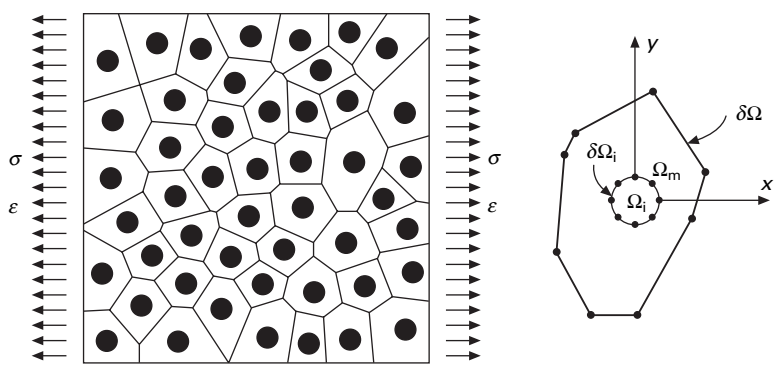
Advances in computing power and parallel developments of new simulation tools have opened the field of computational micromechanics to analyze the RVE of real materials of a size above the critical one. The numerical strategies to solve the boundary value problems have to be computationally efficient and easy to use with complex geometries. Fast Fourier transform algorithms (Moulinec and Suquet, 1994) and the boundary element method (Ingber and Papathanasiou, 1997) fulfill these conditions and have been used occasionally but the finite element method (FEM) is the standard tool for these simulations. The application of the FEM in computational micromechanics has been driven by the development of mesh generation programs, which lead to the quasi-automatic discretization of the RVE, and by the possibility of implementing complex constitutive equations for the different phases and interfaces (finite deformations, crystal plasticity, damage and fracture models, etc.).

In most of the applications of the FEM to computational micromechanics, the phase arrangements are discretized using ‘standard’ continuum elements, the mesh being designed in such a way that element boundaries (and, where

required, special interface elements) are located at all interfaces between constituents. This approach can model any microgeometry, and it is readily available in commercial codes, but the need to resolve the microfields in regions with large strain gradients (for instance, between closely packed reinforcement particles) may lead to very large models.

An alternative approach is given by the Voronoi Cell FEM, which is especially suited to analyze material with matrix inclusion-topology (Moorthy and Ghosh, 1996). The number of degrees of freedom in the analysis is drastically reduced through the use of a special class of hybrid finite elements which are specifically formulated to model the deformation, stress and strain fields in a heterogeneous region consisting of a single inclusion or void together with the surrounding matrix on the basis of some appropriate analytical theory. The mesh for the hybrid elements is obtained by Voronoi tesselations based on the positions of the reinforcements, where each Voronoi cell contains an inclusion embedded in the matrix and stands for a basic hybrid finite element (Fig. 5.3). Large planar multi-inclusion arrangements in the elasto-plastic regime have been analyzed with this methodology (Ghosh *et al.*, 1997), and further developments included damage by fracture of the inclusions. The extension to three-dimensional microstructures has been reported recently (Ghosh and Moorthy, 2004).

Another strategy for the discretization uses a regular three-dimensional array of hexahedral elements or voxels. Each element is assigned to one phase according to the information provided by the three-dimensional reconstruction of the microstructure obtained by serial sectioning or X-ray microtomography (Guldberg *et al.*, 1998; Geandier *et al.*, 2003). Of course, the ragged phase boundaries induced by the cubic discretization can lead to very high local maxima, and good spatial resolution, i.e. a high number of voxels, is necessary for the accuracy of the results, leading sometimes to very large models.



5.3 Two-dimensional Voronoi cell finite-element mesh and detail of one Voronoi finite element.

Finally, it is important to notice that computational micromechanics simulations often involve very large models, with complex non-linear behavior in the phases, so one area of concern is the development of more efficient tools to solve the boundary value problem. This has led to the use of iterative solvers, such as the conjugate gradient method, in which an analytical solution of the stress microfields is given as the starting point to speed up the convergence. Another approach is the domain decomposition method, an iterative strategy which takes advantage of supercomputers based on multiple independent processors (Zohdi and Wriggers, 2001). The model is split into smaller subdomains, whose behavior is solved separately in different computers using approximate boundary conditions. The solutions of all subdomains are assembled and new boundary conditions are assigned to each one. This process is repeated iteratively until continuity in stresses and displacements at the surfaces shared by two subdomains is achieved. More details about the current numerical strategies in the area of computational micromechanics can be found in Zohdi and Wriggers (2005).

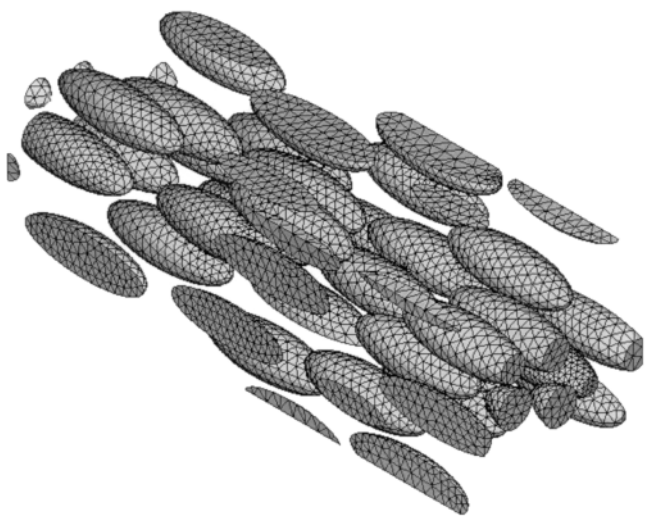
5.4.2 Boundary conditions

The effective properties of an RVE of finite size subjected to far-field homogeneous deformation depend on the actual boundary conditions. For instance, the elastic constants derived under imposed displacements are always higher than those obtained under imposed forces, the differences decreasing as the size of the RVE increases and both converging to the actual values in the limit of an RVE of infinite size. Both types of boundary conditions can be applied to an RVE (whose size is larger than the critical one) to estimate upper and lower bounds for the elastic properties of the heterogeneous material (Zohdi and Wriggers, 2001). Intermediate values for the macroscopic behavior between those provided by imposed displacements or forces can be obtained with the embedded cell method (Dong and Schmauder, 1996): the RVE is embedded in a large cell of a homogeneous medium, which stands for the action of the rest of the material on the RVE, and whose properties are the solution sought. An initial estimation (as given, for instance, by the self-consistent scheme) is used to solve the boundary value problem and to compute the effective behavior of the RVE. This result is used as the new estimation for the effective properties of the homogeneous medium and the process is repeated until convergence is achieved.

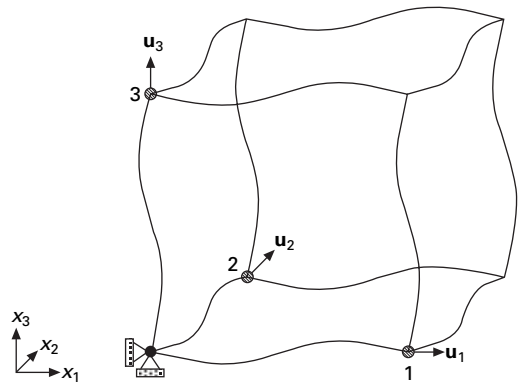
Huet and coworkers (Huet, 1990; Hazanov and Huet, 1994) demonstrated that the effective elastic constants derived under periodic boundary conditions are always bounded by those obtained under imposed forces or displacements, and thus provide a better approximation of the actual effective properties of the heterogeneous material for a given size of the RVE, which is normally limited by the available computational resources. Computational

micromechanics simulations carried out with periodic boundary conditions also require periodicity in the microstructure, and this enabled the approximation of the heterogeneous material by an indefinite extension of a periodic RVE in the three dimensions of space. Simple RVE (made up of one or two features) are able to simulate regular microstructures with bcc, fcc, hcp, etc. symmetries, while materials with random microstructures can be studied by means of RVE containing a random dispersion of the relevant features in the microstructure (grains, inclusions, fibers, etc.). An example of a periodic RVE corresponding to a composite material made up of random and homogeneous dispersion of aligned ellipsoidal inclusions is plotted in Fig. 5.4. The prismatic RVE (with an aspect ratio of 3, equal to the one of the ellipsoids) was built from the computer-generated ellipsoid distribution by splitting the ellipsoids intersecting the prism faces into the appropriate number of parts which were copied to the opposite faces, leading to the 'ellipsoids in box' prism in Fig. 5.4.

Periodic RVE fill the space by translation along three perpendicular axes, and periodic boundary conditions have to ensure that opposite faces of the RVE must fit each other like parts of a jigsaw puzzle in both undeformed and deformed situations. This condition is shown graphically in Fig. 5.5 for a cubic RVE, and it is enforced numerically by imposing the same finite element discretization on opposite faces of the RVE and linking the corresponding degrees of freedom within each pair of faces. In the case of the cubic RVE



5.4 Prismatic periodic RVE of a random and homogeneous dispersion of aligned ellipsoids. The volume fraction of ellipsoids is 25 %. The finite-element discretization is shown on the surface of the ellipsoidal inclusions (Pierard *et al.*, 2007).



5.5 Deformation of a cubic RVE with periodic boundary conditions.

in Fig. 5.5 of volume L^3, if three concurrent edges of the cube stand for the axes of coordinates x_1, x_2 and x_3, the periodic boundary conditions can be expressed as a function of the displacement vector $\mathbf{u}$ as

$$\begin{aligned}
&\mathbf{u}(x_1, x_2, 0) - \mathbf{u}_3 = \mathbf{u}(x_1, x_2, L)\\
&\mathbf{u}(x_1, 0, x_3) - \mathbf{u}_2 = \mathbf{u}(x_1, L, x_3)\\
&\mathbf{u}(0, x_2, x_3) - \mathbf{u}_1 = \mathbf{u}(L, x_1, x_2)
\end{aligned} \qquad [5.13]$$

where $\mathbf{u}_1$, $\mathbf{u}_2$ and $\mathbf{u}_3$ are the displacements of the master nodes. Of course, periodic boundary conditions are more expensive than symmetric ones in terms of computing time and memory requirements because the linking of paired degrees of freedom in opposed faces degrades the band structure of the stiffness matrix. However, symmetric boundary conditions only allow deformation patterns which are compatible with the imposed symmetry while any homogeneous deformation state can be imposed on the periodic RVE by choosing adequately the displacement of the master nodes on each face.

5.4.3 Constitutive equations and damage

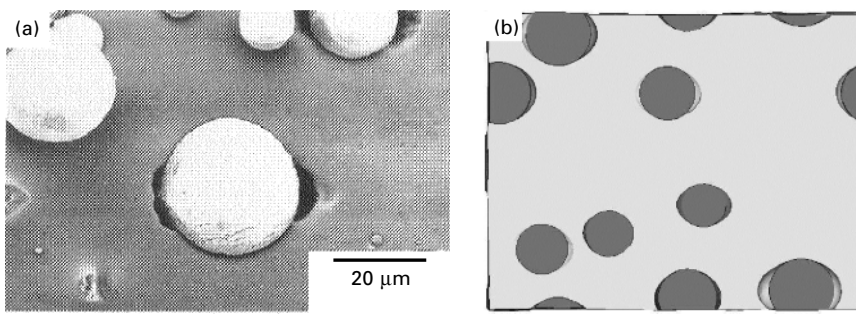
The behavior of the different phases and interfaces in the RVE is a critical factor to determine accurately the macroscopic properties. There is a wide range of constitutive equations available to simulate the thermo-mechanical properties of bulk materials within the framework of the FEM, and very often they can be directly used in computational micromechanics simulations. Nevertheless, the material within the RVE may be very different from that found in bulk samples, and it is sometimes necessary to use more sophisticated models. This includes, for instance, strain gradient plasticity theories to account for size effects in plastic deformation on the μm scale (Hutchinson, 2000) or crystal plasticity models to include the effects of textures in the

behavior of single crystal and polycrystalline metals (Delannay *et al.*, 2006).

In addition, the onset and propagation of damage within the RVE can be simulated by using the appropriate fracture models for the different phases and interfaces. In this context, failure by interface decohesion in a metal–ceramic composite was analyzed by means of interface elements at the particle/matrix interface (Segurado and LLorca, 2005). The interface behavior was simulated by a cohesive crack model with two parameters: the interfacial strength and the interfacial toughness. The patterns of interface decohesion predicted by the numerical model at the microscopic level were in good agreement with experimental findings (Fig. 5.6), and the simulations also revealed the effect of the interface properties (strength and toughness) on the effective tensile properties of the composite. Following this line of research, other authors have studied the influence of ductile matrix failure (LLorca and Segurado, 2004; Drabek and Böhm, 2005) and particle fracture (Eckschlager *et al.*, 2002; Segurado and LLorca, 2002) on the effective properties of heterogeneous materials.

5.5 Multiscale coupling

The techniques based on homogenization methods and the numerical analysis of an RVE can be coupled to carry out multiscale simulations of the behavior of materials within the framework of continuum micromechanics. The most important developments in this area are briefly noted below, together with significant examples which show the potential of these methods to bridge the length scales in the analysis of the mechanical performance of heterogeneous materials.



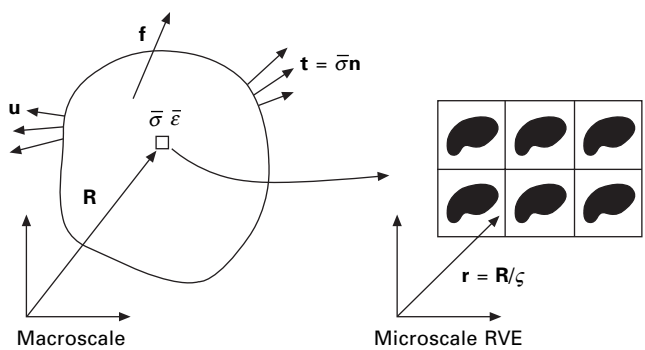
5.6 Interface decohesion in an 6061 Al alloy reinforced with alumina spheres. (a) Experimental observations. (b) Numerical simulations of an RVE containing a random and homogeneous dispersion of elastic spheres embedded in an elasto-plastic matrix (Segurado and LLorca, 2005).

5.5.1 Asymptotic homogenization and finite elements

One of the first multiscale methods was based on the assumptions of multiscale homogenization techniques or multiple scale expansion methods. Detailed theoretical developments can be found in Guedes and Kikuchi (1990) and Fish and co-workers (Fish *et al*., 1997; Wentofr *et al*., 1998; Fish and Shek, 2000; Fish and Yu, 2001) and only the main assumptions and results will be presented here. Basically, the mechanical behavior is studied on the structure and material microstructure length scales. The former is computed by FEM while the latter is determined through an approximate solution for the displacement field and assuming that the microstructure of the heterogeneous material is periodic, as shown in the RVE depicted in Fig. 5.7. Let $\mathbf{R}$ and $\mathbf{r} = \mathbf{R}/\varsigma$ be the position vectors at the macro and microscale, respectively, and $\varsigma << 1$ a small positive number that relates the length scales at the micro and macro levels. The mathematical asymptotic homogenization theory is used to solve a multiscale problem by coupling finite element problems on several length scales. Any macroscopic variables that exhibit explicit dependence on both micro and macro position vectors, such as the macroscopic displacement field on the heterogeneous material, can be asymptotically expanded in the following way

$$\mathbf{u}\left(\mathbf{R}, \mathbf{r} = \frac{\mathbf{R}}{\varsigma}\right) = \mathbf{u}^0(\mathbf{R}) + \varsigma \mathbf{u}^1(\mathbf{R}, \mathbf{r}) + O^2 \qquad [5.14]$$

where $\mathbf{u}^0(\mathbf{R})$ represents the macroscopic displacement field, and $\mathbf{u}^1(\mathbf{R}, \mathbf{r})$ the periodically first-order fluctuations or perturbations due to the microstructure. They are usually known as slow and fast variables, respectively. This asymptotic expansion can be considered an approximate solution to the boundary problem when the ratio between length scales ς is sufficiently small. In that case, $\mathbf{u}(\mathbf{R}, \mathbf{r})$ displays a slow global variation in $\mathbf{R}$ and a fast variation in the vicinity ς



5.7 Schematic strategy of the multiscale simulation based on the finite element method and asymptotic homogenization.

around **R**. Taking into account the mathematical chain rule for spatial differentiation in connection with the small displacement theory framework, the strain field is computed as

$$\boldsymbol{\epsilon}(\mathbf{R}, \mathbf{r}) = \overline{\text{grad}}^{s}\,\mathbf{u} = \text{Grad}^{s}\,\mathbf{u} + \frac{1}{\varsigma}\text{grad}^{s}\,\mathbf{u}$$
$$= \text{Grad}^{s}\,\mathbf{u}^{0} + \text{grad}^{s}\,\mathbf{u}^{1} + \varsigma\,\text{grad}^{s}\,\vec{\mathbf{u}}^{1} + \ldots \quad [5.15]$$

where $\overline{\text{grad}}^{s}\,\mathbf{u} = \text{Grad}^{s}\,\mathbf{u} + \varsigma^{-1}\text{grad}^{s}\,\mathbf{u}$ is the composed gradient operator and $\text{Grad}^{s}\,\mathbf{u}$ and $\text{grad}^{s}\,\mathbf{u}$ stand for the field gradients with respect to the macroscopic and microscopic position vectors, respectively. The asymptotic expansion of the stress field can be computed assuming linear elastic behavior with the constitutive equation in analogy with the former equation leading to

$$\boldsymbol{\sigma}(\mathbf{R}, \mathbf{r}) = \mathbf{L} : \boldsymbol{\epsilon}(\mathbf{R}, \mathbf{r}) = \mathbf{L} : \overline{\text{grad}}^{s}\,\mathbf{u}$$
$$= \boldsymbol{\sigma}^{1}(\mathbf{R}, \mathbf{r}) + \varsigma\boldsymbol{\sigma}^{2}(\mathbf{R}, \mathbf{r}) + \ldots \quad [5.16]$$

where **L** is the local fourth-order stiffness tensor, $\boldsymbol{\sigma}^{1} = \mathbf{L} : (\text{Grad}^{s}\mathbf{u}^{0} + \text{grad}^{s}\mathbf{u}^{1})$ and $\boldsymbol{\sigma}^{2} = \mathbf{L} : \text{Grad}^{s}\mathbf{u}^{1}$. Finally, neglecting ς^{2} and higher-order terms in equation [5.14] and including **f** as the macroscopic body forces vector, the following hierarchical Navier microscopic and macroscopic differential equations can be obtained

$$\text{div}\boldsymbol{\sigma}^{1} = \text{div}[\mathbf{L} : (\text{Grad}^{s}\mathbf{u}^{0} + \text{grad}^{s}\mathbf{u}^{1})] = \mathbf{0} \quad [5.17]$$

$$\text{Div}\boldsymbol{\sigma}^{1} + \text{div}\boldsymbol{\sigma}^{2} + \mathbf{f} = \mathbf{0} \quad [5.18]$$

where $\overline{\text{div}}\boldsymbol{\sigma} = \text{Div}\boldsymbol{\sigma} + \varsigma^{-1}\text{div}\boldsymbol{\sigma}$ is the divergence differential operator obtained again from the chain derivative rule. Averaging equation [5.18] over the volume of Ω leads to the macroscopic equilibrium equation given by

$$\text{Div}\left[\frac{1}{\Omega}\int_{\Omega}\boldsymbol{\sigma}^{1}\text{d}\Omega\right] + \mathbf{f} = \text{Div}\overline{\boldsymbol{\sigma}}^{1} + \mathbf{f} = \mathbf{0} \quad [5.19]$$

as the average value of $\text{div}\boldsymbol{\sigma}^{2}$ disappears due to periodicity of the microstructure at the local level. Finally, microscopic equilibrium equation [5.17] can be decoupled assuming linearity and variable separation for the fast variable $\mathbf{u}^{1}(\mathbf{R}, \mathbf{r}) = \boldsymbol{\chi}(\mathbf{r})\text{Grad}^{s}\mathbf{u}^{0}$ where $\boldsymbol{\chi}(\mathbf{r})$ is an unknown periodic function in Ω which represents the characteristic deformation modes in the RVE. After some manipulation algebra, the homogenized fourth-order stiffness tensor, $\overline{\mathbf{L}}$, can be computed from the mean value of $\overline{\boldsymbol{\sigma}}^{1}$ as

$$\overline{\boldsymbol{\sigma}} = \frac{1}{\Omega}\int_{\Omega}\boldsymbol{\sigma}^{1}\text{d}\Omega$$
$$= \left[\frac{1}{\Omega}\int_{\Omega}\mathbf{L} : \left(\mathbf{I} + \frac{\partial\boldsymbol{\chi}(\mathbf{r})}{\partial\mathbf{r}}\right)\text{d}\Omega\right] : \text{Grad}^{s}\,\mathbf{u}^{0} = \overline{\mathbf{L}} : \text{Grad}^{s}\,\mathbf{u}^{0} \quad [5.20]$$

Each mathematical model on the micro and macroscale can be solved separately using the finite element method. Firstly, $\chi(\mathbf{r})$ is computed by solving the equilibrium equation on the microscopic scale [5.17], and afterwards $\overline{\mathbf{L}}$ can be determined from equation [5.20] and used as the homogenized constitutive stiffness to solve the Navier equations at the macroscopic level [5.15]. Analogous expressions for the homogenized constitutive equation were proposed to incorporate material non-linear effects and damage (Fish *et al*., 1997; Lee *et al*., 1999; Fish and Yu, 2001).

It is important to notice that the scale parameter ς does not appear explicitly in the equilibrium equation at both levels, but the multiscale simulations carried out within this framework are restricted to very small values of ς, when the macroscopic fields are practically constant at the microscopic length scale. These conditions are not fulfilled in presence of high stress and strain gradients (in presence of cracks, corners, stress singularities, etc., …), and more recent developments in multiscale modeling have been directed to accounting for non-uniform macroscopic deformation fields within the microstructural RVE (Fish and Shek, 2000; Kouznetsova *et al*., 2002).

Relevant examples of multiscale modeling using asymptotic homogenization and finite elements can be found in Fish and Shek (2000). These authors studied the behavior of a large structure, a composite diffuser casing, made up of a woven fiber fabric embedded in a non-linear elasto-plastic matrix which obeys the von Mises yield criterion. Three different scales were coupled in the study (structural, local and material) which was solved using an efficient finite element multigrid solver.

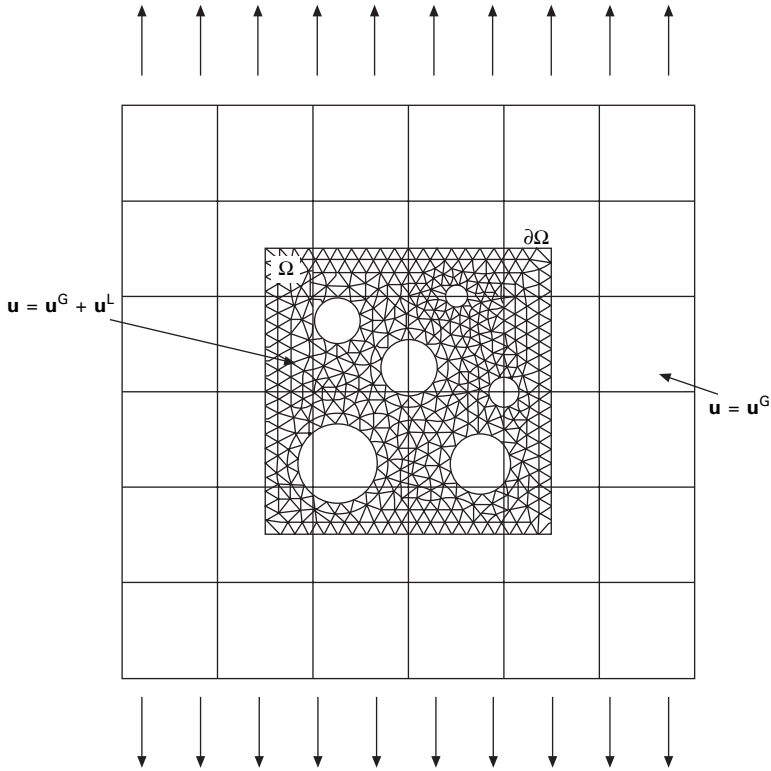
5.5.2 Mesh superposition methods

Another way to solve multiscale problems is the mesh superposition method, which was initially introduced by Fish (1992) to study three-dimensional phenomena in heterogeneous materials, such as the interlaminar stresses that appear in laminated composites. In this method, an additional mesh is superimposed on a global or macroscopic mesh in those regions where detailed information of the stress and strain fields is necessary at the microstructural level, Fig. 5.8. A composed hierarchical displacement field $\mathbf{u}$ can be split into a global displacement $\mathbf{u}^{\mathrm{G}}$ and a local displacement $\mathbf{u}^{\mathrm{L}}$ in the regions of interest

$$\mathbf{u} = \mathbf{u}^{\mathrm{G}} + \mathbf{u}^{\mathrm{L}} \text{ in the local region and} \qquad [5.21]$$

$$\mathbf{u} = \mathbf{u}^{\mathrm{G}} \text{ in the global region.} \qquad [5.22]$$

$\mathbf{u}^{\mathrm{L}}$ can be viewed as a relative displacement of the local mesh with respect to the global mesh or a kinematic enhancement of the displacement field. A necessary condition is to enforce homogeneous displacement compatibility



5.8 Schematic strategy of the mesh superposition technique.

at those points lying on the interface between the meshes, $\mathbf{u}^{L} = \mathbf{0}$ at $\partial\Omega$. The main advantage of this method is the ability to model accurately the heterogeneous material microstructure in regions with high gradients (crack tips, interfaces, etc.) while the rest of the structure is analyzed by coupling the FEM with asymptotic homogenization techniques.

Derivation of the Navier equations from the total displacement field $\mathbf{u} = \mathbf{u}^{G} + \mathbf{u}^{L}$ and the implementation using the weak formulation and the virtual displacement principle is straightforward. This leads to the following system of equilibrium equations

$$\begin{bmatrix} \mathbf{K}^{G} & \mathbf{K}^{GL} \\ \mathbf{K}^{GL} & \mathbf{K}^{L} \end{bmatrix} \begin{pmatrix} \mathbf{u}^{G} \\ \mathbf{u}^{L} \end{pmatrix} = \begin{pmatrix} \mathbf{f} \\ \mathbf{0} \end{pmatrix} \qquad [5.23]$$

where $\mathbf{K}^{L}$ and $\mathbf{K}^{G}$ stand for the stiffness matrices that can be computed from the local and global meshes and $\mathbf{K}^{GL}$ is the stiffness matrix that couples the displacement degrees of freedom of the local and global meshes, and $\mathbf{f}$ is the nodal forces vector.

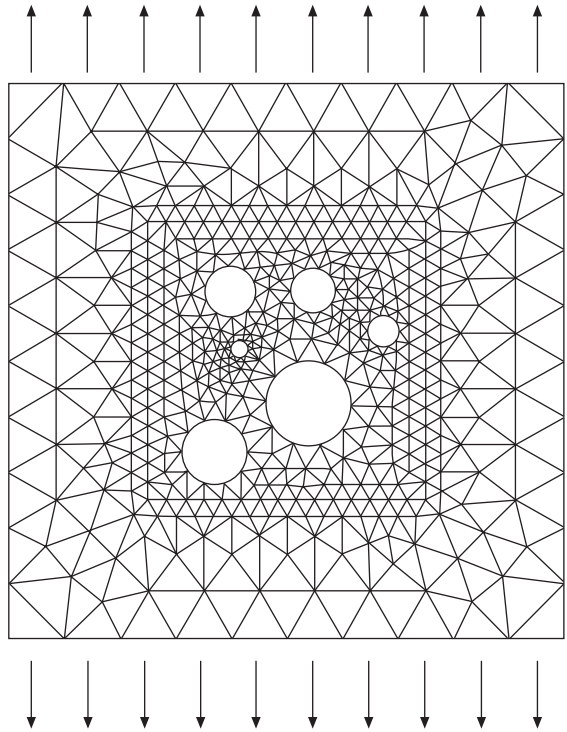
An efficient implementation of the former equations in a finite element code has to take account of several aspects. Firstly, the data structure must be manipulated or renumbered in order to decrease the amount of physical space allocated during the computations. Secondly, the classical sparse structure of the stiffness matrix is destroyed by the presence of the $\mathbf{K}^{GL}$ coupling terms, and efficient iterative solvers must be used instead of direct solvers (e.g. Gauss–Seidel). Detail of additional computational strategies to solve such a system of equations can be found in Fish and Shek (2000).

Takano *et al*. (1996) studied the mechanical behavior of textile composite materials using mesh superposition techniques coupled with damage mechanics to evaluate strength at the mesoscale, while Takano and Okuno (2004) applied this strategy with the asymptotic homogenization theory to study interface cracks in coated materials. Finally, Kawagai *et al*. (2006) used an automatic image-based modeling and finite element mesh superposition technique to study the behavior near the crack tips and/or the interfaces of porous solids.

5.5.3 Embedded cell methods

Embedded cell methods are conceptually similar to mesh superposition techniques. Basically, the model includes a detailed representation of the microstructure in those regions in which it is necessary to have information of the stress and strain fields at the local level. These core regions are embedded within the homogeneous medium, which stands for the homogenized description of the heterogeneous material. The finite element discretization in the core regions is much finer, and is able to capture the fast variations in the fields which occur as a result of the presence of sharp interfaces, cracks, damage, etc. The homogenized material is discretized with a coarse mesh (not superposed on the finer one, as in the mesh superposition method) which only has to capture the slow variations in the fields, Fig. 5.9. The coarse mesh transmits the macroscopic far fields to the core by enforcing the displacement continuity at the interface between the core and the embedding region. Of course, the constitutive equation of the material in the embedding region has to be in accordance with the homogenized behavior of the core region, and it is normally computed by using a suitable homogenization model. Nevertheless, the difference in the constitutive equations between the core and the embedding region leads to the formation of spurious boundary layers in which the stress and strain fields in the core region are affected by the jump in stiffness. The thickness of these boundary layers for elastic materials is of the order of the microstructural details in the core region (particle or grain size) but may be larger in non-linear problems.

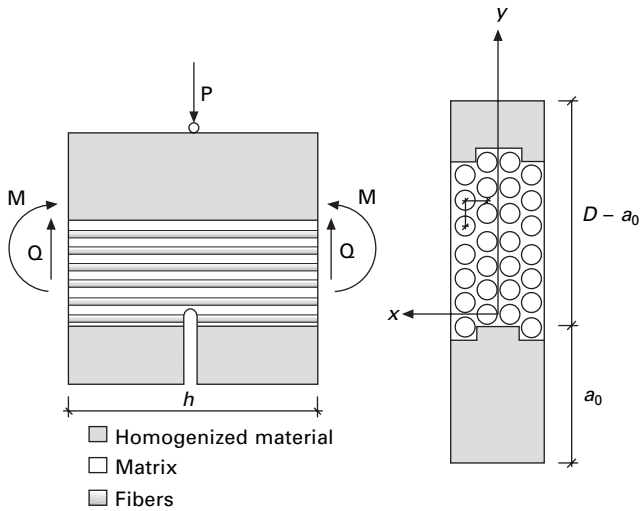
Embedded cell approaches are very useful to simulate fracture because the processes which control the mechanical behavior occur in a very localized region near the crack tip, while the rest of the sample can be easily (and



5.9 Schematic strategy of the embedded cell methodology.

accurately) represented by a homogeneous material. Due to the limitations in computational power, embedded cell approaches to simulate fracture in composite materials were initially limited to two dimensions (Wulf *et al*., 1996; Boselli *et al*., 2001) and they could only provide qualitative information about the interaction of the crack path with the microstructure. More recent analyses have extended these results to three dimensions, providing quantitative results of the influence of the microstructure (properties and spatial arrangement of the phases) on the actual fracture toughness of the material.

An example of this methodology is presented in González and LLorca (2006), which simulated the fracture behavior of a fiber-reinforced composite beam in presence of a notch perpendicular to the fibers by means of a multiscale model based on an embedded cell approach in three dimensions. The beam representation included two zones: one core region around the notch tip, where full details of the composite microstructure (including the matrix, reinforcements and interfaces) were resolved, and another one surrounding this region where the composite was represented by a linear thermo-elastic, transversally-isotropic homogeneous solid (Fig. 5.10). The damage and fracture micromechanisms which controlled the onset of fracture

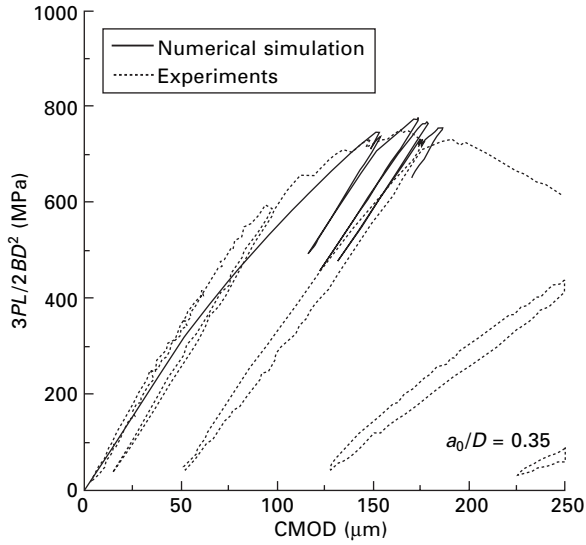


5.10 Embedded cell model to simulate fracture of a fiber-reinforced composite notched beam subjected to three-point bending. *P* is the applied load and *M* and *Q* stand for the corresponding bending moment and shear forces.

(namely, plastic deformation of the matrix, brittle fiber fracture and fiber/matrix frictional sliding) were included in the behavior of the different phases and interfaces around the notch root, and the corresponding micromechanical parameters governing their behavior were measured independently. The mechanical response of the beams was computed through the finite element method, and it was found that the multiscale computational model reproduced – within the available experimental and numerical data – the main deformation and failure micromechanisms. Moreover, the experimental load–CMOD curves (Fig. 5.11) were accurately reproduced by the multiscale model up to the maximum load, which demonstrates the capability of this strategy to simulate the fracture behavior of complex, heterogeneous materials.

5.6 Future directions

Continuum micromechanics is a very active research area because it provides a unique tool with which to establish the relationship between the macroscopic behavior of materials and the microstructural features on the μm length scale. Nevertheless, significant developments are necessary in all the areas outlined in this chapter to attain the final goal of predicting the macroscopic behavior of materials from the properties and spatial distribution of the different phases and interfaces at the microscopic level. They include accurate algorithms to generate realistic microstructures of RVE from the information



5.11 Load *P* (normalized by the beam span, *L*, thickness, *B*, and depth, *D*) vs crack mouth opening displacement of the fiber-reinforced composite notched beam tested in three-point bending.

provided by high-order statistical correlation functions, the extension of homogenization methods to finite deformations and to phases and interphases which present a highly non-linear behavior, and the development of faster and more efficient numerical strategies to analyze larger RVEs, among many others. Multiscale simulations of materials and structures will benefit from all these advances.

5.7 Acknowledgments

This investigation was supported by the Spanish Ministry of Education and Science through grants MAT 2003-5202-C02 and MAT 2006-2602 and by the Comunidad de Madrid through the program ESTRUMAT-CM (reference MAT-0077).

5.8 References

Benveniste Y. (1987), 'A new approach to the application of Mori-Tanaka's theory in composite materials', *Mech. Mater.* **6**, 147–157.

Brockeborough J.R., Suresh S. and Wienecke H.A. (1991), 'Deformation of metal-matrix composites with continuous fibers: geometrical effects of fiber distribution and shape', *Acta Metall. Mater.* **39**, 735–752.

Boselli J., Pitcher P.D., Gregson P.J. and Sinclair I. (2001), 'Numerical modelling of particle distribution effects on fatigue of Al-SiCp composites', *Mater. Sci. Engng.* **A300**, 113–124.

Chaboche J.L., Kanouté P. and Roos A. (2005), 'On the capabilities of mean-field approaches for the description of plasticity in metal matrix composites', *Int. J. Plasticity* **21**, 1409–1434.

Delannay L., Jacques P.J. and Kalidini S.R. (2006), 'Finite element modeling of crystal plasticity with grains shaped as truncated octahedrons', *Int. J. Plasticity* **22**, 1879–1898.

Dong M. and Schmauder S. (1996), 'Modeling of metal-matrix composites by a self-consistent embedded cell model', *Acta Mater*. **44**, 2465–2478.

Drabek T. and Böhm H.J. (2005), Micromechanical finite element analysis of metal matrix composites using nonlocal failure modes, *Comp. Mater. Sci*. **32**, 329–336.

Drugan W.J. and Willis J.R. (1996), 'A micromechanics-based nonlocal constitutive equation and estimates of representative volume element size for elastic composites', *J. Mech. Phys. Solids* **44**, 497–524.

Eckschlager A., Han W. and Böhm H.J. (2002), 'A unit cell model for brittle fracture of particles embedded in a ductile', *Comput. Mater. Sci*. **25**, 85–91.

Eshelby J.D. (1957), 'The determination of the elastic field of an ellipsoidal inclusion and related problems', *Proc. Roy. Soc. London* **A241**, 376–396.

Estevez R., Maire E., Franciosi P. and Wilkinson, D.S. (1999), 'Effect of particle clustering on the strengthening versus damage rivalry in particulate reinforced elastic plastic materials: A 3-D analysis from a self-consistent modeling', *Eur. J. Mech. A/Solids* **18**, 785–804.

Fish J. and Markolefas S. (1992), 'The s-version of the finite element method for multilayer laminates', *Int. J. Num. Methods Eng*. **33**, 1081–1105.

Fish J. and Shek K. (2000), 'Multiscale analysis of composite materials and structures', *Comp. Sci. Techno*. **60**, 2547–2556.

Fish J. and Yu Q. (2001), 'Multiscale damage modeling for composite materials: theory and computational framework', *Int. J. Numer. Meth. Eng*. **52**, 161–191.

Fish J., Shek K., Pandheeradi M. and Shephard M.S. (1997), 'Computational plasticity for composite structures based on mathematical homogenization: theory and practice', *Comput. Methods Appl. Mech. Eng*. **148**, 53–73.

Geandier G., Hazotte A., Denis S., Mocellin A. and Maire E. (2003). Microstructural analysis of alumina chromium composites by X-ray tomography and 3-D finite element simulation of thermal stresses', *Scripta Mater*. **48**, 1219–1224.

Ghosh S. and Moorthy S. (2004), 'Three dimensional Voronoi cell finite element model for microstructures with ellipsoidal heterogeneities', *Comput. Mech*. **34**, 510–531.

Ghosh S., Novak Z. and Lee K. (1997), 'Quantitative characterization and modeling of composite microstructures by Voronoi cells', *Acta Metall. Mater*. **46**, 2215–2234.

González C. and LLorca J. (2000), 'A self-consistent approach to the elasto-plastic behavior of two-phase materials including damage', *J. Mech. Phys. Solids*. **48**, 675–692.

González C. and LLorca J. (2006), 'Multiscale modeling of fracture in fiber reinforced composites', *Acta Mater*. **54**, 4171–4181.

González C., Segurado J. and LLorca J. (2004), 'Numerical simulation of elasto-plastic deformation of composites: evolution of stress microfields and implications for homogenization models, *J. Mech. Phys. Solids* **48**, 1573–1593.

Guedes J.M. and Kikuchi N. (1990), 'Preprocessing and postprocessing for materials based on the homogenization method with adaptative finite element methods', *Comput. Methods Appl. Mech. Eng*. **83**, 143–198.

Guldberg R.E., Hollister S.J. and Charras G.T. (1998), 'The accuracy of digital image-based finite element models', *J. Biomech. Eng*. **120**, 289–295.

Gusev A. (1997), 'Representative volume element size for elastic composites: a numerical study', *J. Mech. Phys. Solids* **45**, 1449–1459.

Hashin Z. and Shtrikman S. (1963), 'A variational approach to the theory of elastic behavior of multiphase materials', *J. Mech. Phys. Solids* **11**, 127–140.

Hazanov S. and Huet C. (1994), 'Order relationships for boundary condition effects in heterogeneous bodies smaller than the representative volume', *J. Mech. Phys. Solids* **42**, 1995–2011.

Hill R. (1963), 'Elastic properties of reinforced solids: some theoretical principles', *J. Mech. Phys. Solids* **11**, 357–372.

Hill R. (1965), 'A self-consistent mechanics of composite materials', *J. Mech. Phys. Solids* **13**, 213–222.

Hine P.J., Lusti H.R. and Gusev A.A. (2002), 'Numerical simulation of the effects of volume fraction, aspect ratio and fibre length distribution on the elastic and thermoelastic properties of short fibre composites', *Comp. Sci. Techno*. **62**, 1445–1453.

Huet C. (1990), 'Application of variational concepts to size effects in elastic heterogeneous bodies', *J. Mech. Phys. Solids*. **8**, 813–841.

Hutchinson J.W. (2000), 'Plasticity at the micron scale', *Int. J. Solids Struct*., **37**, 225–238.

Ingber M.S. and Papathanasiou T. (1997), 'A parallel supercomputing investigation of the stiffness of aligned short fiber reinforced composites using the boundary element method', *Int. J. Numer. Meth. Eng*. **40**, 3477–3491.

Kawagai M., Sando A. and Takano N. (2006), 'Image-based multi-scale modelling strategy for complex and heterogeneous porous microstructures by mesh superposition method', *Model. Simul. Mater. Sci. Eng*. **14**, 53–69.

Kouznetsova V., Geers M.G.D. and Brekelmans W.A.M. (2002), 'Multi-scale constitutive modeling of heterogeneous materials with a gradient-enhanced computational homogenization scheme', *Int. J. Numer. Meth. Eng*. **54**, 1235–1260.

Kröner E. (1958), 'Berechnung der elastischen Konstanten des Vielkristalls aus den Konstanten der Einkristalls', *Z. Physik* **151**, 504–518.

Lee K., Moorthy S. and Ghosh S. (1999), 'Multiple scale computational model for damage in composite materials' *Comput. Methods Appl. Mech. Eng*. **172**, 175–201.

LLorca J. and Segurado J. (2004), 'Three-dimensional multiparticle cell simulations of deformation and damage in sphere-reinforced composites', *Mater. Sci. Eng.* **A365**, 267–274.

LLorca J., Needleman A. and Suresh S. (1991), 'An analysis of the effects of matrix void growth on deformation and ductility in metal-ceramic composites', *Acta Metall. Mater*. **39**, 2317–2335.

Monetto I. and Drugan W.J. (2004), 'A micromechanics-based nonlocal constitutive equation for elastic composites containing randomly oriented spheroidal heterogeneities', *J. Mech. Phys. Solids* **52**, 359–393.

Moorthy S. and Ghosh S. (1996), 'A model for analysis of arbitrary composite and porous microstructures with Voronoi cell finite elements', *Int. J. Numer. Methods Eng*. **39**, 2363–2398.

Mori T. and Tanaka K. (1973), 'Average stress in the matrix and average elastic energy of materials with misfitting inclusions', *Acta Metall. Mater*. **21**, 571–574.

Moulinec H. and Suquet P. (1994), 'A fast numerical method for computing the linear and nonlinear properties of composites', *C. R. Acad. Sci. Paris II* **318**, 1417–1423.

Nemat-Nasser S. and Hori M. (1999), *Micromechanics. Overall Properties of Heterogeneous Materials*, Amsterdam: North-Holland.

Pierard O., Friebel C. and Doghri I. (2004), 'Mean-field homogenization of multi-phase thermo-elastic composites: a general framework and its validation', *Comp. Sci. Tech.* **64**, 1587–1603.

Pierard O., González C., Segurado J., LLorca J. and Doghri I. (2007), 'Micromechanics of elasto-plastic materials reinforced with ellipsoidal inclusions', *Int. J. Solids Struct.*, **44**. In press.

Ponte Castañeda P. and Suquet P. (1998), 'Nonlinear composites,' *Adv. Appl. Mech.* **34**, 171–301.

Segurado J. and LLorca J. (2002), 'A numerical approximation to the elastic properties of sphere-reinforced composites', *J. Mech. Phys. Solids* **50**, 2107–2121.

Segurado J. and LLorca J. (2005), 'A computational micromechanics study of the effect of interface decohesion on the mechanical behavior of composites', *Acta Mater.* **53**, 4931–4942.

Segurado J., González C. and LLorca J. (2003), 'A numerical investigation of the effect of particle clustering on the mechanical properties of composites', *Acta Mater.* **51**, 2355–2369.

Sidhu R.S. and Chawla N. (2006), 'Three-dimensional (3D) visualization and microstructure-based modeling of deformation in a Sn-rich solder', *Scripta Mater.* **54**, 1627–1631.

Takano N. and Okuno Y. (2004), 'Three-scale finite element analysis of heterogeneous media by asymptotic homogenization and mesh superposition methods', *Int. J. Solids Struct.* 2004, 4121–4135.

Takano N., Uetsuji Y., Kashiwagi Y. and Zako M. (1996), 'Hierarchical modelling of textile composite materials and structures by the homogenization method', *Model. Simul. Mater. Sci. Eng.* **7**, 207–231.

Tewari A., Gokhale A.M., Spowart J.E. and Miracle D.B. (2004), 'Quantitative characterization of spatial clustering in three-dimensional microstructures using two-point correlation functions', *Acta Mater.* **52**, 307–319.

Torquato S. (2001), *Random Heterogeneous Materials*. New York: Springer.

Wentofr R., Collar R., Shephard M.S. and Fish J. (1998), 'Automatic modeling for complex woven mesostructures', *Comput. Methods Appl. Mech. Eng.* **172**, 273–291.

Wulf J., Steinkopff T. and Fishmeister H. (1996), 'FE-simulation of crack paths in the real microstructure of an Al(6061)/SiC composite' *Acta Mater.* **44**, 1765–1779.

Youssef S., Maire E. and Gaertner R. (2005), 'Finite element modelling of the actual structure of cellular materials determined by X-ray', *Acta Mater.* **53**, 719–730.

Zohdi T.I. and Wriggers P. (2001), 'Aspects of the computational testing of the mechanical properties of heterogeneous material samples', *Int. J. Numer. Method Eng.* **50**, 2573–2599.

Zohdi T.I. and Wriggers P. (2005), *Introduction to Computational Micromechanics*, Berlin: Springer.

6

Grain–continuum modelling of material behaviour

MAX O. BLOOMFIELD and TIMOTHY S. CALE,
Rensselaer Polytechnic Institute, USA

6.1 Introduction

6.1.1 Overview

This chapter summarizes how we have approached introducing microstructure, or 'granularity', into microelectronics processing and reliability modeling and simulation. We think it is a good approach to consider for introducing 'grain-focused' modeling and simulation into other application domains as well. As discussed herein, we think grain-focused modeling will play an important role; it complements continuum and discrete/atomistic modeling efforts and is particularly relevant to *multiscale modeling* studies. The material in the chapter is somewhat parochial, in that the majority of the background references are to our efforts. Other groups have done similar work, e.g. in process simulation and in including microstructure in models, but our current grain-focused work grew out of our experiences and modeling efforts in microelectronics relevant modeling, so our cited background is more directly relevant to the material presented.

This section motivates the chapter and provides some microelectronics context. Section 6.2 discusses some approaches to modeling the formation and/or evolution of structures that might be of interest in microelectronics, particularly those that are useful in motivating and defining our approach to grain-focused modeling. Section 6.3 presents our approach to grain-focused modeling; what we call the *grain-continuum* (GC) approach. Section 6.4 summarizes some examples that demonstrate the usefulness of the GC approach. Section 6.5 presents some opportunities in GC modeling.

6.1.2 A virtual wafer fab

The performance of an integrated circuit (IC), as with many products, depends upon the details of its structure, i.e. the shapes, sizes, compositions and microstructures of the sub-structures that constitute the product after fabrication

and packaging. ICs are largely designed based upon the lessons learned over the years, e.g. how processing and materials interact with circuit design to determine what can be accomplished. It is well-known that the functional density of ICs has been scaling according to an observation made by Moore[1], which is referred to as *Moore's Law*. In general, the designs for each new IC generation drive the introduction of processes and materials into IC fabrication. Processes and materials are introduced only after extensive, expensive, largely experimental R&D efforts verify that their introduction will result in the expected IC performance and required IC reliability[2,3].

In the IC design process, the performance and reliability of the yet-to-be-fabricated ICs are predicted using two classes of models. The first class consists of models developed using experience and data from previous IC generations, then extrapolated to the new geometries. The second class consists of models that are developed during the process and materials R&D needed for the IC under development. The models use *as-drawn* circuit structures in general, with experimental information introduced to account for the realities of the process line. Considerable expertise is required to evaluate the suitability of an introduced process or material, because decisions as to their implementation need to be made well before the ICs are manufactured: preferably before the design has been finalized. These decisions are largely made on the basis of information obtained from *short flows* and reliability studies on circuit structures, e.g. transistors, test circuits and interconnects. A short flow consists of a small number of process steps designed to test a proposed process sequence and/or to generate a circuit structure to be used in reliability testing. Interconnects are the wires that connect devices (transistors) to make circuits and transfer information between the outside world and the IC. These interconnects are formed using a structure that contains many layers of metal and dielectric materials above the *active region* (where the transistors are)[3,4], in what is called the *multilevel metallization* (MLM)[4–6].

The semiconductor industry established a roadmapping effort in the mid-1990s in order to organize the development of each new generation of ICs, and to reduce the burden on individual IC companies. The International Technology Roadmap for Semiconductors (ITRS)[2,5] has organized the funding of research and development needed to continue IC performance improvement according to Moore's Law. Even with the ITRS, the introduction of a new generation of ICs is extremely expensive. In addition, the delays experienced in the introduction of low dielectric permittivity materials show that even an organized, multifaceted, worldwide R&D effort is no guarantee of unqualified success. Advances become particularly difficult when a goal pushes component properties to incongruous limits. New IC generations, as well as the expected increase in specialized IC designs within each generation, will involve introducing new materials and processes at greater rates than can be realized

using experimental approaches. Analogous concerns are seen in other technology domains.

A 'holy grail' for ICs has been the *Virtual Wafer Fab* (VWF): software that designs, fabricates, and predicts the performance and reliability of ICs[7,8]. A VWF would use the materials' properties and structures to predict the performance/reliability of designed ICs that are generated *in silico* by process models. Development using a VWF would clearly be much less expensive than the design–process–performance–reliability feedback loop that is currently used to advance to a new IC generation. In addition to being far less expensive, a VWF would allow more design options to be evaluated during the time period between IC generations.

Progress towards a VFW is being made, though there is no roadmap. A sequence of changes in *wafer state*, or process steps, can be assembled schematically to demonstrate an idealized or as-drawn IC (or its parts)[9]. Each change in wafer state is achieved in a process or *unit operation*. Unit operations that modify the *topography* of the wafer surface, on the scale of microns, are common, e.g. one unit operation may etch sub-micron *features* into the surface, whereas another might deposit material to fill part or all of existing features. Models, some more physics-based than others, exist for many of the hundreds of unit operations that are used to fabricate an IC, i.e. they predict the change in wafer state due to that unit operation. To simplify discussions, we usually refer to a change in wafer state, though there are across-wafer variations and across-die variations at each process step. (A die is the 'IC-to-be', and can be seen as the centimeter-scale rectangular patterns on an in-process wafer's surface.) Not surprisingly, a range of IC performance is expected, even on the same wafer. Though the focus of this chapter is on the micron and sub-micron scales of features on wafer surfaces and grains in films, it is important to note that models for variations at the wafer and die scales are also important[2]. Process simulations can provide reasonable predictions of the geometries, dimensions, and compositions of IC sub-structures that result from short flows, i.e. a small number of unit operations. With enough effort, models, data, and computational power, such predictions can be extended to more steps and larger simulation domains. For a review of process simulations, see for example Refs 10–13.

Microstructures, properties, and the performance and reliability of IC structures cannot be predicted in the general case. We are particularly far from understanding interfaces, e.g. their roles in the performance and stability of structures[14–16]. Models that relate process model outputs to properties and performance/reliability are just being developed. Process modeling results are still useful, as empirical information exists through which crucial properties and performance/reliability can be estimated from what can be computed or measured. The industry has made it this far using such empirical models, in combination with physics-based models, to develop increasingly functional

ICs. It will make sense at some point in time to build a software structure that can be populated with a mixture of physics-based models, empirical models' and databases. At a high enough level of predictability, this would become a VWF that would then be updated as models and information improve.

6.1.3 Grain-focused modeling

Though realization of a VWF will require significant advances in several areas, we focus on introducing 'granularity' into models[17], i.e. start including microstructural information as well as shapes, dimensions, and compositions into process, property, performance, and reliability models. In turn, microstructures are determined by the materials and fabrication processes used. Product development time and expense would be significantly reduced through the use of simulation tools that are able to reliably predict microstructural characteristics and the resulting performance properties. Accurate predictions by these simulation tools would also require reliable models for the proposed materials and processes. While considerable progress in process and materials modeling has been made in recent years, there are many challenges that remain before we can accurately predict the properties of processed materials and the performance of products that rely on them.

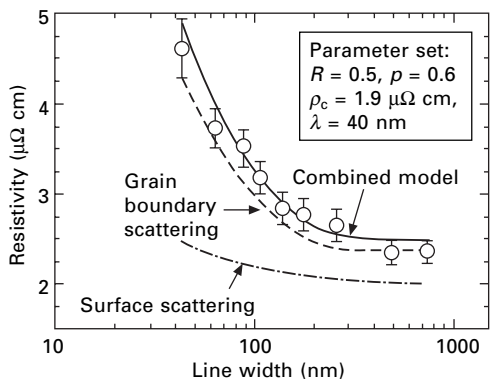
One well-recognized hurdle to the development of predictive processing, property and performance models and software is our limited ability to predict film microstructure. Recent meetings (e.g. see Refs 18–20 devoted to modeling materials and processes across length and timescales have made it clear that atomic scale (atomistic) simulators need to be combined with continuum simulators to form multiscale simulators that can predict microstructure. Implementing accurate materials and process models in such simulators will accelerate the development of products that depend upon microstructured materials, e.g. in microelectronics, aerospace, structural materials, and consumer products. Atomic scale simulations can in some situations provide enough information to explain product performance, e.g. if performance is limited by very localized phenomena that can be identified and analyzed. On the other hand, the structures to be analyzed are often on the order of tens to hundreds of microns or even larger (*mesoscale* to *macroscale*), and are not reasonably addressed by atomistic simulators. The timescales involved in product fabrication and reliability (lifetime) are also usually beyond what can be reasonably addressed using atomistic simulators. In turn, larger scale, continuum models and simulators can in some situations provide enough information to explain product performance; however, continuum approaches usually do not represent all of the phenomena needed for accurate predictions. So, the consensus is that atomistic and continuum simulators should exchange information and be solved simultaneously. Finally, the combined atomistic

and continuum simulators should be able to run on computational facilities that can be routinely accessed in a corporate environment, as well as on computers used to develop process and materials models.

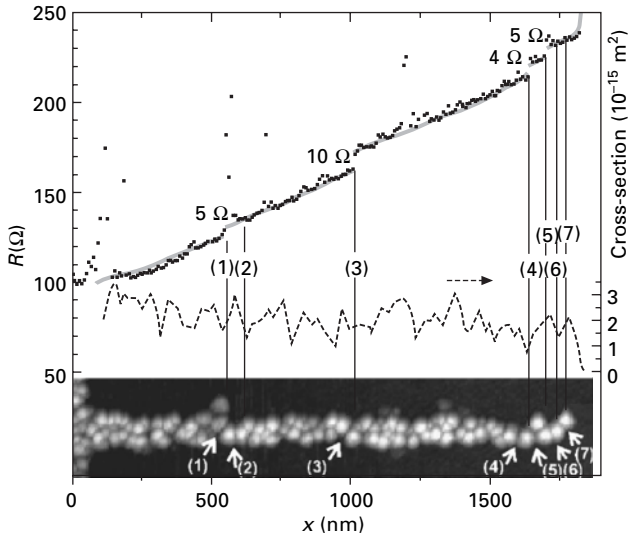
Software capable of combined continuum and discrete simulations relevant to the microelectronics industry is not available[2]. Our research has focused on a representation that we consider to lie between continuum and discrete approaches. This chapter is intended to motivate, define, and demonstrate a grain-focused approach to process and materials modeling, which we label GC modeling. The goal is to provide engineers with microstructure-centric models that can deal with the large volumes often needed to interpret performance and reliability results. In addition, the GC model ideally should provide procedures to include atomistic information.

6.1.4 Focus on copper

To focus this chapter, we deal with copper-based MLM structures used in ICs. The global experience and expectations of Cu-based MLM make it clear that the microstructures of the Cu and associated (barrier, adhesion, and seed) films are critical[2,5]. As indicated in Fig. 6.1[21], the resistivity of Cu interconnects is increasing (a negative for performance) as ICs feature sizes shrink, because shrinking the lines (a) increases the relative size of the surface scattering term, and (b) increases the size of the grain boundary scattering contribution[21–23]. The latter occurs because the grain size is constrained by the line width. Figure 6.1 shows that at 45 nm line widths, grain boundary scattering is the largest contribution to the electrical resistivity of Cu[21]. This is because the grain sizes are getting to be same size as the mean free path of electrons in Cu (~ 40 nm). Figure 6.2 highlights this issue for very small metal interconnects. Note the significant jumps in resistance



6.1 Measured resistivity vs line width, compared to resistivity models that include surface or grain boundary scattering (from Ref. 21).

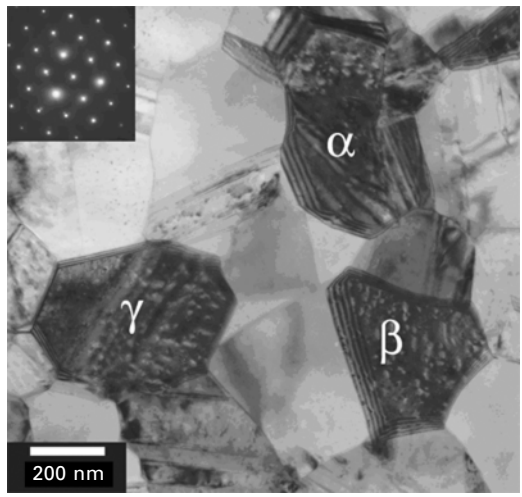


6.2 Resistance vs probe position along a nanoscale gold line, showing resistivity jumps between grains[24].

upon moving a probe from one grain to an adjacent grain[24], which indicates that electrons are scattering significantly at that particular grain boundary. Grain structures can also impact the reliability of the Cu interconnects, i.e. how long they last during use in products[25–27]. In addition to these motivations, there are phenomena related to grain structure that affect both processing and reliability, e.g. Cu grain growth after electrochemical deposition[28–31] (see Section 6.4.3). In short, grains and grain boundaries are playing ever larger roles as IC components shrink.

To impact IC design and fabrication, the models we use should improve our understanding of how grain structures arise; how they evolve in time during processing, storage, and use; and how they might be tuned to improve performance. Experimental observations of grain structures are possible[32,33], as shown in Fig. 6.3[34]; however, they are difficult. Characterization techniques are often destructive, require sample preparation that can alter grain structures, and only give limited information about shape and size. This is an ideal situation for computer-based modeling to play a role in a process and/or material optimization cycle. Further, simulation of grain structures can help in the interpretation of measured data, e.g. the reconstruction of 3D structures from 2D measurements[35], such as sequences of cross-sectional images like that in Fig. 6.3b. By reducing the need for experiments, grain-focused simulation tools will increase the efficiency of material, process, and performance model development.

(a)



(b)

6.3 (a) Top and (b) side cross-sectional TEM images of a Cu thin film, showing individual grains (from Ref. 34). (The Greek letters are not relevant to this work.)

6.2 Representations and models

6.2.1 Overview

From our perspective, structural modeling and simulation can be broken down into two major subtasks. First, there must be a way to represent the structure, both at one time and as it evolves in time. Second, a model of the appropriate physical process must be applied in the case of an evolving structure. In this section we review the aspects of structure and process modeling that are particularly relevant to this chapter. It is broken down first by the approach to the first task, that of representation, and then further subdivided by the model physics.

6.2.2 Discrete representations

Discrete models treat solids as atoms[36–39] or aggregates of atoms ('discs'[40–43] or clusters[44]), and can provide information on film microstructure and its evolution. Atomistic approaches to film evolution are based upon either solving a dynamics problem for each atom during a brief period of time (molecular dynamics (MD)[38,45]), or using hopping probabilities for atoms as they move to find lower energy states (Monte Carlo (MC)[46,47]).

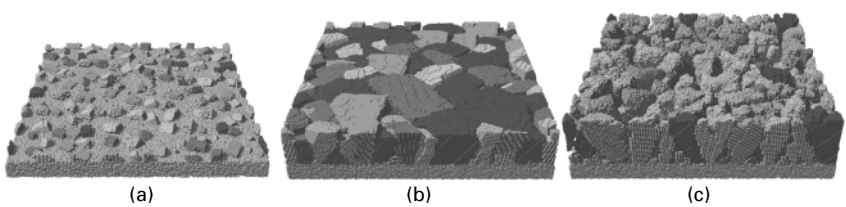
MD methods can be used to calculate parameters that directly affect structural evolution, such as grain boundary migration rates or vacancy cluster diffusion coefficients[48]. Alternatively, MD methods can be applied to both deposition and evolution of existing structures directly. However, such MD structure simulations are computationally very expensive and are thus not suitable for typical processing times (minutes to hours) or for the length scale of typical manufactured structures (tens of nanometers and up, even for modern transistors). Consequently, in MD deposition simulations, fluxes must be chosen to be unrealistically high in order to simulate growth of several atomic layers, and artifacts such as large roughnesses result from the high deposition rates. Moldovan *et al.*[49] have applied MD methods to direct simulation of nano-grained structures of up to 100 nm square by only treating a few atomic layers periodically in the third dimension. Huang and coworkers have studied atomic diffusion on and mechanical deformation of surfaces using a combination of molecular dynamics and density-functional theory based *ab initio* calculations[50–52].

MC techniques are often applied to atomic representations directly. Kinetic lattice MC (KLMC) simulations are used to study deposition and diffusion on lattices or templates. Whereas MD methods model systems on the timescale of atomic vibrations, KLMC deals with the far greater timescale of atomic jumps. KLMC can simulate much larger systems (billions of atoms) for longer times, and is used to simulate thin film growth under more realistic deposition rates than MD. The essence of KLMC is to determine the jump rate of all atoms in the system, then to pick one atom (using a distribution weighted by its jump rate), and then to move it appropriately. MC models based on transport and deposition of many-atom aggregate particles (clusters) have been used to model film growth over length scales representative of IC topographies[40–44,53,54]. These models yield understanding of film profiles, but their contribution to understanding microstructure is not clear. Potts models, a kind of MC approach, in which clusters of atoms are treated as entities with their own orientation and updated based on transition probabilities, have been used in both 2D[55,56] and 3D[57,58] to examine grain structure evolution. Potts models, as well as other types of cellular automata-based models, do not extend well to systems with coupled fields, including those with applied voltages and externally imposed strains.

The validity of KLMC models depends on the underlying assumptions regarding atomic mobility, including both the type and rates of the allowed jump processes. This information can be experimentally or theoretically determined. A number of studies of diffusion of atoms on surfaces exist, especially for metal on metal and Si on Si (e.g. see Refs 59–61). There have been many studies of Cu on Cu, e.g. see Ref. 62 for a review, and estimates are available for most of the jump parameters needed to model film growth and evolution.

Because of the importance of the initial stages of deposition, nucleation has been the subject of numerous studies. Most attention has been given to nucleation during *physical vapor deposition* (PVD), in which deposited atoms impinge and condense directly onto a surface, as during deposition by evaporation or most sputtering processes. Many groups have reported MC simulation models for nucleation[38,63–71]. Reviews of both experimental and theoretical studies of nucleation phenomena can be found in Refs 38 and 72. Less attention has been given to the nucleation and initial phases of growth by *chemical vapor deposition* (CVD)[73–81]. This is largely due to the complexity of modeling gas phase and surface chemistries in CVD. (See Section 6.4.2 for additional discussion of PVD and CVD.)

Huang and co-workers developed ADEPT[47,82–86] to simulate thin film deposition. Figure 6.4 summarizes an ADEPT-based study of the trend in texture formation during Al deposition[82,87]. ADEPT has gone through several stages of development: (a) single lattice in three dimensions[85], (b) dual lattices in three dimensions[82], (c) multiple lattices in two dimensions[47], and (d) multiple lattices in three dimensions[87]. ADEPT is now capable of simulating texture competition during thin film deposition, under realistic deposition rates, at the atomic level, and in three dimensions[87]. In the current implementation, an efficient numerical algorithm of three consecutive 2D mappings has been implemented. Simulations of 0.2 micron thin film deposition are possible on standalone PCs.



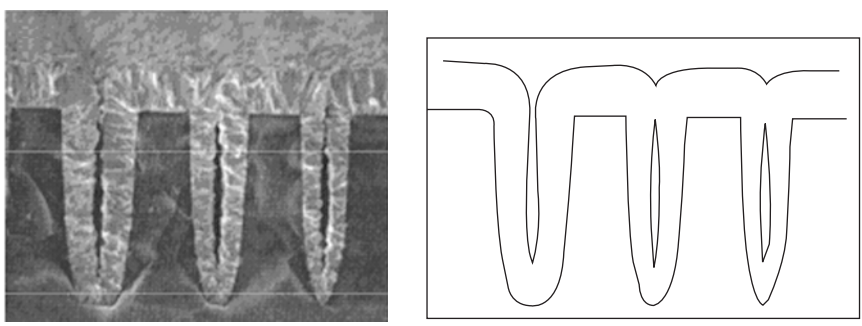
6.4 ADEPT simulation results of (a) grain nucleation, (b) low-energy texture dominance, and (c) columnar growth. Lighter shades indicate lower energies.

6.2.3 Continuum representations

These approaches represent films as continua, with each continuum being a single material, perhaps with spatially varying properties. The material interfaces which bound each continuum can be moved according to materials fluxes. A common use of this representation is for 'feature scale' topography evolution during processes used to fabricate ICs. Continuum representations are particularly useful in the context of multiscale simulation, because they interface naturally with the representations of phenomena in reactor scale simulations, which are often also continuous in nature, being finite-element or finite-difference based.

Several (IC-oriented) continuum simulators have been developed to study the evolution of thin films[10,88–92]. EVOLVE, developed by Cale and co-workers[7,13,93–95] has been used to simulate a variety of deposition, etch, and reflow processes, i.e. to determine the topography and composition of processed films and surfaces[10,16,94]. EVOLVE uses explicit representations of materials in 2D, while calculating transport of materials in 3D[92]. One of its strengths is the ability to incorporate complex chemistries[7,95]. Figure 6.5 shows an example of EVOLVE's predictive ability for the CVD of tungsten using the hydrogen reduction of tungsten hexafluoride[96]. Discussions of continuum process integration studies, in which several unit operations are simulated, to correspond to a short flow in the process line, have been reported on[97,98].

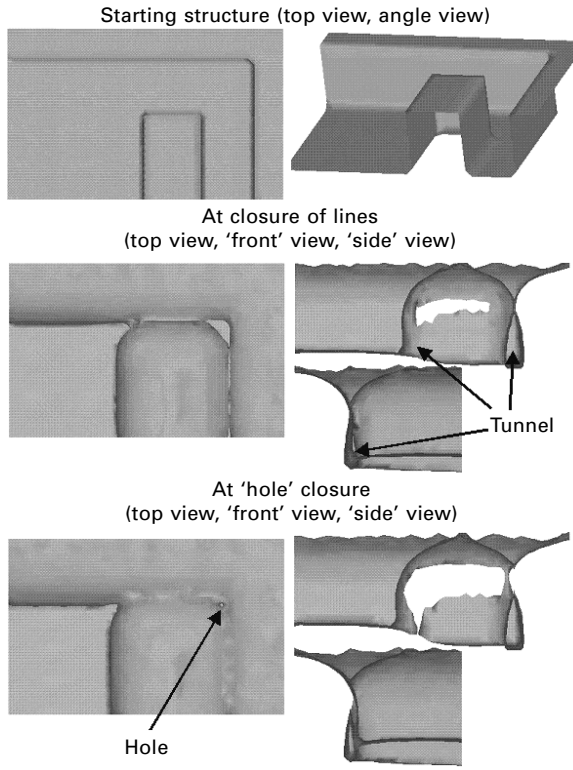
Topography simulations on 3D surfaces are sometimes required. Most 3D topography codes use *level-set* methods[99–101] to avoid numerical difficulties due to changes in topology. Level-set methods are based upon representing an interface implicitly, as a contour or level set in a scalar field, and reducing the explicit motion of the interface to the time evolution of the scalar field under a differential equation[99–101]. The computational overhead of representing a field in d dimensions to track a $(d-1)$ dimensional interface is deemed to



6.5 (a) SEM of W lines deposited using the H_2 reduction of WF^6. (b) EVOLVE result[96].

be made up for by not having to deal explicitly with several geometrically troublesome issues, such as topological changes in the structure. Borucki and co-workers developed a 3D topography code that uses level sets for surface evolution[7,11,102–105].

For example, Figure 6.6 shows the results of a simulated plasma enhanced CVD of silicon dioxide to fill gaps between lines. Details can be found in Refs 11, 104 and 105. A 3D topography simulator was required because the focus of that study was the corner region and tunnel formed during this fill process. It is quite reasonable that a corner region, being wider on the diagonal, requires a thicker deposited film to close relative to a straight run of 'trench' between lines. The more interesting reliability issue explained using this simulation is that the tunnel formed by the void in the deposit is open to the environment. This 'tunnel entrance' is due to one line ending while the other continues. Process fluids (gases/liquids) can enter the tunnel during downstream



6.6 3D/3D PETEOS simulation over a structure designed to demonstrate the closure of holes in corner regions and particularly the tunnels formed when one line ends and an adjacent line continues. Side views of closed regions are transparent. Figure taken from Ref. 11.

processes and fill the gap formed during deposition. This caused the dielectric to 'pop off' at these regions of the IC.

The phase field method[106,107] is another implicit method for representing continua, that is conceptually similar to level-set methods in some ways. Phase field methods conceptualize the boundary of materials as having a finite thickness (ε). This interface thickness is advantageous in formulating field-based models of certain processes, such as solidification from melts[106,107] and phase separation of binary materials[106]. One disadvantage of the finite interface thickness is that the underlying mesh or grid must have a length scale of the order of ε in the neighborhood of the interface.

6.2.4 Grain-continuum representation

Continuum simulators have been extremely useful for process design; however, they capture only the evolution of surfaces and interfaces. Microstructures such as nano-voids and grain boundaries are not represented. This limitation becomes more serious in applications of semiconductor materials processing because of the ever smaller feature sizes in ICs. A GC representation retains more information than a continuum representation of a microstructured system. One key feature is that each grain is a continuum surrounded by a grain boundary that is 'honored' by the mesh that represents the films and structures.

One way to approach GC representation is through an explicit representation of the boundaries between continua, by either segments in 2D or polygons in 3D. Frost and Thompson[108] used such an explicit network model in 2D to study the evolution of grain size distributions. Zhang and Adams developed FACET[109], which explicitly tracks grain boundaries and surfaces in 2D through the use of line segments. Roosen and Carter[110] and Moldovan *et al.*[49] performed coarsening simulations on polygonal grains, using network models. Kuprat *et al.*[111] reported on an explicit, 3D, moving boundary grain structure evolution code (GRAIN3D) that uses the Los Alamos Grid Toolbox (LaGriT)[112] and has demonstrated it for grain boundary motion under mean curvature[113]. Many complex operations are needed to maintain an explicit topology during evolution[114–116], and implicit approaches (e.g. level sets) are used by most codes to evolve 3D structures of any complexity.

Phase field methods that use multiple order parameters[117] to designate poly-phase systems are implicit methods that we include under the heading of GC. As in the two-phase variant, poly-phase field interfaces have finite thicknesses that must be resolved spatially[118]. Moreover, the formulation and solution of the appropriate Cahn–Hilliard[119] and Ginzberg–Landau[120] equations for the evolution of the order parameters become more difficult with additional phases[117,121], with the required energy parameters relying on phase diagrams of higher order systems that are more difficult to calculate[117].

Finally, level-set methods can implicitly represent the boundaries of grains.

The traditional level-set method[96] relies on a single level set to track one type of interface. However the multiple level-set method, as will be discussed in detail in the next section, allows for the use of multiple scalar functions to represent many grains in a structure.

6.2.5 Discrete and continuum

There are a few reports in the literature in which discrete and continuum approaches have been combined to some extent. The trade-offs made deal with the degree of detail and explicitness with which the atomic scale information is determined and 'passed to' the continuum simulator. The goal is to get physics from atomic scale discretizations to the micron scale, mesoscale (10 to 100 micron) and macroscale (100 micron to millimeters and larger). In general, it seems reasonable to do this by representing small volumes of the system under study at the atomic level. These volumes will be at places where structures and/or processes at the atomic scale significantly influence the structure and/or process being studied.

Ortiz and co-workers[122–126] have solved model problems using a *quasi-continuum* approach[123,124,126]. In carefully constructed systems of crack-tips in single crystals, atomistic representations in the form of kinetic Monte Carlo simulations were linked directly to finite-element simulations using highly refined meshes and identifying 'key' atomic locations. The resulting simulations predict dislocation shedding from crack-tips, and crack-tip propagation. An extension of the method to finite temperatures has also been proposed[127]. These multiscaling procedures average sampled atomistic information, and hand it to larger scale solution procedures; note that this approach takes advantage of the fact that finite-element representations of continua are discrete. Refs 122 and 124 contain reviews of quasi-continuum methods.

Level-set methods have been used to study epitaxy in '2 + 1 dimensional' types of simulations by Ratsch and co-workers[128,129] and by Chopp[130]. In these island dynamics studies, level sets are used to represent a continuum in two dimensions, on a small number of discrete atomic layers that act as a third dimension. The continuum representation is modeled based on rate equations, while the discrete information (namely the monolayer information) is injected into the simulation in the form of nucleation events, with probabilities based on the continuum information such as concentrations of adatoms.

Bloomfield *et al.*[131] used an *encapsulation* procedure on the discretely represented islands that resulted from KLMC simulations to obtain continuum representations of the islands. This set of islands is the starting point for GC vapor phase deposition simulations that result in poly-granular films. Encapsulation is discussed further in Section 6.4.2.

6.3 Grain-continuum approach

6.3.1 Overview

The heart of the GC approach is the view that a microstructured material consists of multiple phases that are distinct from each other and each can be modeled as a continuum. As used here and in the discussion to follow, the term *phase* indicates any material in which a property of interest is continuous. To apply this terminology to a polycrystalline material, the phase property is chosen to be crystallographic orientation, and a phase becomes identified with a grain or perhaps a set of grains with a common orientation. The desire to retain the internal continuum nature of each phase seems to make the discrete building-block approaches, such as the lattice or aggregate representations[38–46] discussed in Section 6.2.2, unsuitable. The phase field and level-set techniques discussed in Section 6.2.3 allow for only one type of interface in a given model, and seem to limit their applications to two phases.

6.3.2 Multiple level sets

We think a 'true' GC representation can be achieved by extending the level-set method to include multiple scalar fields, as done by Merriman *et al.*[132]. In this multiple level-set (MLS) representation, each phase has its own scalar field represented on a grid or mesh throughout the computational domain. This MLS representation allows each phase to have its own properties represented as fields, distinct from each other, but continuous in space. As with the traditional level-set method, the sign of the scalar variable at a given point indicates whether that point is inside or outside a phase. It is worth noting that there can be parts of the zero contour for a given φ that are not connected to each other, which would indicate a distributed phase. Whether using one level-set field or multiple level-set fields, the mapping of the motion of an explicit phase to the implicit representation of that phase is accomplished by the applying the following level-set equation, of the Hamilton–Jacobi type[99–101], to the scalar representing that phase:

$$\frac{\partial \varphi}{\partial t} + v \cdot \nabla \varphi = 0 \qquad [6.1]$$

Here, φ represents the level-set field associated with the phase being evolved, and v is the velocity of the boundary of that phase. The $(d-1)$-dimensional boundary of the phase moves in d-dimensional space as the scalar function evolves and changes position of the contour where φ equals zero. This equation reduces the task of tracking the evolving phase to determining the velocities throughout the model domain, finding the gradients of φ, and integrating φ

in time. Equation [6.1] must be solved for each phase with non-zero velocities. For the common situation in which the velocity is only known as the speeds in the normal directions on the surface(s), the least intuitive part of the task is probably constructing a velocity field away from interface, but there are well-defined procedures for *extending the velocities* at the interface into the domain[133,134].

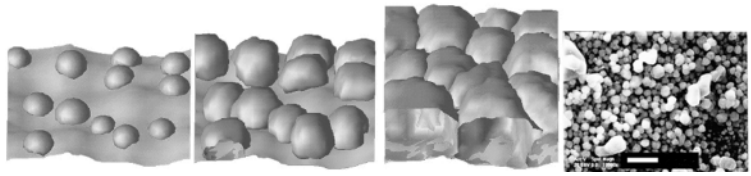
There are several highly efficient integration schemes for Equation [6.1] that account for the fact that the physical situation represented by the scalar fields requires that information move only downwind, i.e. away from boundary of the phase. When the fields are represented on regular grids, *fast-marching* methods are often employed to update points on the grid in the correct order[101,133]. When finite-element meshes are used to represent the fields, *streamline upwinding Petrov–Galerkin* (SUPG)[135] methods can be employed to ensure the correct information flow.

Non-physical situations can be represented when using MLS fields[132,136], such as a point being inside multiple materials at the same time. To prevent such a situation from occurring, the scalar variables must be subject to a set of constraints. These constraints are usually applied by integrating the individual fields for a short time, then *reconciling* the fields to get rid of any overlaps that occur. Ideally, whatever velocity field is computed for each side of a given interface will be equal and opposite (as they represent the same interface, just from different points of view) and overlaps can only occur due to numerical error.

Numerical errors can build up in the field away from the zero levels and may cause instabilities in the integration schemes. Usually some sort of *redistancing* scheme is used to keep the fields well-conditioned away from the zero level contours without moving the position of the zero level contours themselves[137,138]. With an appropriate choice of scalar field, such as that of the signed distance to the interface, the same Hamilton–Jacobi solver used to integrate Equation [6.1] can be used to maintain the condition of the field[137,138].

6.3.3 PLENTE

PLENTE[139] is software that implements a grain-focused, MLS method to represent and track distinct phases as they change in time (similar to Ref 132). PLENTE is a parallel, finite-element code that uses an unstructured tetrahedral mesh to represent multiple phases, all or some of which might be grains in a microstructure. PLENTE discretizes the n scalar functions that represent the n phases in a given simulation, using the unstructured mesh to allow local refinement of the representation to achieve finer resolution of the geometric details in interesting regions of the structure. In order that all the information in the system comes from the interfaces and moves into the domain, we use a SUPG technique. The SUPG method used to solve the

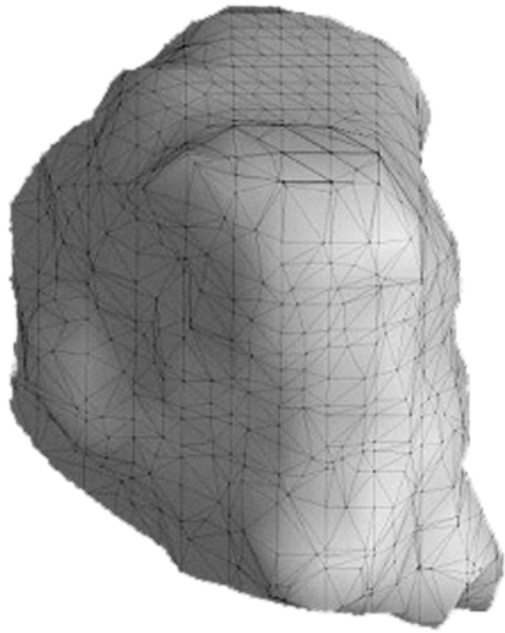


6.7 (a) Initial structure for ELD simulation, showing distribution of hemispherical nuclei on 5 nm RMS rough surface. (b) Onset of coalescence during ELD simulation as described in the text. (c) Fully coalesced grain structure. Note that only the grain surfaces are shown here. (d) Micrograph[141] showing spherical nuclei during electroless deposition of copper.

Hamilton–Jacobi equation (Equation [6.1]) was first formulated for triangular elements (in 2D)[135] and extended to tetrahedral meshes (in 3D)[136].

PLENTE interacts with process simulators to determine the rates of evolution of 3D structures. Several examples are discussed in Section 6.4. To illustrate the basic concept, we consider the simple electroless deposition (ELD) model (process simulator) that was used to generate the simulation summarized in Fig. 6.7*. We start by generating hemispherical nuclei distributed randomly along a rough surface (see Fig. 6.7a). PLENTE then gets the specified linear growth rate from the model. Here the model is simply that the free surface of the nuclei grows with a uniform, constant velocity and that sub-surface interfaces have zero growth. PLENTE then evolves the islands for a small time step and represents the new structure (see Fig. 6.7b). It then gets new growth rates (in this case, the same growth rate on triangles that are exposed to the source), and the cycle is repeated until a given film thickness is attained. Figure 6.7c shows the resulting 3D polycrystalline film. To retain the polycrystalline nature of the resulting film requires a GC representation In addition, PLENTE stores information regarding the grains. For example, suppose the by-products of deposition build up in the solution and act as contaminants. The contaminant concentrations can be tracked as a function of position inside each grain, and used in further process simulation, or even in a performance or reliability calculation. Figure 6.8 shows an example of a grain that has been ‘pulled out’ of the polycrystalline film generated in Fig. 6.7 and can act as an input for such analysis.

*Although the reason for discussing this simulation here is to demonstrate what PLENTE does, we also note that the model used is a reasonable start toward a more complete model. Based on evidence of spherical grains, such as those seen in the micrograph shown in Fig. 6.7d, we concluded that the ELD process could be adequately modeled by assuming isotropic deposition.



6.8 Example grain, 'pulled' for inspection from a polycrystalline film during a deposition simulation using PLENTE.

Computational burden is an important issue in any proposed simulation effort. PLENTE employs the Message Passing Interface (MPI) libraries[140] to spread its computational workload across many processors in parallel. Each phase (or set of phases) may be placed on a separate processor, allowed to evolve in time, and then reconciled on a master processor. A primary advantage of this parallelization is observed when the information needed to calculate the grain boundary velocities on a given grain is contained in the shape of that grain, and so simulation time for evolving the structure remains approximately constant as more grains are added to the simulation. Interaction with other codes can be done either from the master, such as is done when PLENTE uses the process simulators discussed in Section 6.4, or can be done by each slave as necessary, such as is done when employing a separate mesh coarsening code.

Most process simulation codes that PLENTE needs to interact with require explicit representations of the system. PLENTE must be able to extract an explicit representation of the system geometry, in the form of a surface mesh or a volume mesh. The normal procedure in the single level-set method is to find the zeros of the level-set scalar functions on the grid lines or mesh edges of the grid or mesh used to represent the function, and then connect those points appropriately with triangles or quadrilateral patches (or with line

segments in 2D problems). This procedure is quick and gives good answers; however, it should be noted that it is an over-specified system because the MLS method represents each interface twice (once for each phase that makes up the interface). When the errors associated with the finite resolution of the underlying field representation are not exactly equal, direct extraction of the interface from the scalar fields can create two different representations of the interface. There are various techniques to resolve these differing points of view on where the interface lies, and this is still a field of ongoing research, as discussed in Section 6.5.3.

6.4 Grain-continuum examples

6.4.1 Overview

It is common to make a distinction between processes that involve grain structure formation and processes that involve grain structure evolution. In the context of this chapter, formation problems involve processes in which grains are being grown as material is being added to a free surface, whereas grain evolution problems involve processes that move existing grain boundaries. Most process simulations fall into the category of formation, in which granular films are deposited from a gas or a liquid that contains a source of material. Most reliability simulations fall into the category of evolution, in which existing grains change shape and the boundaries between grains move. Although in the examples below, the distinction between formation and evolution is kept for the sake of clarity, we note that this distinction is not necessary within a GC simulator such as PLENTE. Because a gas or liquid is represented as just another phase, albeit with significantly different properties than the solid phases, the same data structures may be used for both types of problems and grain formation becomes a special case of evolution. Thus, PLENTE can deal with cases in which both formation and evolution are important aspects of the same system. We also summarize an example GC simulation on a microstructured, but not polycrystalline, material.

6.4.2 Grain formation

Deposition model

In microelectronic processing, deposition of polycrystalline thin films is often done at low pressure from a vapor phase. In the case of PVD, the vapor contains atoms or clusters of atoms of the material to be deposited that adsorb directly onto the growing film. In the case of CVD, the vapor contains a metastable precursor that first adsorbs and then undergoes a reaction that produces the deposition material. By-products of this reaction step are volatile and leave the surface. In PVD, material from the vapor phase tends to stick

(close to) where it hits the substrate; the sticking probability is close to 1. In CVD, depending upon the reactivity of the precursor and the temperature of the substrate, the precursor may desorb and leave the surface before having a chance to react and deposit. In a typical CVD process, the probability that a given precursor molecule reacts on any given visit to the substrate surface is 0.001 or less.

Though the impetus behind this chapter is to explain the role of grain-focused modeling, it is worthwhile summarizing a basic transport model, which will be essentially the same for continuum and GC representations. Consider low-pressure deposition processes at the feature and/or grain scale. At conditions typical of low-pressure deposition processes, e.g. PVD or CVD, the gas is rarefied enough that particles are likely to strike one of the feature 'walls' or other solid surface (grain boundary) before striking another particle. Thus, transport is 'line-of-sight' or *ballistic*. More details of the basis for and limitations of the ballistic transport and reaction model (BTRM) can be found in Refs 7, 94 and references therein. The mathematical representation of the basic BTRM, which suffices for the purposes of this chapter, is a set of integro-differential equations that represent diffuse transport and chemical reactions. Ignoring surface diffusion to keep things simple, the integral equations

$$\eta_i^{a,t}(\mathbf{x}) = \eta_i^{a,1}(\mathbf{x}) + \int_{\partial R} q(\mathbf{x}', \mathbf{x})\, g_i(\theta, \phi)[\eta_i^{a,t}(\mathbf{x}') + R_i(\boldsymbol{\eta}^{a,t}, \xi, T)]\mathrm{d}\mathbf{x}' \quad \text{for} \quad \mathbf{x}', \mathbf{x} \in R \qquad [6.2]$$

predict the growth rates of surfaces (∂R are surfaces in the domain R) as functions of species fluxes from a source[7,91,142]. This integral equation expresses the total flux of a species ($\boldsymbol{\eta}^{a,t}$) to a point ($\mathbf{x}$) on the surface due to the flux of that species both from the source ($\boldsymbol{\eta}^{a,1}$) and from other parts of the surface ($\mathbf{x}'$) due to re-emission of material that did not react (stay at $\mathbf{x}'$). The functions R_i represent the rate that the i-th species is produced by all the heterogeneous reactions in the system, based on the fluxes ($\boldsymbol{\eta}^{a,t}$), the coverages (ξ) of adsorbed species, and the temperature. The function $q(\mathbf{x}', \mathbf{x})$ is called the transmission probability and is non-zero only if $\mathbf{x}$ and $\mathbf{x}'$ can see each other, and it is a function of the distance between x and x' and the angles each surface makes relative to the line segment that connects the two points (analogous to radiation heat transfer). The function $g_i(\theta, \phi)$ is a distribution function that gives the relative number of molecules leaving $\mathbf{x}'$ at the specified angles, i.e. it indicates the fraction of flux re-emitted in a differential solid angle around the direction (θ, ϕ)[7,92,142]. It is a constant for *diffuse* (also known as *cosine* or *random*) re-emission.

Accurate predictions of film deposition and topography evolution for specific processes rely on having knowledge of the distributions of the fluxes

of incoming species (flux distributions) and the relevant reaction rate expressions. Once estimates are available for these constitutive relations, film profile evolution can be simulated. Several flux distributions have been used; simple ones are discussed in Refs 7, 92 and 142. References 7, 95 and 143 demonstrate the use of complex reaction rate expressions, that may involve several surface species and several surface reactions. Often, for simple, single-component systems such as PVD, the reaction rate is well-characterized by a sticking probability or *sticking factor*. This corresponds to a reaction rate expression

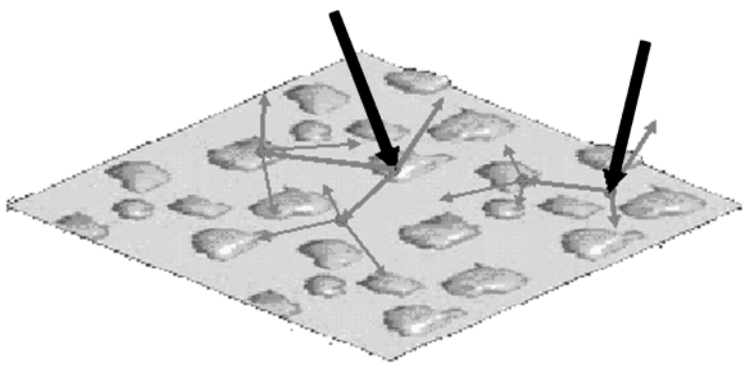
$$R_i = -\sigma_i \eta_i^{at} \qquad [6.3]$$

where σ_i is the sticking factor for species i, ranging from 0 to 1.

PLENTE has been used to guide the simulation of deposition from the vapor phase by interfacing with code based upon EVOLVE to solve the BTRM's integral equations, i.e. it has served as part of a traditional continuum 3D/3D process simulator to simulate deposition and etch processes. More specifically, PLENTE has been used as a 3D/3D topography simulator much like ASSET was used (see Section 6.2.3). Though representing and solving such problems on 3D surfaces is computationally expensive, it is really just a (difficult, expensive, but easy to understand) extension of work done using process simulations on 2D surfaces (surfaces that can be represented by cross-sections). The problem gets much more difficult when polycrystalline nature of films is studied.

Grain formation

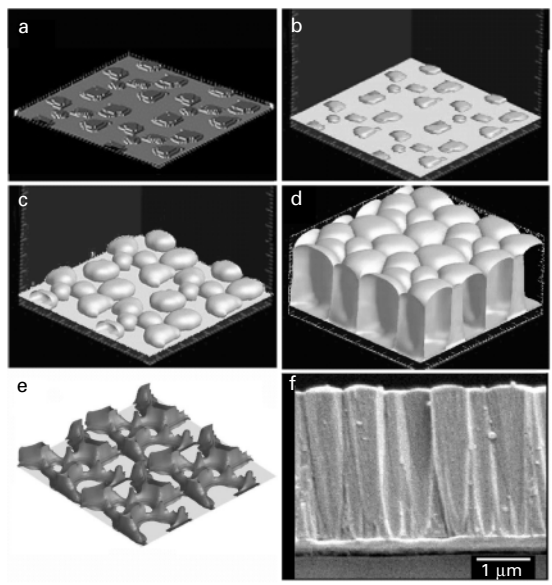
Consider using the BTRM to model grain growth from post-nucleation through polycrystalline film formation. Figure 6.9 shows the concept behind the



6.9 Schematic representation of a key concept behind the BTRM, redistribution of unreacted species.

BTRM, as applied to growing grains, i.e. moieties (atoms/molecules) from the source strike the surface and either react or desorb, perhaps multiple times, before either reacting or returning to the source. There is also line of sight transport between grains, and between the bare substrate and the grains.

PLENTE can create new nuclei in the simulation from structural information from experiments, from atomistic simulations, and/or from statistical instructions about the times, sizes, shapes, and locations for insertion. Figure 6.10 summarizes a PVD simulation that PLENTE guided, starting from the set of discretely represented nuclei shown in Fig. 6.10a, which were generated using a Monte Carlo code[63,70,71]. The first step in the process is to convert the discrete islands into continuum representations, by encapsulating[136] them. Encapsulation forms a boundary representation fitted to an atomistic representation (such as an island, or piece of a crystal/grain). The boundary needs to be consistent with the other elements in the structure. This procedure results in a GC phase, which represents an island or grain. This procedure is repeated until all atomistic representations are represented as continua; we then have a GC representation. We find the easiest stage in a deposition simulation to do this is as soon as the individual crystallites reach a length scale that is large enough to be represented by the level-set representation on a mesh that is appropriate for the larger scale problem under study. For this

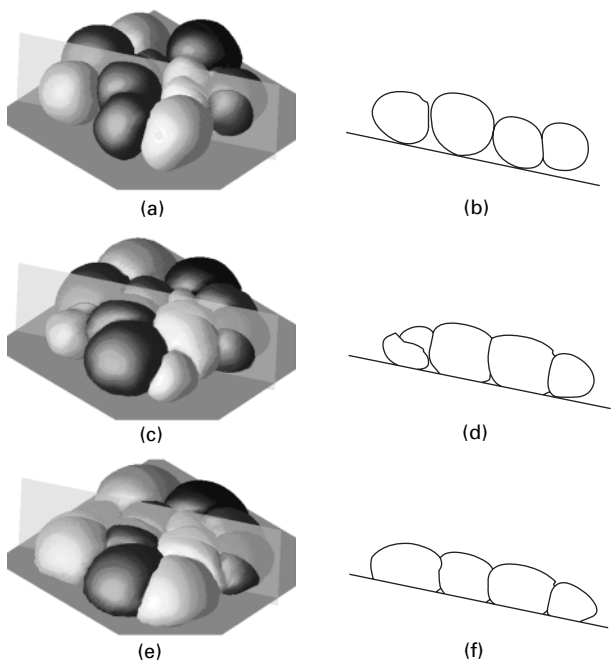


6.10 (a) Islands from an MC simulation. (b) The 'encapsulated' islands. (c) Impinging islands during growth. (d) Final grain-continuum film. (e) Void volume at the substrate surface. (f) SEM cross-section of PVD tantalum oxide film[144].

reason, it can be advantageous to refine the level-set meshes in PLENTE in regions of grain nucleation and island formation.

The result of encapsulating the islands shown in Fig. 6.10a is shown in Fig. 6.10b. Then PVD simulation starts. A 3D view factor code (FVIEW)[105] is used to determine the transmission probabilities needed to compute how much of the source flux arrives at each discretized area of the surface[105]. Finally the islands impinge upon each other to form a continuous film, as in Fig. 6.10d. Figure 6.10e shows the void volume near the original substrate. Figure 6.10f is a cross-section of a PVD tantalum oxide film after flashing with gold to provide conductivity[144]. The existence of voiding near the original substrate surface is indicated.

The void formation noted in the previous paragraph will reasonably be a function of the sticking factor of the reacting species. Figure 6.11 shows the result of deposition simulations using PLENTE to look at the effect of sticking factor[145]. The reactivity of a precursor is varied from extremely high (very fast reactions or essentially unity) sticking factor to very small sticking factor, corresponding to σ of Equation [6.3] taking on values of 1.0, 0.5, and 0.0. Deposition takes place onto a series of small nuclei on a flat substrate.



6.11 Simulated grain structure and cross-sections of Cu CVD: (a, b) high (σ = 1.0), (b, c) moderate (σ = 0.5) and (d, e) low (σ = 0.001) precursor reactivity as noted in text. Void volume decreases as deposition rates become more isotropic.

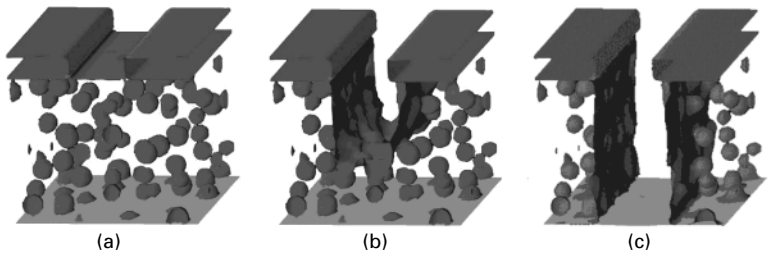
In this simulation, both FVIEW and the feature scale simulator EVOLVE are required, in order to calculate the redistribution of unreacted precursor. Notice that the film becomes denser as the reactivity of the precursor goes down. The reason for this phenomenon is that when the precursor is allowed to desorb before it has a chance to react, it may go on to visit other parts of the structure that are normally shadowed in the higher sticking factor case. This phenomenon explains why there are significant voids near the substrate in the high sticking factor case but none in the low sticking factor.

6.4.3 Porous dielectrics

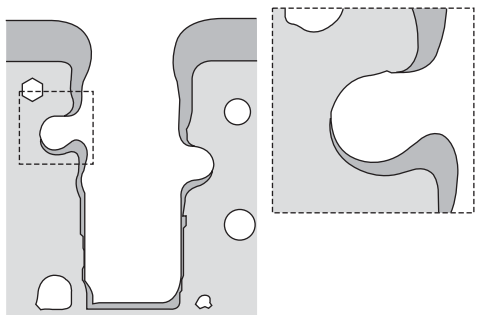
Using a GC simulator such as PLENTE, we can address issues on non-homogeneous substrates as well as the monolithic substrates in the above examples. Though the focus of this chapter is polycrystalline films, the GC approach and PLENTE really address a wider ranger of problems, such as those that deal with more general microstructure. One of the recent goals in the IC industry has been to reduce the relative dielectric permittivities of the dielectric materials used in MLM stacks. This decrease can be done by including empty pores in the material, bringing the effective dielectric constant of the material closer to unity. However, the presence of pores and pore networks can raise processing issues. As examples, we look first at etching* porous substrates followed by a barrier deposition into the resulting trenches cut into the porous substrates.

Figure 6.12 shows an example of the etching of a porous substrate. In this simulation, a porous dielectric layer is being etched through a mask by high-energy argon ions (Ar^+) extracted from a plasma in a process called ion milling. This type of physical etch process has the advantage of being quite anisotropic while not using chemically reactive species that can contaminate the pore structure. In our PLENTE simulation of this phenomenon, the incoming argon ions are modeled as being mono-directional, coming in perpendicular to the surface of the wafer. Because the dielectric material is so soft, the selectivity towards the dielectric versus the underlying etch stop is chosen to be quite good for a physical etch (20:1) and 1:0 for the substrate versus the mask. As can be seen in Fig. 6.12, the substrate is ablated away by the incoming ions. When a pore is opened up, it is treated as exposed surface and the interior surface of the pore is etched based on its angle with the incoming ions. Several pores are opened up on the sidewall creating overhangs that, as we will see, present some problems for further processing, as well as the

*In contrast to deposition, etching processes are subtractive, removing material. However, we include etching in the same category as deposition because both involve the motion of a free surface, but do not include the motion of solid–solid interfaces, such as the evolution of grain boundaries.



6.12 A porous dielectric etching simulation completed using PLENTE, at various times. The process is ion milling using Ar^+ ions, with a 20 to 1 selectivity of the dielectric to the etch stop. Even with this reasonably good selectivity, impressions of the pore structure can still be seen on the etch stop after etching is complete.



6.13 A cross-section view of a simulation of the subsequent vapor deposition of a liner layer into a trench etched into a porous substrate after mask removal. As can be seen by the magnified section, etching these types of pore structures can open up areas that are difficult to deposit onto effectively.

opportunity for trapping process materials. The pore structure here is only 13% pore fraction with a non-connected pore network. Any problems associated with opening up individual pores would be compounded for interconnected pore networks. Additionally, due to the small but finite selectivity of the substrate with respect to the etch stop layer, impressions of the pores in the etched material can still be seen on the etch stop after completion (Fig. 6.12c).

As seen above, etching through porous substrates can create re-entrant surfaces. This phenomenon is problematic for several reasons, among them that the etched trench must be sealed with a barrier layer, just as in the example above of PVD into a dual damascene trench. Figure 6.13 shows a PLENTE PVD simulation into just such an etched structure (after removing the mask). Although the simulation is performed in three dimensions, a cross-section is shown to expose the pore structure. The deposited material

is tantalum, and deposited to a thickness of 165 nm on the flats surrounding this 1 micron deep trench. This thickness corresponds to a minimum thickness on the sidewall of 2.5 nm. However, as can be seen in the magnified section of Fig. 6.13, showing a pore, parts of the interior of the pore can be completely shadowed, having no deposition at all. The failure to create a continuous diffusion barrier can be a severe problem in terms of producing working ICs.

6.4.4 Grain evolution

Although examples of grain evolution do arise during processing, some of the most important examples of evolution involve its impact on reliability. Of particular interest are the phenomena of *curvature-driven grain boundary migration*[31] and *electromigration*[146]. As an example, we look at curvature-driven grain boundary motion. Lifshitz and Slyozov[147] showed theoretically that the grain boundary velocity is related to the pressure difference caused by the curvature, with the grain boundary mobility taken as the proportionality constant. In this case, atoms are transported only the short distance across a grain boundary (on the order of a few lattice constants) moving from one lattice to another where it is energetically favorable due to the curvature. This phenomenon can be thought of as being driven by the increase in the coordination number of adatoms as they move from convex to concave surfaces. Alternatively it can be thought of as being driven by reducing the total area of grain boundary surfaces in a volume, trading surface energy for a lower volume energy.

Using the approach of Turnbull[148], an expression for the migration rate (v) of a small section of grain boundary can be written as follows:

$$v = p\frac{bf\,\Omega}{kT}\exp\left(-\frac{\Delta G}{kT}\right) \qquad [6.4]$$

In this equation, b is the boundary displacement associated with an atomic exchange, Ω is the volume change associated with the atomic exchange, f is the Debye frequency, ΔG is the free energy change associated with the exchange, k is the Boltzmann factor, T is the absolute temperature, and p is the driving 'pressure' for the exchange. The right hand side may be expressed as a product of p and a mobility M, itself a product of a fundamental mobility M_0 and an Arrhenius factor

$$v = pM = pM_0\exp\left(-\frac{Q}{kT}\right) \qquad [6.5]$$

Here Q is the activation enthalpy for atomic motion across the grain boundary. The driving pressure is the heart of understanding the grain boundary migration phenomenon. From an energetics viewpoint, the driving pressue is the ratio of the change in free energy to change in position of the boundary[149]

$$p = \frac{\delta G}{\delta r} \qquad [6.6]$$

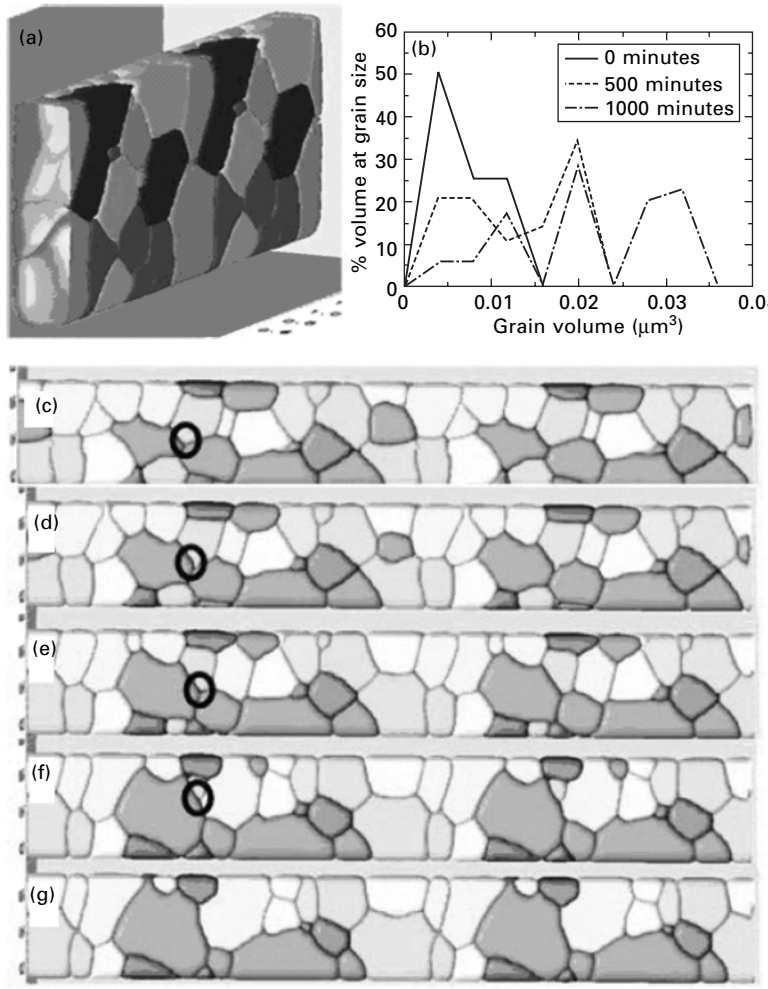
Here δG is the change in energy when a unit area moves a distance δr. For the case of an isotropic grain boundary, in which the energy per unit area (γ) is constant, the Gibbs–Thompson relation $p = \gamma\kappa$ can be applied to write

$$v = M\gamma\kappa = \gamma\kappa M_0 \exp\left(-\frac{Q}{kT}\right) \qquad [6.7]$$

Using this formulation of grain-boundary migration to obtain migration rates and using PLENTE to integrate the motion of grain structure geometry, we have performed a study of coarsening due to curvature-driven grain-boundary migration in micron-sized Cu lines[145]. The parameter values used in the evolution model are taken from the literature (or calculated from known values) and are representative of copper. Initial grain structures were unrelaxed polycrystalline Cu structures grown in simulation using an isotropic ELD model[136] (see Section 6.3.3). There are 30 grains per periodic length in the initial structure, as shown in Fig. 6.14a. As the simulation progresses (Figs 6.14c to g), grain boundaries move and the structure coarsens, with larger grains eventually 'swallowing' smaller grains. (The small black circles in Figs 6.14c to f are discussed in Section 6.5.3) As each grain is swallowed, the number of grains in the simulation decreases by one, until at 1800 minutes of simulation time, there are only 13 grains left in the line.

The phenomenon of larger grains growing at the expense of smaller grains has the effect of reducing intermediate sized grains from the simulation at first, splitting the distribution into increased numbers of large and small grains, as seen in the grain size distribution curves in Fig. 6.14b. As the simulation progresses further, the balance continues to shift to the larger grains, until at 1000 minutes, the number of grains is reduced to the point that the statistics of the distribution break down due to sample size.

Fair agreement was found when the grain size distributions at selected times were compared against experimental data on similarly processed lines[31]. On the other hand, experimental Cu films indicate highly faceted grains developing, as shown in Fig. 6.15; the model needs to be improved. Curvature-driven grain boundary motion is driven by interfacial energy, much like the coarsening of soap froths is driven by the surface tension of the walls between bubbles[108]. However, unlike the case of air trapped in soap bubbles, the crystalline material that makes up grains has an orientation, so not all interfacial planes have the same surface energy. Because anisotropic surface energies, as functions of grain boundary parameters, are not well known, the model used in the simulation that yielded Fig. 6.14 used an isotropic energy and thus does not predict facet formation.

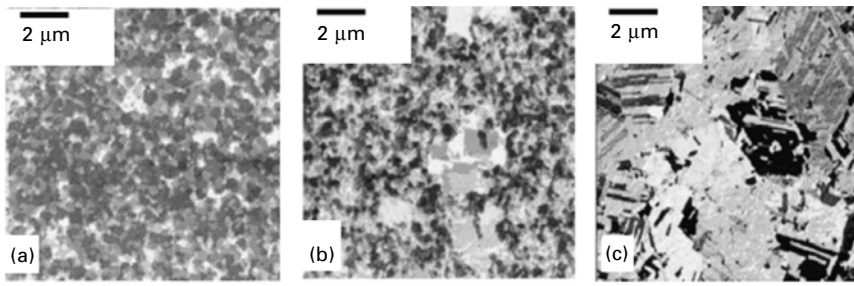


6.14 (a) Starting structure for a curvature-driven grain boundary simulation. (b) grain-size distributions at various times throughout the simulation. (c, d, e, f, g) Snapshots taken at selected times during the simulation[31,145].

6.5 Opportunities

6.5.1 Overview

Although much of the formalism for grain continuum simulation is in place, there are still challenges to meet and opportunities to address new phenomena on several fronts. Some of the biggest challenges are technological and have been touched upon briefly in the previous sections. These are software needs and include the ability do the following:



6.15 Electrochemically deposited Cu films (1 μm thick), after self-annealing at room temperature at (a) 0.75, (b) 8, and (c) 48 hrs[31].

1. Quickly and robustly extract consistent interfaces from the over-specified system represented by our MLS approach.
2. Construct volume meshes in structures of complex, non-manifold topologies in order to calculate process and evolution models on them.
3. Rapidly and accurately calculate and efficiently store the large numbers of view factors needed for ballistic transport calculations

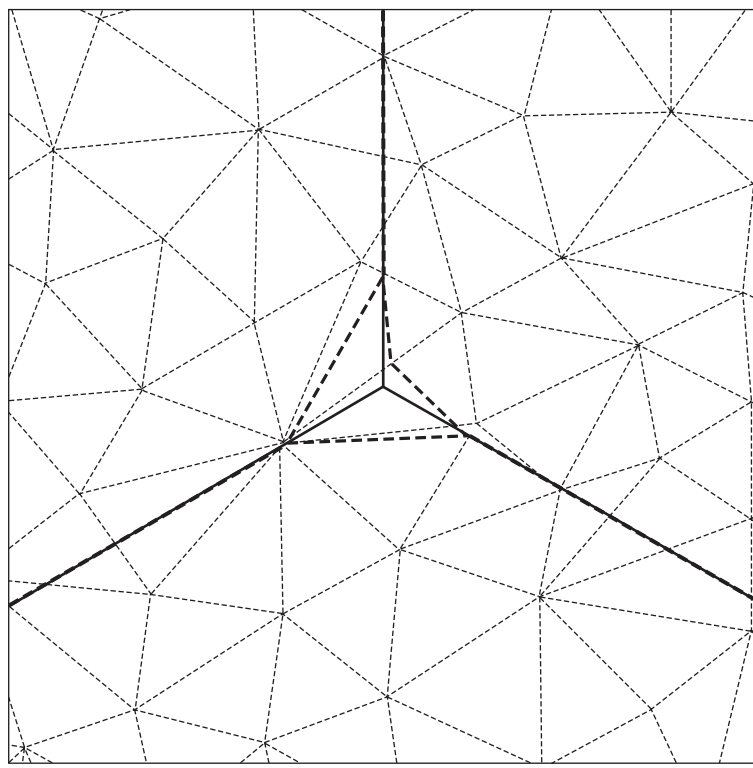
In addition to the above GC/PLENTE-focused software opportunities and introducing other process models (both inside and outside of microelectronics), perhaps the opportunity with the highest risk and highest reward is interfacing PLENTE with atomistic simulators. Such a marriage would allow more accurate information regarding local physical phenomena, e.g. surface and interface diffusivities and energies, to be introduced into grain-continuum simulations.

6.5.2 Software opportunities

Efficient and robust meshers are required by both PLENTE and many of the supplemental codes that we desire to use, such as finite-element solvers. Current mesh generation methods are labor intensive and can account for half of the time and cost of an analysis. Shephard and co-workers[150–154] have developed techniques to automatically generate and adapt meshes for general non-manifold 3D domains. These procedures are well suited to dealing with the multi-material domains to be considered in this research. Parallel versions capable of generating meshes with hundreds of millions of elements are on the horizon and preliminary versions are under development[151,152].

One of the issues that the MLS method faces is that of *consistent extraction*. In the MLS method, each phase is associated with a level-set scalar, with the zero level indicating the boundary of that phase. Because an interface is, by definition, the boundary of two phases, the exact position of an interface is determined by more than one zero level. When these zero levels agree within

machine precision, the interface is single valued. However, when three or more phases (two or more interfaces) come together (a third- or higher-order junction), it is not uncommon for the precise position of the zero levels to diverge, as shown in Fig. 6.16 for a 2D case. This is not a problem of accuracy, as the errors are within the previously established bounds, i.e. below the resolution of the mesh. The problem is that the representation errors are not identical in all of the representations of the interfaces. Upon extraction of the interfaces, the resulting explicit representation can be left with regions of space that do not correspond to any particular phase, including not belonging to a 'vacuum phase'. Also, the discretizations of the involved phase boundaries may differ, making an explicit 'knitting together' of the problematic regions very difficult. These extractions are referred to as *inconsistent*. This situation is troublesome because, when the explicit structure

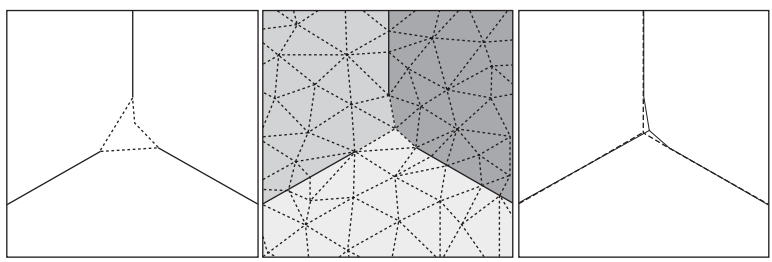


6.16 An explicit set of interfaces (black solid lines) can be represented as contours in scalar fields, themselves represented on a finite-element mesh (light dotted lines). However the re-extracted interfaces (dashed lines) do not come together at triple points and higher order junctions, leaving regions that do not correspond to any physical situation.

is needed by a process simulation code, applying boundary conditions to inconsistent parts of the structure will be problematic.

To circumvent this problem, Bloomfield *et al.*[136] proposed a *voxel extraction* (VE) method, in which the extraction from the implicit representation to an explicit surface mesh first passes through an explicit volume representation. This method gives consistent extractions as well as producing volume meshes of all the phases, but the results are limited to 'Manhattan' geometries, i.e. are constructed of faces that lie only in a limited number of planes, resembling a city skyline. The same authors later proposed the *conformal voxel extraction* (CVE) method[155] that avoids the Manhattan geometry limitations of the original VE method. In the CVE method, a volume mesher is used to create a space-filling mesh that meets certain internal constraints, making use of information from the inconsistent extraction. The elements in this mesh are then used to map the implicit representation to an explicit volume representation, and finally an explicit, consistent surface representation is extracted from the explicit volume representation. In Fig. 6.17, the CVE method is applied in two dimensions to the problematic structure from Fig. 6.16. As a side benefit of this extraction method, the space-filling voxelation mesh is a volume mesh (in 3D problems) for the explicit extraction, addressing the software need listed above of completing the (perhaps problematic) task of creating an interface fitted mesh for the extraction.

The third technological challenge listed becomes an issue in low pressure simulation, when applying the BTRM (See Section 6.4.2). It should be noted that, although the application of the BTRM to simulate the evolution of 3D surfaces is a straightforward extension of its application to 2D surfaces, the problem can quickly become huge. For non-unity sticking factors, Equation



6.17 The conformal voxel extraction procedure, as applied to the implicit representation shown in Fig. 6.16. First, consistent elements from the directly extracted sets (left) are identified (solid lines). Next, a space filling mesh (middle) is created using these elements as internal constraints and each voxel is assigned to a phase (shading). Finally, the conformal voxel extraction (solid line, right) is derived by comparing neighboring voxels, and compared to the structure used to initialize the implicit representation (dashed line, right).

[6.2] is solved iteratively, using a fixed-point iteration in the fluxes (and surface coverages when appropriate). To solve this problem, the view factor matrix q is required for whatever geometry is being considered. It is this view factor matrix which places a computational limit on the size of problems that can be reasonably addressed. Note that this matrix has size N^2 where N is the number of points sampled on the surface, often corresponding to the number of patches or triangles (in 3D) or segments (in 2D) that are used to represent the surface. The number of non-zero entries in this matrix is highly dependent upon the shape of the surface being represented. The problem of determining and representing view factors efficiently is beyond the scope of this chapter.

Recently a Boltzmann transport equation (BTE) based alternative to solving low-pressure transport and reaction models has been presented[156,157]. Gobbert *et al.* introduced the kinetic transport and reaction model (KTRM), formulated as a discontinuous Galerkin problem. BTEs are solved over a domain as a boundary value problem using the heterogeneous chemical reaction rate expressions and re-emission models as boundary conditions. The solution to the KTRM is the particle density functions at each point in the system, which contain the information about transport in the system. Two strong advantages of the KTRM, for systems for which it applies, are that collisional flows can be studied and process transients can be followed. However, solutions require discretizing the empty space between the surfaces as well as the surfaces. For collisonless transport, this is a computational expense not required by the BTRM. For cases in which both the BTRM and KTRM are valid, it is unclear whether the KTRM (recently developed) or the BTRM will be preferred. With their different strengths, it is likely that both may find a long-term place in 3D vapor deposition simulation.

6.5.3 Interface between atomistic and grain-continuum simulators

To get atomistically derived information into PLENTE, parts of the simulation domain need to be represented atomistically. We term the flip side of the encapsulation procedure discussed in Section 6.4.2 as *atomation*. We have implemented an atomation functionality in PLENTE that can take a set of GC representations as input. It constructs a set of atomic positions that represent the occupied points for the lattices associated with the grain(s) in a neighborhood of a requested sample point. During the atomation procedure, a distribution of defects such as vacancies or substitutional impurities can be imposed upon the atomated structure, based on information stored in the GC representation. One approach to incorporating the effect of atomic scale phenomena into PLENTE, or another GC implementation, is indicated by the black rings in the first four ‘still shots’ in Fig. 6.14c through Fig. 6.14f.

At any one of these time steps, PLENTE would call support software to identify key locations, such as these points on grain boundaries. These points (small volumes on the GC scale) would be atomated and passed to an atomistic simulator, e.g. ADEPT, to get improved rate parameters or actual rates. In Fig. 6.14c, the black ring identifies a higher-order grain junction where atomistically-derived information might be needed. As time progresses, the junction indicated by the black ring in Fig. 6.14 moves in space, and finally disappears as one of the grains is consumed. The decision making code clearly needs to be adaptive.

A large portion of any effort to interface GC with atomistic simulators efficiently will be to provide the interface code with the ability to make decisions and guide the atomistic simulations using information from the GC simulation. The code that decides when atomistic information is needed should also manage, cache, and interpolate simulation results using a database, so that multiple requests for similar information can be filled without repeated atomistic computation. Ideally, if a similar atomistic simulation had already been run then the information desired should be looked up or interpolated from previous results. A large part of such an interface code would be the ability to decide if existing information in its database is similar enough to use in a 'current' request for atomistically derived information. The design of the database itself would be important for both efficiency and accuracy[158,159].

This marriage of GC (PLENTE) and atomistic (ADEPT) codes will allow concurrent simulation from the atomistic to the meter scales. For example, PLENTE interacts with programs that deal with larger scales (e.g. COMSOL[160] for structural or equipment scale models), taking inputs such as fluxes, reactant compositions, or mechanical forces from these simulations (simultaneous solution is preferred[143,161,162]). ADEPT takes its inputs from models, tools (e.g. VASP), and previous results.

In addition to spanning a wide range of the spatial scales of interest to materials scientists, process engineers, and product engineers, this approach may help to deal with timescale issues associated with bridging discrete and continuum simulators. PLENTE uses structural and rate information at one time to predict small changes in structure, which occur over a time step that in turn depends upon the rate of structural evolution[136]. PLENTE then evolves the structure under study, and re-computes the structure and rates for the next time step. Consider again a specific time step in a PLENTE evolution simulation, such as one shown in Fig. 6.14. To determine how the system evolves during the next time step, PLENTE seeks structural and rate information. PLENTE calls code that decides what part of that information needs to be atomistically derived, allocates the needed resources, then gathers and returns the information to PLENTE. PLENTE then takes a time step, and the process is repeated. The short timescales typical of atomistic simulations are appropriate for such an application; i.e. instantaneous rates and structures

needed at a specific time during the solution of transient partial differential equations.

One of the research efforts needed to develop such an environment is to determine the level of detail that should be represented in the GC, i.e. what should be represented atomistically, for a given material and process. This balance will depend upon the targeted film characteristic or performance property. As more experimental information on grain boundaries becomes available, such a combined GC/atomistic approach will help determine how to represent grain-boundaries to the problem-dependent level needed.

6.6 References

1. G.E. Moore, 'Cramming more components onto integrated circuits', *Electronics* **38**(8), 114–117 (1965).
2. The International Technology Roadmap for Semiconductors, *http://public.itrs.net/*, accessed February 2006.
3. S. Wolf, *Microchip Manufacturing*, Sunset Beach, CA: Lattice Press, 2003.
4. R. Pierret, *Advanced Semiconductor Fundamentals*, Upper Saddle River, NJ: Prentice Hall, 2003.
5. S. Wolf, *Silicon Processing for the VLSI Era, Vol. 4: Deep-Submicron Process Technology*, Sunset Beach, CA: Lattice Press, 2002.
6. S.A. Campbell, *The Science and Engineering of Microelectronic Fabrication*, Oxford: Oxford University Press, 2nd Edn, Oxford, 2001.
7. T.S. Cale, T.P. Merchant, L.J. Borucki and A.H. Labun, 'Topography simulation for the virtual wafer fab', *Thin Solid Films* **365**(2), 152–175 (2000).
8. K.C. Saraswat, 'Programmable factory IC manufacturing for the 21st century', in *IEEE/SEMI International Semiconductor Manufacturing Science Symposium*, IEEE, 1993, 2–6.
9. R.S. Blewer, personal communication to T.S. Cale 2003.
10. *Modeling of Film Deposition for Microelectronics Applications*. Thin Films, ed. S. Rossnagel and A. Ulman, Vol. 22, Academic Press, San Diego, CA, 1996.
11. T.S. Cale, B.R. Rogers, T.P. Merchant and L.J. Borucki, 'Deposition and etch processes: continuum film evolution in microelectronics', *Comput. Mater. Sci.* **12**, 333–353 (1998).
12. J.D. Plummer, M.D. Deal, and P.B. Griffin, *Silicon VLSI Technology: Fundamentals, Practice, and Modeling*, Upper Saddle River, NJ: Prentice Hall, 2000.
13. T.S. Cale, M.O. Bloomfield, D.F. Richards, K.E. Jansen, and M.K. Gobbert, 'Integrated multiscale process simulation', *Comput. Mater. Sci.* **23**(1–4), 3–14 (2002).
14. M. Lane, 'Interface fracture', in *Ann. Rev. Mater. Res.* **33**, 29–54 (2003).
15. C.S. Litteken, S. Strohband and R.H. Dauskardt, 'Residual stress effects on plastic deformation and interfacial fracture in thin-film structures', *Acta Mater.* **53**(7), 1955–1961 (2005).
16. M.W. Lane, J.M. Snodgrass and R.H. Dauskardt, 'Environmental effects on interfacial adhesion', *Microelectron Reliab.* **41**(9–10), 1615–1624 (2001).
17. P. Cottrell and W. Haensch, 'Device/Process Simulation, from atoms to applications', presented at Rensselaer Polytechnic Institute, 2006.

18. N. Ghoniem, E. Busso and H. Huang, 'Nano/micro Mechanics of Materials', in *Phil. Mag*. 31–34, 2003, special issue.
19. Z.X. Guo, H. Huang, S. Iwata, O. Pankratov and S. Yip, 'Multiscale Materials Modeling', in *Comput. Mater. Sci*. 23, 2002, special issue.
20. H. Huang, *Computational Nanomechanics of Materials: Cross Fertilization of Physics, Chemistry, Materials Science, Mechanics, and Computation*, 2004, Workshop Report; submitted to National Science Foundation and Army Research Office and circulated among workshop participants.
21. W. Steinhoegl, G. Schindler, G. Steinlesberger, M. Traving, and M. Engelhardt, 'Scaling laws for the resistivity increase of sub-100 nm interconnects', in *IEEE International Conference on Simulation of Semiconductor Processes and Devices*, IEEE, 2003, 27–30.
22. J.A. Davis, R. Venkatesan, A. Kaloyeros, M. Beylansky, S.J. Souri, K. Banerjee, K.C. Saraswat, A. Rahman, R. Reif and J.D. Meindl, 'Interconnect limits on gigascale integration (GSI) in the 21st century', *Proc. IEEE* **89**(3), 305–324 (2001).
23. P. Kapur, J.P. McVittie and K.C. Saraswat, 'Realistic copper interconnect performance with technological constraints' in *Proc. IEEE 2001 International Interconnect Technology Conference*, IEEE, 2001, 233–235.
24. A. Bietsch and B. Michel, 'Size and grain-boundary effects of a gold nanowire measured by conducting atomic force microscopy', *Appl. Phys. Lett.* **80**(18), 3346–3348 (2002).
25. V. Sukharev, 'Physically-based simulation of electromigration induced failures in copper dual-damascene interconnect', in *5th International Symposium on Quality Electronic Design*, IEEE, 2004, 225–230.
26. E. Zschech, E. Langer and M.A. Meyer, 'Failures in copper interconnects – localization, analysis and degradation mechanisms', in *IEEE International Symposium on the Physical and Failure Analysis of Integrated Circuits*, IEEE, 2003, 37–44.
27. J.R. Lloyd, M.R. Lane, X.-H. Liu, E. Liniger, T.M. Shaw, C.-K. Hu and R. Rosenberg 'Reliability challenges with ultra-low k interlevel dielectrics', *Microelectron. Reliab.* **44**(9–11), 1835–41.
28. P. Andricacos, 'Electrochemical technology in microelectronics', *IBM J. Res. Dev*. **49** (1), editor's note (2005).
29. C. Lingk, M.E. Gross and W.L. Brown, 'Texture development of blanket electroplated copper films', *J. Appl. Phys.* **87**(5), 2232–2236 (2000).
30. K. Ueno, T. Ritzdorf and S. Grace, 'Seed layer dependence of room-temperature recrystallization in electroplated copper films', *J. Appl. Phys*. **86**(9), 4930–4935 (1999).
31. S. Lagrange, S.H. Brongersma, M. Judelewicz, A. Saerens, I. Vervoort, E. Richard, R. Palmans and K. Maex, 'Self-annealing characterization of electroplated copper films', *Microelectron. Eng.* **50**(1–4), 449–457 (2000).
32. J.M. Howe, *Interfaces in Materials: Atomic Structure, Thermodynamics and Kinetics of Solid-Vapor, Solid-Liquid and Solid-Solid Interfaces*, New York: Wiley-Interscience, 1997.
33. D.M. Saylor, J. Fridy, B.S. El-dasher, K.-Y. Jung and A.D. Rollett, 'Statistically representative three-dimensional microstructures based on orthogonal observation sections', *Metall. Mater. Trans. A* **35A**(7), 1969–1979 (2004).
34. H.L. Wei, H. Huang, C.H. Woo, R.K. Zheng, G.H. Wen and X.X. Zhang, 'Development of <110> texture in copper thin films', *Appl. Phys. Lett.* **80**(13), 2290–2292 (2002).

35. E. Lifshin, J. Evertsen, E. Principe and J. Friel, 'Three dimensional imaging of microelectronic devices using a cross beam FIB', in *ISTFA 2004 – Proceedings of the 30th International Symposium for Testing and Failure Analysis*, Materials Park, OH: ASM International, 429–435.
36. G.H. Gilmer and P. Bennema, 'Simulation of crystal growth with surface diffusion', *J. Appl. Phys.* **43**(4), 1347–1360 (1972).
37. F. Abraham and G. White, 'Computer simulation of vapor deposition on two-dimensional lattices', *J. Appl. Phys.* **41**, 1841–1849 (1970).
38. B.W. Dodson, 'Molecular dynamics modeling of vapor-phase and very-low-energy ion-beam crystal growth processes', *CRC Crit. Rev. Solid State Mater. Sci.* **16**(2), 115–130 (1990).
39. P.K. Swaminathan, C.S. Murthy and M.J. Redmon, 'Computer simulations of epitaxial growth', *Phys. Rev. B* **B39**, 4541–4547 (1989).
40. S.K. Dew, T. Smy and M.J. Brett, 'Simulation of elevated temperature aluminum metallization using SIMBAD', *IEEE Trans. Electron Devices* **39**(7), 1599–1606 (1992).
41. L.J. Friedrich, S.K. Dew, M. Brett and T. Smy, 'Thin film microstructure modelling through line-segment simulation', *Thin Solid Films* **266**(1), 83–88 (1995).
42. T. Smy, R.N. Tait and M.J. Brett, 'Ballistic deposition simulation of via metallization using a quasi-three-dimensional model', *IEEE Trans. Computer-Aided Design of Integrated Circuits and Systems* **10**(1), 130–135 (1991).
43. K.L. Westra, T. Smy and M.J. Brett, 'Simulation by ballistic deposition of local density variation and step coverage for via metallization', *IEEE Electron Device Lett.* **10**(5), 198–199 (1989).
44. D.G. Coronell and K.F. Jansen, 'Monte Carlo simulations of very low pressure chemical vapor deposition', *J. Comput.-Aided Mater. Des.* **1**(1), 3–26 (1993).
45. A.F. Voter, 'Hyperdynamics: accelerated molecular dynamics of infrequent events', *Phys. Rev. Lett.* **78**(20), 3908–3911 (1997).
46. P. Meakin, P. Ramanlal, L.M. Sander and R.C. Ball, 'Ballistic deposition on surfaces', *Phys. Rev. A* **34**(6), 5091–5103 (1986).
47. H. Huang and G.H. Gilmer, 'Multi-lattice Monte Carlo model of thin films', *J. Comput.-Aided Mater. Des.* **6**(2–3), 117–127 (1999).
48. M. Basham, P.A. Mulheran and F. Montalenti, 'Diffusion and stability of small vacancy clusters on Cu(100) – a simulation study', *Surf. Sci.* **565**(2–3), 289–299 (2004).
49. D. Moldovan, D. Wolf and S.R. Phillpot, 'Linking atomistic and mesoscale simulations of nanocrystalline materials: quantitative validation for the case of grain growth', *Phil. Mag.* **83**(31–34), 3643–3659 (2004).
50. J. Wang and H. Huang, 'Shockley partial dislocations to twin: another formation mechanism and generic driving force', *Appl. Phys. Lett.* **85**(24), 5983–5985 (2004).
51. S.J. Liu, H. Huang and C.H. Woo, 'Schwoebel-Ehrlich barrier: from two to three dimensions', *Appl. Phys. Lett.* **80**(18), 3295–3297 (2002).
52. L.G. Zhou and H. Huang, 'Are surfaces elastically softer or stiffer?' *Appl. Phys. Lett.* **84**(11), 1940–1942 (2004).
53. T. Smy, R.V. Joshi and S.K. Dew, 'Full 3D microstructural simulation of refractory films deposited by PVD and CVD', in *Advanced Metallization Conference 1999 (AMC 1999)*, Warrendale, PA: Materials Research Society 2000, 669–675.
54. F. Baumman and G. Gilmer, '3D modeling of sputter and reflow processes for interconnect metals', in *International Electron Devices Meeting. Technical Digest*, 1995, 89–92.

55. D.J. Srolovitz, G.S. Grest and M.P. Anderson, 'Computer simulation of grain growth. V. Abnormal grain growth', *Acta Metall*. **33**(12), 2233–2247 (1985).
56. E.A. Holm, N. Zacharopoulos and D.J. Srolovitz, 'Nonuniform and directional grain growth caused by grain boundary mobility variations', *Acta Mater.* **46**(3), 953–964 (1998).
57. B. Radhakrishnan and T. Zacharia, 'Simulation of curvature-driven grain growth by using a modified Monte Carlo algorithm', *Metall. Mater. Trans. A* **26A**(1), 167–180 (1995).
58. Q. Yu and S.K. Esche, 'Three-dimensional grain growth modeling with a Monte Carlo algorithm', *Mater. Lett.* **57**(30), 4622–4626 (2003).
59. T. Fordell, P. Salo and M. Alatalo, 'Self-diffusion on fcc (100) metal surfaces: Comparison of different approximations', *Phys. Rev. B* **65**(23), 233408/1–4 (2002).
60. C.L. Liu, J.M. Cohen, J.B. Adams and A.F. Voter, 'EAM study of surface self-diffusion of single adatoms of fcc metals Ni, Cu, Al, Ag, Au, Pd, and Pt', *Surf. Sci.* **253**(1–3), 334–344 (1991).
61. S. Toyoshima, T. Kawamura, S. Nishida and A. Ichimiya, 'Surface diffusion during decay of nano-island on Si(100) at high temperature', *Surf. Sci.* **572**(1), 84–92 (2004).
62. J.B. Adams, Z. Wang and Y. Li, 'Modeling Cu thin film growth', *Thin Solid Films* **365**(2), 201–210 (2000).
63. A. Challa, J. Drucker and T.S. Cale, 'Morphology of Ag islands grown on GaAs(110) at low coverage: Monte Carlo simulations', in *Evolution of Epitaxial Structure and Morphology*, eds A. Zangwill, D. Jesson, D. Chambliss and R. Clarke, Warrendale, PA: Materials Research Society, 1996, 103–108.
64. S.V. Ghaisas and A. Madhukar, 'Role of surface molecular reactions in influencing the growth mechanism and the nature of nonequilibrium surfaces: A Monte Carlo study of molecular-beam epitaxy', *Phys. Rev. Lett.* **56**(10), 1066–1069 (1986).
65. D. Kaschiev, J.P.V.d. Eerden and C. Van Leeuwen, 'Transition from island to layer growth of thin films: A Monte Carlo simulation', *J. Cryst. Growth* **40**, 47–58 (1977).
66. A.I. Michaels, G.M. Pound and F.F. Abraham, 'Monte Carlo simulation of the kinetics of heterogeneous nucleation', *J. Appl. Phys*. **45**(1), 9–16 (1974).
67. R.A. Outlaw and J.H. Heinbockel, 'Potential energy scaling Monte Carlo simulation of thin film nucleation and growth', *Thin Solid Films* **108**(1), 79–86 (1983).
68. J.P. van der Eerden, G. van Leeuwen, P. Bennema, W.L. van der Kruk and B.P.T. Veltman, 'Crystal growth: a comparison of Monte Carlo simulation nucleation and normal growth theories', *J. Appl. Phys.* **48**(6), 2124–2130 (1977).
69. J. Salik and H.M. Ghose, *Computer Simulation of Thin Film Nucleation and Growth – The Volmer-Weber Case*, NASA Technical Report NASA-TM-86968, 1985.
70. A. Challa, 'Microstructural evolution during the early stages of deposition: Monte Carlo simulations', Arizona State University, Phoenix, AZ, 1996.
71. A. Challa, T.S. Cale and J. Drucker, 'Growth morphology of Ag islands on GaAs (110) at low coverage: Monte Carlo simulations', in *Evolution of Thin Film and Surface Structure and Morphology*, eds B.G. Demczyk, E. Garfunkel, B.M. Clemens, E.D. Williams and J.J. Cuomo, Warrendale, PA: Materials Research Society 101–106.
72. J.A. Venables, G.D.T. Spiller and M. Hanbuecken, 'Nucleation and growth of thin films', *Rep. Prog. Phys.* **47**(4), 399–459 (1984).
73. R.A. Levy, M.L. Green and P.K. Gallagher, 'Characterization of LPCVD Aluminum For VLSI Processing', *J. Electrochem. Soc*. **131**(9), 2175–2182 (1984).

74. D. Yang, J. Hong, D.F. Richards and T.S. Cale, 'Nucleation and film growth during copper chemical vapor deposition using the precursor Cu(TMVS)(hfac)', *J. Vac. Sci. Technol. B* **20**(2), 495–506 (2002).
75. B.R. Rogers, Z. Song, R. Geil, V. Wahlig and D. Crunkleton, 'UHV-CVD of ZrO_2 for deeply scaled CMOS gate dielectric applications', 2003, presented at AIChE 2003 Annual Meeting, 11/19/2003.
76. W.A.P. Claassen and J. Bloem, 'Nucleation of CVD silicon on SiO_2 and Si_3N_4 substrates. I. The SiH_4-HCl-H_2 system at high temperatures', *J. Electrochem. Soc.* **127**(1), 194–202 (1980).
77. W.A.P. Claassen and J. Bloem, 'Nucleation of CVD silicon on SiO_2 and Si_3N_4 substrates. II. The SiH_2Cl_2-H_2-N_2 system', *J. Electrochem. Soc.* **127**(8), 1836–1843 (1980).
78. W.A.P. Claassen and J. Bloem, 'Nucleation of CVD silicon on SiO_2 and Si_3N_4 substrates', *J. Electrochem. Soc*. **128**(6), 1353–1359 (1981).
79. J. Bloem, 'Nucleation and growth of silicon by CVD', *J. Cryst. Growth* **50**(3), 581–604 (1980).
80. D.E. Kotecki and I.P. Herman, 'A real time Monte Carlo simulation of thin film nucleation in localized-laser chemical vapor deposition', *J. Appl. Phy.* **64**, 4920–4942 (1988).
81. J. Bloem and W.A.P. Claassen, 'Nucleation and growth of silicon films by chemical vapour deposition', *Philips Tech. Rev.* **41**(2), 60–69 (1983).
82. G.H. Gilmer, H. Huang, T.D. de la Rubia, J. Dalla Torre and F. Baumann, 'Lattice Monte Carlo models of thin film deposition', *Thin Solid Films* **365**(2), 189–200 (2000).
83. G.H. Gilmer, H. Huang and C. Roland, 'Thin film deposition: Fundamentals and modeling', *Comput. Mater. Sci.* **12**(4), 354–380 (1998).
84. H. Huang and G. Gilmer, 'Texture competition during thin film deposition-effects of grain boundary migration', *Comput. Mater. Sci.* **23**, 190–196 (2002).
85. H. Huang, G.H. Gilmer and T.D. de la Rubia, 'An atomistic simulator for thin film deposition in three dimensions', *J. Appl. Phys.* **84**, 3636–3649 (1998).
86. H. Huang and G.H. Gilmer, 'Atomistic simulation of texture competition during thin film deposition', *J. Comput. Aided Mater. Des.* **7**(3), 203–216 (2000).
87. H. Huang and L.G. Zhou, 'Atomistic simulator of polycrystalline thin film deposition in three dimensions', *J. Comput.-Aided Mater. Des.* **11**, 59–74 (2004).
88. W.G. Oldham, S.N. Nandgaonkar, A.R. Neureuther and M. O'Toole, 'A general simulator for VLSI lithography and etching processes. I. Application to projection lithography', *IEEE Trans. Electron Devices* **ED-26**(4), 717–722 (1979).
89. J.P. McVittie, J.C. Rey, A.J. Bariya, M.M. IslamRaja, L.Y. Cheng, S. Ravi and K.C. Saraswat, 'SPEEDIE. A profile simulator for etching and deposition', in *Advanced Techniques for Integrated Circuit Processing*, eds T. Bondur and T.R. Turner, Bellingham, WA: SPIE, 1991, 126–138.
90. Y.-W. Kim, J. Moser, I. Petrov, J.E. Greene and S.M. Rossnagel, 'Directed sputter deposition of AlCu: film microstructure and microchemistry', *J. Vac. Sci. & Technol. A* **12**(6), 3169–3175 (1994).
91. T.S. Cale, V. Mahadev, Z. Tang, G. Rajagopalan and L.J. Borucki, 'Topography Evolution During Semiconductor Processing', in *Plasma Processing of Semiconductors*, ed. F. Williams, NATO ASI Series E, Vol. 336, Dordrecht: Kluwer Academic Publishers, 1997, 109–124.
92. G. Rajagopalan, V. Mahadev and T.S. Cale, 'Surface evolution during semiconductor processing', *VLSI Design* **6**(1–4), 379–384 (1998).

93. EVOLVE, T.S. Cale, EVOLVE, the low pressure module of EVOLVE was written by T. S. Cale, 1990–2006.
94. T.S. Cale and V. Mahadev, 'Modeling of film deposition for microelectronic applications', in *Modeling of Film Deposition for Microelectronics Applications*, eds S. Rossnagel and A. Ulman, San Diego, CA: Academic Press, 1996, 175–276.
95. A.H. Labun, H.K. Moffat and T.S. Cale, 'Mechanistic feature-scale profile simulation of SiO_2 low-pressure chemical vapor deposition by tetraethoxysilane pyrolysis', *J. Vac. Sci. Technol. B* **18**(1), 267–278 (2000).
96. T.S. Cale, G.B. Raupp, M.B. Chaara and F.A. Shemansky, 'Reaction mechanism discrimination using experimental film profiles in features', *Thin Solid Films* **220**(1–2), 66–72 (1992).
97. Y.H. Im, M.O. Bloomfield, C.P. Sukam, J.A. Tichy, T.S. Cale and J. Seok, 'Integrated multiscale multistep process simulation', in *IEEE International Conference on Simulation of Semiconductor Processes and Devices*, IEEE, 2003, 307–310.
98. M.O. Bloomfield, L.J. Borucki, Y.H. Im and T.S. Cale, 'Integrated multi-step process simulation with chip-scale structures', in *Advanced Metallization Conference 2003 (AMC 2003)*, Warrendale, PA: Materials Reseach Society, 2004, 775–779.
99. S. Osher and J.A. Sethian, 'Fronts Propagating with Curvature Dependent Speed: Algorithms Based on Hamilton-Jacobi Formulations', *J. Comput. Phys.* **79**, 12–49 (1988).
100. S. Osher and R. Fedkiw, 'Level set methods: an overview and some recent results', *J. Comput. Phys.* **169**(2), 463–502 (2001).
101. J.A. Sethian, *Level Set Methods and Fast Marching Methods: Evolving Interfaces in Computational Geometry, Fluid Mechanics, Computer Vision, and Materials Science*, 2nd edn, Cambridge Monographs on Applied and Computational Mathematics, eds P.G. Ciarlet, A. Iserles, R.V. Kohn and M.H. Wright, Cambridge: Cambridge University Press, UK, 1999.
102. ASSET, L.J. Borucki, ASSET is a 3D simulation environment written at Motorola.
103. L.J. Borucki, T.P. Merchant and T.S. Cale, 'Low Pressure Service for ASSET Simulation Environment', LPS-ASSET is a 3D ballistic transport module for ASSET written at Motorola.
104. T.S. Cale, T.P. Merchant and L.J. Borucki, 'The roles of "3d/2d" and "3d/3d" topography simulators in process development: model development and process integration', in *Advanced Metallization Conference in 1998 (AMC 1998)*, Warrendale, PA: Materials Research Society, 1999, 737–741.
105. T.S. Cale, T.P. Merchant and L.J. Borucki, 'Roles of '3d/2d' and '3d/3d' topography simulators in virtual wafer fabs' in *Semiconductor Process and Device Performance Modelling*, eds ST. Dunham and J.S. Nelson, Warrendale, PA: Materials Research Society, 1998, 201–212.
106. H. Emmerich, *The Diffuse Interface Approach in Materials Science: Thermodynamic Concepts and Applications of Phase-Field Models*, Lecture Notes in Physics, Vol. M, New York: Springer-Verlag, 2003.
107. W. Miller, I. Rasin and F. Pimentel, 'Growth kinetics and melt convection', *J. Cryst. Growth* **266**(1–3), 283–288 (2004).
108. H.J. Frost and C.V. Thompson, 'Computer simulation of microstructural evolution in thin films', *J. Electron. Mater.* **17**(5), 447–458 (1988).
109. J. Zhang and J.B. Adams, 'FACET: a two dimensional simulator of polycrystalline thin film growth', in *Multiscale Modeling of Materials*, eds V.V Bulatov, E. Kaxinas,

N. Ghoniem, R. Phillips and T. Diazde la Rubia, Warrendale, PA: Materials Research Society 2000, Z10.1.1–6.
110. A.R. Roosen and W.C. Carter, 'Simulations of microstructural evolution: anisotropic growth and coarsening', *Physica A* **261**(1–2), 232–247 (1998).
111. A. Kuprat, D. George, E. Linnebur, H. Trease and R.K. Smith, 'Moving adaptive unstructured 3-D meshes in semiconductor process modelling applications', *VLSI Design* **6**(1–4), 373–378 (1998).
112. LaGriT User's Manual, *http://lagrit.lanl.gov/*, accessed February 2006.
113. D. George, N. Carlson, J.T. Gammel and A. Kuprat, '3D modeling of metallic grain growth', in *Proceedings of Modeling and Simulation of Microsystems, Semiconductors, Sensors and Actuators*, Boston: Computational Publications, 1999, 463–466.
114. A. Kuprat, *Inspecting and Repairing Physical Topology in a Moving Grid Grain Growth Simulation*, Los Alamos National Laboratory, LANL Report LA-UR-00-3475, 2002.
115. A. Kuprat and D. George, 'Maintaining tetrahedral mesh quality in response to time-dependent topological and geometrical deformation', in *Proc. 6th Int. Conf. on Num. Grid Generation in Comp. Field Simulations*, 1998, p. 589–598.
116. A. Kuprat, A. Khamayseh, D. George and L. Larkey, 'Volume conserving smoothing for piecewise linear curves, surfaces, and triple lines', *J. Computat. Phys.* **172**, 99–118 (2001).
117. L. Granasy, T. Pusztai and J.A. Warren, 'Modelling polycrystalline solidification using phase field theory', *J. Phys. Condensed Matter* **16**(41), 1205–1235 (2004).
118. L.-Q. Chen and Y. Wang, 'Continuum field approach to modeling microstructural evolution', *JOM* **48**(12), 13–18 (1996).
119. J.W. Cahn, 'On spinodal decomposition', *Acta Metall.* **9**(9), 795–801 (1961).
120. P.C. Hohenberg and B.I. Halperin, 'Theory of dynamic critical phenomena', *Rev. Mod. Phys.* **49**(3), 435–479 (1977).
121. G. Lu and E. Kaxiras, 'Overview of multiscale simulations of materials', in *Handbook of Theoretical and Computational Nanotechnology*, eds M. Rieth and W. Schommers, Los Angeles, CA: American Scientific Publishers, 2005, 1–33.
122. R. Miller, M. Ortiz, R. Phillips, V. Shenoy and E.B. Tadmor, 'Quasicontinuum models of fracture and plasticity', *Eng. Fract. Mech.* **61**(3–4), 427–444 (1998).
123. V.B. Shenoy, R. Miller, E.B. Tadmor, R. Phillips and M. Ortiz, 'Quasicontinuum models of interfacial structure and deformation', *Phys. Rev. Lett.* **80**(4), 742–745 (1998).
124. V.B. Shenoy, R. Miller, E.B. Tadmor, D. Rodney, R. Phillips and M. Ortiz, 'Adaptive finite element approach to atomic-scale mechanics – the quasicontinuum method', *J. Mech. Phys. Solids* **47**(3), 611–642 (1999).
125. E.B. Tadmor, R. Miller, R. Phillips and M. Ortiz, 'Nanoindentation and incipient plasticity', *J. Mater. Res.* **14**(6), 2233–2250 (1999).
126. E.B. Tadmor, M. Ortiz and R. Phillips, 'Quasicontinuum analysis of defects in solids', *Phil. Mag. A* **73**(6), 1529–1563 (1996).
127. D.J. Diestler, Z.-B. Wu and X.C. Zeng, 'An extension of the quasicontinuum treatment of multiscale solid systems to nonzero temperature', *J. Chem. Phys.* **121**(19), 9279–9282 (2004).
128. S. Chen, B. Merriman, M. Kang, R.E. Caflisch, C. Ratsch, L.T. Cheng, M. Gyure, R.P. Fedkiw, C. Anderson and S. Osher, 'A level set method for thin film epitaxial growth', *J. Comput. Phys.* **167**(2), 475–500 (2001).

129. M.F. Gyure, C. Ratsch, B. Merriman, R.E. Caflisch, S. Osher, J.J. Zinck and D.D. Vvedensky, 'Level-set methods for the simulation of epitaxial phenomena', *Phys. Rev. E* **58**(6), R6927–R6930 (1998).
130. D.L. Chopp, 'A level-set method for simulating island coarsening', *J. Comput. Phys.* **162**, 104–122 (2000).
131. M.O. Bloomfield, Y.H. Im and T.S. Cale, 'Microstructure development and evolution', in *IEEE International Conference on Simulation of Semiconductor Processes and Devices*, IEEE, 2003, 19–22.
132. B. Merriman, J.K. Bence and S.J. Osher, 'Motion of multiple junctions: a level set approach', *J. Comput. Phys*. **112**, 334–363 (1994).
133. D. Adalesteinsson and J.A. Sethian, 'The fast construction of extension velocities in level set methods', *J. Comput. Phys.* **148**, 2–22 (1999).
134. D.F. Richards, M.O. Bloomfield, S. Sen and T.S. Cale, 'Extension velocities for level set based surface profile evolution', *J. Vac. Sci. Technol. A* **19**(4, pt. 1–2), 1630–1635 (2001).
135. T.J. Barth and J.A. Sethian, 'Numerical schemes for the Hamilton-Jacobi and level set equations on triangulated domains', *J. Comput. Phys.* **145**(1), 1–40 (1998).
136. M.O. Bloomfield, D.F. Richards and T.S. Cale, 'A computational framework for modelling grain-structure evolution in three dimensions', *Phil. Mag.* **83**(31–34), 3549–3568 (2003).
137. M. Sussman and E. Fatemi, 'An efficient, interface-preserving level set redistancing algorithm and its application to interfacial incompressible fluid flow', *SIAM J. Sci. Comput*. **20**(4), 1165–1191 (1999).
138. M. Sussman, P. Smereka and S. Osher, 'A level set approach for computing solutions to incompressible two-phase flow', *J. Comput. Phys.* **114**(1), 146–159 (1994).
139. PLENTE: Parallel Level-set Environment for Nanoscale Topography Evolution, M.O. Bloomfield and T.S. Cale, PLENTE is a 3D parallel simulation enviroment for multiple material systems.
140. The Message Passing Interface (MPI) standard, *http://www-unix.mcs.anl.gov/mpi/*, accessed February 2006.
141. W.-T. Tseng, C.-H. Lo and S.-C. Lee, 'Electroless deposition of Cu thin films with $CuCl_2$-HNO_3 based chemistry: I. Chemical formulation and reaction mechanisms', *J. Electrochem. Soc*. **148**(5), C327–C332 (2001).
142. T.S. Cale, 'Flux distributions in low pressure deposition and etch models', *J. Vac. Sci. Technol. B* **9**(5), 2551–2553 (1991).
143. M.K. Gobbert, T.P. Merchant, L.J. Borucki and T.S. Cale, 'A multiscale simulator for low pressure chemical vapor deposition', *J. Electrochem. Soc.* **144**(11), 3945–3951 (1997).
144. P. Jain, personal communication to M. Bloomfield 2003.
145. M.O. Bloomfield and T.S. Cale, 'Formation and evolution of grain structures in thin films', *Microelectron. Eng.* **76**, 195–204 (2004).
146. L. Arnaud, G. Tartavel, T. Berger, D. Mariolle, Y. Gobil and I. Touet, 'Microstructure and electromigration in copper damascene lines', *Microelectron Rel.* **40**(1), 77–86 (2000).
147. I.M. Lifshitz and V.V. Slyozov, 'The kinetics of precipitation from supersaturated solid solutions', *J. Phys. Chem. Solids* **19**, 35–50 (1961).
148. D. Turnbull, Theory of grain boundary migration ratio, *Trans. Am. Inst. Min. Engrs*, **191**, 661–665 (1951).

149. P.S. Sahni, D.J. Srolovitz, G.S. Grest, M.P. Anderson and S.A. Safran, 'Kinetics of ordering in two dimensions. II. Quenched Systems', *Phys. Rev. B* **28**(5), 2705–2716 (1983).
150. M.S. Shephard, 'Meshing environment for geometry-based analysis', *Int. J. Numer. Methods Eng*. **47**(1), 169–190 (2000).
151. H.L. de Cougny and M.S. Shephard, 'Parallel refinement and coarsening of tetrahedral meshes', *Int. J. Numer. Methods Eng.* **46**(7), 1101–1125 (1999).
152. J.-F. Remacle, O. Klaas, J.E. Flaherty and M.S. Shephard, 'Parallel algorithm oriented mesh database', *Eng. Comput.* **18**(3), 274–84 (2002).
153. M.S. Shephard, J.E. Flaherty, K.E. Jansen, X. Li, X. Luo, N. Chevaugeon, J.-F. Remacle, M.W. Beall and R.M. O'Bara, 'Adaptive mesh generation for curved domains', *Appl. Numer. Math.* **52**(2–3 Special issue), 251–271 (2005).
154. M.W. Beall, J. Walsh and M.S. Shephard, 'A comparison of techniques for geometry access related to mesh generation', *Eng. Comput.* **20**(3), 210–221 (2004).
155. M.O. Bloomfield, D.F. Richards and T.S. Cale, 'The use of conformal voxels for consistent extractions from multiple level-set fields', *Computational Science – ICCS 2005*, in *Lecture Notes in Computer Science*, eds V.S. Sunderam, G.D.v. Albada, P.M.A. Sloot, and Jack J. Dongarra, Berlin and Heidelberg: Springer – Verlag, 2005, 49–56.
156. M.K. Gobbert, S.G. Webster and T.S. Cale, 'A Galerkin method for the simulation of the transient 2-D/2-D and 3-D/3-D linear Boltzmann equation', *J. Sci. Comput*., **30**(2), 237–273 (2007).
157. M.K. Gobbert and T.S. Cale, 'A kinetic transport and reaction model and simulator for rarefied gas flow in the transition regime', *J. Comput. Phys*. **213**, 591–612 (2006).
158. S. Lei and A. Grama, 'Extended consistent hashing: an efficient framework for object location', in *IEEE International Conference on Distributed Computing Systems*, IEEE, 2004 254–62.
159. E.J. Otoo, D. Rotem and S. Seshadri, 'Efficient algorithms for multi-file caching', in *Database and Expert Systems Applications. 15th International Conference, DEXA 2004*, Boston: Springer-Verlag, 2004, 707–719.
160. Comsol Multiphysics 3.2, Comsol, Inc, *http://www.comsol.com*.
161. M.K. Gobbert, C.A. Ringhofer and T.S. Cale, 'Mesoscopic scale modeling of microloading during low pressure chemical vapor deposition', *J. Electrochem. Soc.* **143**(8), 2624–2631 (1996).
162. T.P. Merchant, M.K. Gobbert, T.S. Cale and L.J. Borucki, 'Multiple scale integrated modeling of deposition processes', *Thin Solid Films* **365**(2), 368–375 (2000).

7

Coupled atomistic/continuum modelling of plasticity in materials

R. E. MILLER, Carleton University, Canada

7.1 Introduction

7.1.1 Mobile dislocations in ductile metals

Understanding the mechanisms of plasticity on the nanoscale is a subject of much recent interest. For example, nanoscale plasticity is relevant in the interpretation of nano-indentation experiments (Gerberich *et al.*, 1996; de la Fuente *et al.*, 2002), and in the understanding of the enhanced hardness that comes from nano-crystalline materials (Valiev, 2002) and nano-layered coatings (Yashar and Sproul, 1999). The great difficulty associated with obtaining detailed experimental observations from such small-scale processes has led to a reliance on atomistic modeling, either with empirical interatomic potentials or with more accurate density functional theory (DFT) calculations. Such modeling ranges from the determination of the details of dislocation core structures (from which certain plastic behaviour can be explained or predicted) to large-scale simulations of plastic deformation. Examples include simulations of single crystal ductile fracture (Abraham *et al.*, 1998b, 2002), nano-indentation into single crystals (Kelchner *et al.*, 1998; Lilleoden *et al.*, 2001, 2003; Zimmerman *et al.*, 2001; de la Fuente *et al.*, 2002; Gerberich *et al.*, 2002; Li *et al.*, 2002; Van Vliet *et al.*, 2003) and nano-layered materials (Saraev and Miller, 2005, 2006), and the yielding of a nano-crystalline sample (Schiotz *et al.*, 1998; Van Swygenhoven, 2002; Yamakov *et al.*, 2002, 2004; Schiotz and Jackobsen, 2003; Van Swygenhoven *et al.*, 2004).

In many of these simulations, there exists a significant challenge to obtaining accurate results. Specifically, the long-ranged stress fields associated with the boundary conditions make it difficult to perform an atomistic simulation that is not in some way affected by artificial constraints imposed by practical limits on the system size. For example, small atomistic models or the use of periodic boundary conditions to mimic an infinite system are not suitable for nano-indentation simulations. First, the finite size of the atomistic system means the model is of an unrealistically thin layer, leading to an artificial

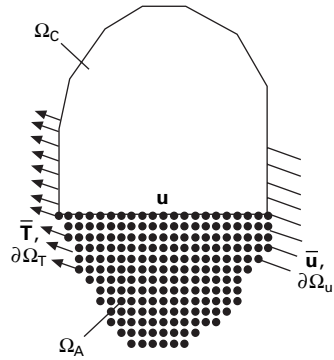
'substrate' effect. Second, periodic boundary conditions lead to a model, not of a single indenter, but of a periodic array of indenters whose interactions can be significant[1].

This challenge has led to so-called 'concurrent' or 'coupled' multiscale approaches, which are the focus of this chapter. In this class of multiscale approaches, a certain critical part of a body is modeled atomistically and coupled directly to a 'coarse-grained' or 'continuum' region elsewhere. The goal of such an approach is to retain fully atomistic detail only where it is needed, while modeling a large enough system to provide realistic boundary conditions and long-ranged elastic interactions. Such a multiscale approach is distinct from so-called 'hierarchical' approaches in which modeling is done at one length scale to extract key data for use in coarse-grained models at larger scales. Reviews of coupled methods may be found in articles by Curtin and Miller (2003) and Liu *et al.* (2004).

In a coupled approach, the underlying atomistic model may be a simple pair potential, an effective medium theory like the embedded atom method (EAM) or glue models, or a full DFT treatment. In addition, the atomistic model may be a zero-temperature equilibrium structure calculation or a molecular dynamics model. In the interest of brevity, we will refer simply to the 'atomistic' model, and make any additional important details clear from the context. It is obviously the case that the accuracy of a concurrent model is limited by the accuracy of the underlying atomistic model, and so it is common to think of the fully atomistic treatment of the same problem (if such a treatment were computationally tractable) as the 'exact' solution that the multiscale approach seeks to represent. The hope of accurately predicting real material behaviour rests on the shoulders of the underlying inter-atomic model, and hence on the subject matter of other chapters. The book by Finnis (2003) is an excellent treatment of the subject of inter-atomic force models.

Most coupled methods take the generic form illustrated in Fig. 7.1, where there is a clear partition of the model into an atomistic region Ω_A, a continuum region Ω_C, and an interface between them, $\partial\Omega_I$. In general, there will be some prescribed displacements, $\mathbf{u} = \overline{\mathbf{u}}$ along part of the boundary $\partial\Omega_u$ and possibly some applied tractions $\mathbf{T} = \overline{\mathbf{T}}$ on $\partial\Omega_T$. The atomistic region is such that in the limit where it becomes large, the exact underlying atomistic model is recovered. The continuum is typically modeled by a linear elastic or hyperelastic finite-element mesh with a strain energy function derived from the underlying atomistics. The key to the accuracy of such approaches, however, is the design of the atomistic/continuum interface, where forces are transmitted to and from the atoms. By necessity, such an interface represents a degree of approximation in the model, and the nature of this approximation

[1]See Saraev and Miller (2006) for discussion of a specific example of such a spurious effect.



7.1 The generic coupled multiscale boundary value problem. Adapted from Shilkrot *et al.* (2002).

will determine the magnitude and extent of so-called 'ghost forces' that arise in the vicinity of the interface. Additional considerations are necessary if the model is dynamic: spurious wave reflections from the interface can render the model useless, and much effort has been spent on developing interfaces with suitably low impedance (Rudd and Broughton, 1998, 2005; Cai *et al.*, 2000; E and Huang, 2001; Park *et al.*, 2005; Qu *et al.*, 2005) or on using a thermostat to mitigate the effects of the reflected waves (Dupuy *et al.*, 2005; Shiari *et al.*, 2005).

However, even within the framework of a coupled model with a well-designed interface, an additional limitation persists when we try to model deformation in ductile solids. In many such problems, there are a large number of mobile dislocations in the model that are either initially present or that are generated during the deformation. Often these dislocations experience high, long-ranged stresses within a material with relatively low lattice resistance to motion. As such, dislocations want to move over long distances, but they cannot be accommodated by the continuum regions of most coupled methods. The apparent solution to this problem is simply to make the atomistic region very large, encompassing the whole domain of potential dislocation mobility. This unfortunately defeats the very purpose of using the coupled model in the first place. It is this detail of coupled models on which we want to focus in this chapter: How do we deal with dislocations in a coupled atomistic/continuum framework?

In passing, we mention that there are several coupled methods for which the atomistic/continuum interface is less clearly defined (see, for example, Wagner and Liu, 2003; Datta *et al.*, 2004; Park *et al.*, 2005b, c) because the continuum and atomistic regions fully overlap and a projection operation is used to prescribe the connection between the two. These methods would face the same challenges discussed here with respect to dislocations, arising from the boundary of the atomistic region instead of from an explicit atomistic/

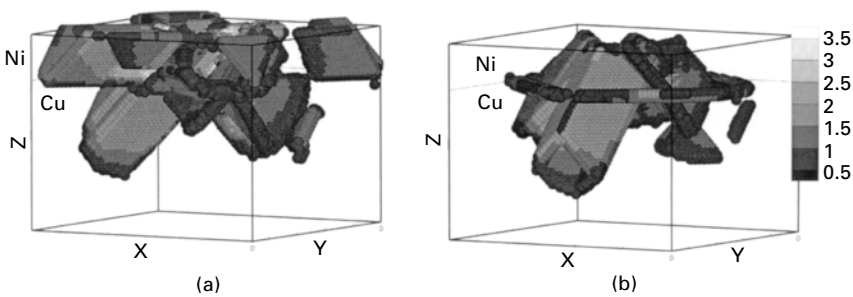
continuum interface. We focus our attention on the two methods that have been developed to solve this specific problem of mobile dislocations: the quasicontinuum (QC) method (Tadmor *et al*., 1996; Shenoy *et al*., 1998 a, b) and the coupled atomistic and discrete dislocation (CADD) method (Shilkrot *et al*., 2002, 2004).

7.1.2 Illustrating the problem

Figure 7.2 shows two MD simulation results for nano-indentation into a copper single crystal coated with a thin single crystal of nickel (Saraev and Miller, 2006). The crystals were initially defect-free, with the only dislocations present in the model being a crisscrossing pattern of misfit dislocations in the Ni–Cu interface. The only difference between the two simulations is the location of the center of the indenter relative to the positions of the misfit dislocations in the interface; however, we can see that there is a substantial qualitative difference in the final dislocation structure which is formed.

The challenge to modeling this indentation example with a multiscale technique is that it is not possible to know in advance where the dislocations will form and move during the simulation. If the atomistic region of the simulation is too small, it will artificially affect how the dislocation structure evolves. If it is too large, on the other hand, it will quickly render the multiscale approach intractable.

Another slightly more quantitative example comes from fracture mechanics. Suppose we have a 2D crack loaded in mode I, which has just emitted a dislocation from its tip. How large does the atomistic region around the crack tip need to be to allow this dislocation to move to its position of equilibrium? We can estimate this based on the stress field around the crack tip, which we know from linear elasticity to go as $\tau \propto K_{\mathrm{I}}/\sqrt{2\pi r}$ where K_{I}



7.2 Molecular dynamics simulations of nano-indentation into a Cu single crystal coated with a thin single crystal of Ni. Atoms are only shown if they are in a region where slip has taken place, as measured by a slip vector analysis (Zimmerman *et al.,* 2001) and shown by the contours. Adapted from Saraev and Miller (2006).

is the stress intensity factor due to the loading and r is the distance from the crack tip. This is roughly the stress driving the dislocation, which will move away from the crack tip until this stress has decayed to below the Peierls stress of the crystal. A typical bulk, polycrystalline aluminum alloy has a fracture toughness in the range of $K_{IC} = 20$ MPa $\sqrt{m}$, but perhaps more appropriate in this case is a value of about $K_{IC} = 0.5$ MPa $\sqrt{m}$ reported by Miller *et al.* (1998) for the advance of an atomistically modeled crack tip in a single crystal of Ni. Assume the crack is loaded well below this critical value, such that $K_I = 0.1\ K_{IC} = 0.05$ MPa $\sqrt{m}$. Now, take a typical value of the Peierls stress to be $\tau_{Peierls} = 20$ MPa as reported by Olmsted *et al.* (2001) and we find that the equilibrium distance for the dislocation is at $r = 1\ \mu$m. Requiring an atomistic region with a 1 μm radius, which even in 2D contains on the order of 10^7 atoms, is clearly not computationally efficient.

In this chapter, we discuss two specific solutions that have been developed to address the issue of mobile dislocations in coupled methods. The first is to use a model that can continually adapt the size and shape of the atomistic region as necessary to accommodate the evolving defect structure. In this way, a fully atomistic treatment can follow dislocation cores as they emerge into the model. This approach was developed within the framework of the QC method, and so we discuss it here in that context. The second method to treat dislocations is to keep the atomistic region fixed, but use an algorithm that automatically detects dislocations as they approach the atomistic/continuum interface and converts them to linear elastic defects. These are then passed into the continuum region and treated using the methods of discrete dislocation (DD) models (see Chapter 2). As such, this approach assumes that once the dislocations have left the initial atomistic region, their important effect is their elastic interactions with the rest of the model, and the details of their atomistic core can then be ignored. The CADD method has been developed to effect this coupling between an atomistic region and a DD region, and we review the method here with particular emphasis on the dislocation detection and passing components.

Throughout the chapter, we will focus on a very simple 2D indentation example to illustrate the problem and solution associated with each method. We leave the discussion of the many applications of these methods to other recent reviews (Miller and Tadmor, 2002; Curtin and Miller, 2003; Tadmor and Miller, 2006) and the original references (Shenoy *et al.*, 1998a; Shilkrot *et al.*, 2002). An open-source implementation of the QC method is available for download from www.qcmethod.com.

7.2 Automatic adaption: the QC method

As we have said, the focus in this chapter will be on ways to deal with dislocations in multiscale approaches. The adaptive meshing strategies

implemented in the QC approach and discussed in Section 7.2.3 can, in principle, be applied to other multiscale approaches such as those of Rudd and Broughton (2000, 2005) and Abraham *et al*. (1998a). However, because this adaptive methodology is currently only implemented in QC, we will focus on the QC formulation specifically. Before turning to the question of adaption, it is helpful to briefly review the details of the QC method itself, and for that purpose we follow closely the review presented by Tadmor and Miller (2006).

7.2.1 The 'exact' atomistic problem

For the time being, we focus on *lattice statics* solutions. We are looking for equilibrium atomic configurations for a given model geometry and externally imposed forces or displacements. For the interested reader, there has been much recent work in developing a finite temperature, dynamic QC implementation (Shenoy *et al*., 1999; Dupuy *et al*., 2005).

We assume that there is some reference configuration of N atomic nuclei, initially forming a lattice. Thus, the reference position of the ith atom in the model $\mathbf{X}_i$ is found from an integer combination of lattice vectors and an origin atom position, $\mathbf{X}_0$

$$\mathbf{X}_i = \mathbf{X}_0 + l_i\mathbf{A}_1 + m_i\mathbf{A}_2 + n_i\mathbf{A}_3 \qquad [7.1]$$

where (l_i, m_i, n_i) are integers and $\mathbf{A}_j$ is the jth Bravais lattice vector. In principle, the QC method does not depend on having a crystalline reference configuration, but rather depends on having an efficient and unambiguous method for determining the locations of all the atoms in the reference body. The Bravais lattice is perhaps the simplest such prescription.

The deformed position of the ith atom $\mathbf{x}_i$ is then found from a unique displacement vector $\mathbf{u}_i$ for each atom

$$\mathbf{x}_i = \mathbf{X}_i + \mathbf{u}_i \qquad [7.2]$$

The displacements $\mathbf{u}_i$, while only having physical meaning on the atomic sites, can be treated as a continuous field $\mathbf{u}(\mathbf{X})$ throughout the body with the property that $\mathbf{u}(\mathbf{X}_i) \equiv \mathbf{u}_i$. This approach, while not the conventional one in atomistic models, is useful in effecting the connection to continuum mechanics. Note that for brevity we will sometimes refer to the field $\mathbf{u}(\mathbf{X})$ to represent the set of all atomic displacements $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_N\}$ where N is the number of atoms in the body.

We assume that the underlying atomistic model is such that we can write the total potential energy of the system of atoms as a sum over well-defined site energies E_i associated with each atom i. Specifically,

$$E^{\text{tot}} = \sum_{i=1}^{N} E_i(\mathbf{u}) \qquad [7.3]$$

where E_i is the site energy of atom i, which depends on the displacements $\mathbf{u}$ through the relative positions of all the atoms in the deformed configuration. In addition to the potential energy of the atoms, there may be energy due to external loads applied to atoms. Thus, the total potential energy of the system (atoms plus external loads) can be written as

$$\Phi(\mathbf{u}) = E^{\text{tot}}(\mathbf{u}) - \sum_{i=1}^{N} \mathbf{f}_i \mathbf{u}_i \qquad [7.4]$$

where $\mathbf{f}_i\mathbf{u}_i$ is the potential energy of the applied load $\mathbf{f}_i$ on atom i. Thus the 'exact' fully atomistic solution is the set of displacements $\mathbf{u}$ that minimize the function Φ subject to the external loads $\mathbf{f}_i$ and constraints $\mathbf{u}_i = \bar{\mathbf{u}}_i$ on a subset of the atoms in the body.

7.2.2 The continuum region

The essential idea of the QC method is to re-write the energy of equation [7.3] such that it remains acceptably accurate but computationally more efficient. First, we can think of the sum as being re-written as a sum over atoms in the 'atomistic' region and another over the atoms in the 'continuum' region as illustrated in Fig. 7.1. Practically speaking, such a decomposition will initially be made by fiat during the setup of the model, but will then be able to evolve based on criteria (to be discussed later) that quantitatively determine whether or not a certain region needs to be treated atomistically. Thus, we have the decomposed sum

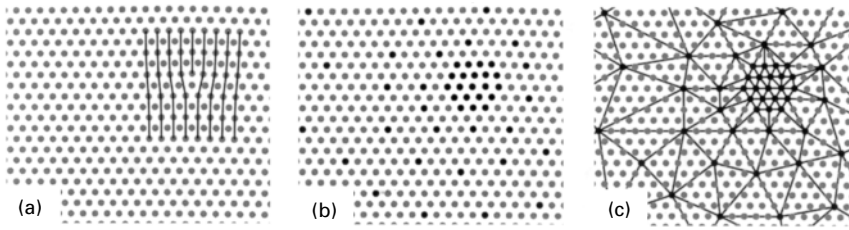
$$E^{\text{tot}} = \sum_{i \in \Omega_{\text{A}}} E_i(\mathbf{u}) + \sum_{i \in \Omega_{\text{C}}} E_i(\mathbf{u}) \qquad [7.5]$$

The first sum must be performed exactly as in the fully atomistic model. The second sum, which is generally over a very large number of atoms, is made computationally tractable by use of the finite-element method, as we will outline presently.

A small subset of the atoms in the problem are chosen as so-called 'representative atoms' or 'repatoms'. In the atomistic region, all atoms are chosen, while in the continuum region a relatively small fraction are chosen, as shown in Fig. 7.3. A finite-element mesh is drawn between the repatoms, and the constraint is imposed that only the displacements of the repatoms are degrees of freedom in the problem. For any atom lying between the repatoms, the displacements are constrained to follow an interpolated displacement field found from the finite-element shape functions.

This first approximation of the QC, then, is to replace the energy E^{tot} by $E^{\text{tot,h}}$:

$$E^{\text{tot,h}} = \sum_{i \in \Omega_{\text{A}}} E_i(\mathbf{u}^{\text{h}}) + \sum_{i \in \Omega_{\text{C}}} E_i(\mathbf{u}^{\text{h}}) \qquad [7.6]$$



7.3 (a) A dislocation in an otherwise perfect crystal experiences severe deformation variations only around the dislocation core. (b) Schematic representative atom selection, with the region around the core treated atomistically. (c) Mesh for the QC method.

In this equation the atomic displacements are now found through the interpolation functions and take the form

$$\mathbf{u}^{\mathrm{h}} = \sum_{i=1}^{N_{\mathrm{rep}}} S_i \mathbf{u}_i \qquad [7.7]$$

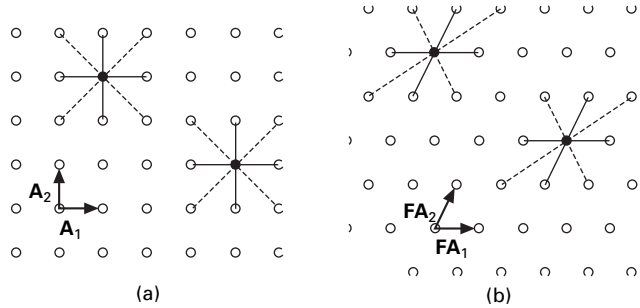
where S_i is the interpolation (shape) function associated with repatom i, and N_{rep} is the number of repatoms, $N_{\mathrm{rep}} << N$. Note that the formal summation over the shape functions in equation [7.7] is in practice much simpler due to the compact support of the finite-element shape functions. Specifically, shape functions are identically zero in every element not immediately adjacent to a specific repatom. Also note that in the atomistic region, this interpolation is actually not used since there are no atoms *inside* elements in that region.

Introducing this kinematic constraint on most of the atoms in the body will achieve the goal of reducing the number of degrees of freedom in the problem, but notice that for the purpose of energy minimization, equation [7.6] still requires a sum over all the atoms in Ω_{C} to compute the total energy. It is necessary to reduce the computational effort required to find the total energy.

A key feature of a simple Bravais lattice is that every atomic site sees the same neighbor environment as any other site. Figure 7.4 shows a Bravais lattice undergoing a uniform deformation. It is useful to characterize the deformation by the *deformation gradient*, $\mathbf{F}$, defined by

$$\mathbf{F}(\mathbf{X}) \equiv \frac{\partial \mathbf{x}}{\partial \mathbf{X}} = \mathbf{I} + \frac{\partial \mathbf{u}}{\partial \mathbf{X}} \qquad [7.8]$$

where $\mathbf{I}$ is the identity tensor. For a uniform deformation, $\mathbf{F}$ is constant throughout the body. Clearly, even if the resulting deformation is quite large, the equivalency of all the atomic neighbor environments is preserved as shown in Fig. 7.4(b). To compute the energy of all the atoms in some region of this deformed crystal (assuming it to be infinite in extent), we need only



7.4 A simple Bravais lattice in (a) deforms uniformly under deformation gradient in (b). Every atom in either the reference or the deformed lattice sees an equivalent neighbor environment, as shown for two atoms in each lattice.

compute the energy of one atom, and multiply by the number of atoms in the region.

In QC, we make use of this fact and employ finite-element shape functions that ensure a uniform deformation gradient within each element. Thus, we approximate the energetic contribution of all the atoms in the element as the number of atoms in the element multiplied by the energy of one such atom, embedded in an infinite crystal experiencing the deformation gradient of the element. The energy is therefore approximated by

$$E^{\mathrm{tot,h}} \approx E^{\mathrm{QC}} = \sum_{i \in \Omega_{\mathrm{A}}} E_i(\mathbf{u}^{\mathrm{h}}) + \sum_{\mu=1}^{M_{\mathrm{C}}} n_\mu \hat{E}(\mathbf{F}_\mu) \qquad [7.9]$$

where M_{C} is the number of elements filling the space between the repatoms of the continuum region. The weighting function n_μ is the number of atoms within element μ, and $\hat{E}(\mathbf{F})$ is the energy of a single atom in an infinite crystal deformed by the deformation gradient $\mathbf{F}$. The energy $\hat{E}(\mathbf{F})$ can be efficiently computed by building a small crystallite of atoms, deformed according to $\mathbf{F}$ and just large enough to completely fill the cutoff sphere of an atom at its center.

Note that there is an additional level of approximation here, because atoms near the boundaries of the elements will generally see a non-uniform environment if some of their neighbors reside in other elements. In the multiscale literature, this is sometimes referred to as the 'local' approximation, reflecting the fact that it assumes that we can compute the energy of an atom armed only with the knowledge of the local (pointwise) value of the deformation gradient. In fact, the nature of atomic models (at least those that are more sophisticated than simple near-neighbor pair potential descriptions), is that they are inherently *non-local*. That is to say, the energy of an atom depends on neighbors a finite distance away, typically up to some cutoff distance on the order of the third or fourth neighbors for metallic systems. Clearly, if the

deformation of a region is uniform up to this distance from the atom in question, then this atom's energy will be accurately obtained from the local approximation. However, if there are significant deviations from the uniform deformation gradient within that range, the local approximation will breakdown. The calculation of the total energy of a uniformly deformed region of a crystal in this way is related to the Cauchy–Born rule (Ericksen, 1984) which is essentially the assumption that a uniform deformation on the continuum length scale prescribes the same uniform deformation of the underlying atomic lattice.

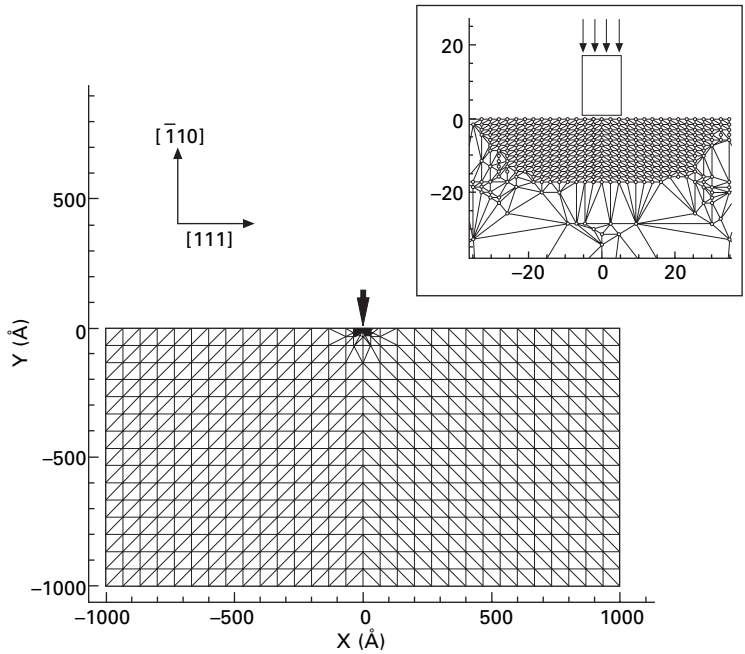
Equation [7.9] provides the essential framework for the static (zero T) QC method. Using it to approximate the total potential of the system in equation [7.4] leads to a function of only the positions of the repatoms, and so differentiation with respect to the repatom positions permits minimization of Φ, and ultimately one can find the displacements for the equilibrium solution.

7.2.3 Automatic adaption

The key aspect of QC that we wish to focus on here is mesh adaption. As a specific illustration, we have run the example indentation simulation provided with the QC code V1.2 at www.qcmethod.com ('Punch-Example'). A block of single crystal fcc aluminum, with the dimensions and orientation shown in Fig. 7.5, is modeled using the embedded atom potentials of Ercolessi and Adams (1994). A flat square indenter of width 10 Å is modeled by simply prescribing the displacement of five atoms along the top surface of the crystal. This displacement is incremented by 0.2 Å between relaxations of the model to simulate quasi-static loading. The inset in Fig. 7.5 shows the initial atomistic region below the indenter.

As we shall see, the result of this simulation will be the nucleation of a pair of Shockley partials from beneath each corner of the indenter. The nucleated defects want to move some distance from the indenter, perhaps into the region that is currently modeled using the continuum approximation. However, the continuum region cannot support dislocations, because the deformation around a dislocation core varies too rapidly on the atomic scale. At the very least, we must adapt the mesh in the continuum region down to the atomic scale, in order to accurately capture the displacement field around a dislocation core. Even this is not enough, since the local approximation inherent to the continuum region is no longer valid. The goal, then, is to implement an algorithm within QC that can automatically update both the degree of mesh refinement and the status of a repatom as either 'atomistic' (and hence fully non-local) or 'continuum' (and therefore reasonably approximated by the local assumption).

We first turn our attention to the question of mesh refinement. To this end, the QC makes use of the finite-element literature, where considerable attention



7.5 The initial mesh for the illustrative indentation example using QC.

has been given to adaptive meshing techniques for many years. A scalar measure is defined to quantify the error introduced into the solution by the current density of repatoms. Elements in which this error estimator is higher than a prescribed tolerance are targeted for adaption, while at the same time the error estimator can be used to remove unnecessary nodes from the model. The error estimator of Zienkiewicz and Zhu (1987), originally posed in terms of errors in the stresses, is re-cast for the QC in terms of the deformation gradient. Specifically, we define the error estimator to be

$$\varepsilon_e = \left[\frac{1}{\Omega_e} \int_{\Omega_e} (\overline{\mathbf{F}} - \mathbf{F}_e) : (\overline{\mathbf{F}} - \mathbf{F}_e) \, \mathrm{d}\Omega \right]^{1/2} \qquad [7.10]$$

where Ω_e is the volume of element e, $\mathbf{F}_e$ is the QC solution for the deformation gradient in element e, and $\overline{\mathbf{F}}$ is the L_2-projection of the QC solution for $\mathbf{F}$, given by

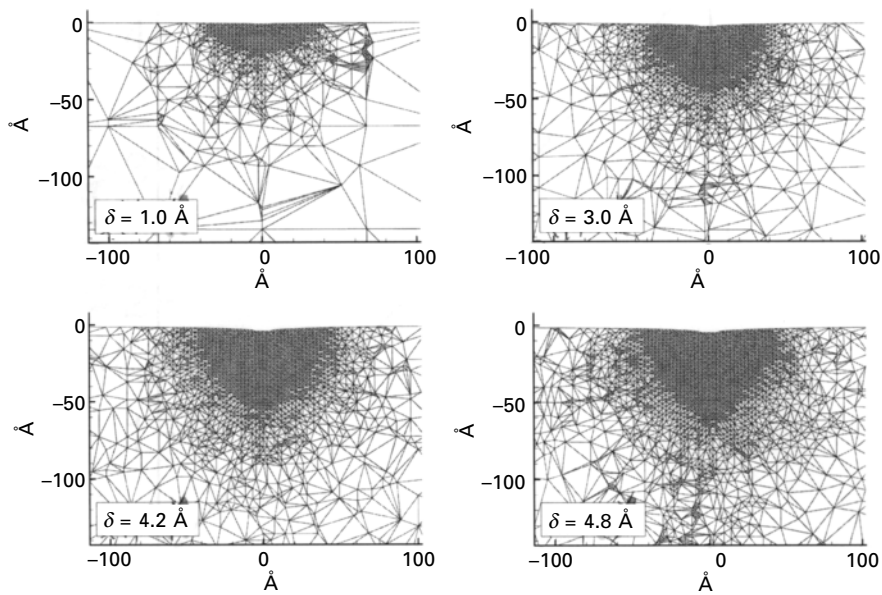
$$\overline{\mathbf{F}} = \sum_{i=1}^{N_{\mathrm{rep}}} S_i \mathbf{F}_i^{\mathrm{avg}} \qquad [7.11]$$

Here, S_i is the shape function for node i, and $\mathbf{F}_i^{\mathrm{avg}}$ is the matrix of nodal values of the projected deformation gradient $\overline{\mathbf{F}}$. Because the deformation gradients are constant within the linear elements used in the QC, the nodal

values $\mathbf{F}_i^{\text{avg}}$ are simply computed by averaging the deformation gradients found in each element touching a given repatom. Clearly, if neighboring elements all have the same $\mathbf{F}_e$, then $\mathbf{F}_{\text{avg}} \approx \mathbf{F}_e$ and ε_e is small. On the other hand, if the deformation gradient varies rapidly between neighboring elements, adaption is triggered by this criterion as required.

Refinement then proceeds by adding new repatoms within the targeted elements, depending on the type of element considered. For example, in a 2D, 3-node triangular element, new repatoms are added at the nearest atomic sites to the midpoints on the three element sides. Notice that since repatoms must fall on actual atomic sites in the reference lattice, there is a natural lower limit to element size; if the nearest atomic sites to the mid-sides of the elements are the atoms at the element corners, the region is fully refined and no new repatoms can be added. In Fig. 7.6, we see a number of adaption steps due to the increased variation in the deformation gradient as the indentation proceeds.

The question of whether a repatom should be treated locally or non-locally is probed using the so-called *non-locality criterion* in QC. As we have stressed above, simply having a large deformation in a region does not in itself require a non-local repatom, as the local formulation will exactly describe the energy of any *uniform* deformation, regardless of the severity. The key feature that should trigger a non-local treatment of a repatom is a significant *variation* in the deformation gradient on the atomic scale in the



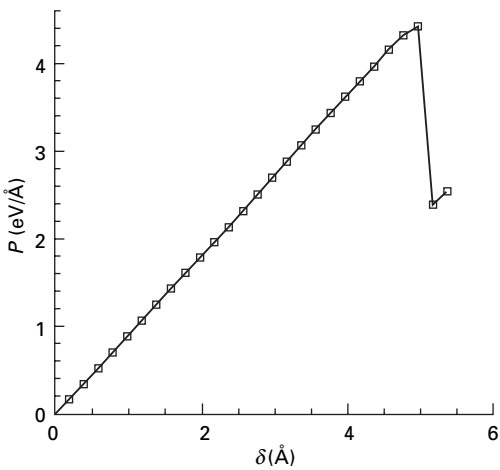
7.6 Adapted QC mesh for various indentation depths, δ.

repatom's proximity. Thus, the non-locality criterion is implemented as follows. A cutoff, r_{nl}, is empirically chosen to be between two and three times the cutoff radius of the inter-atomic potentials. The deformation gradients in every element within this cutoff of a given representative atom are compared, by looking at the differences between their eigenvalues. The criterion is then:

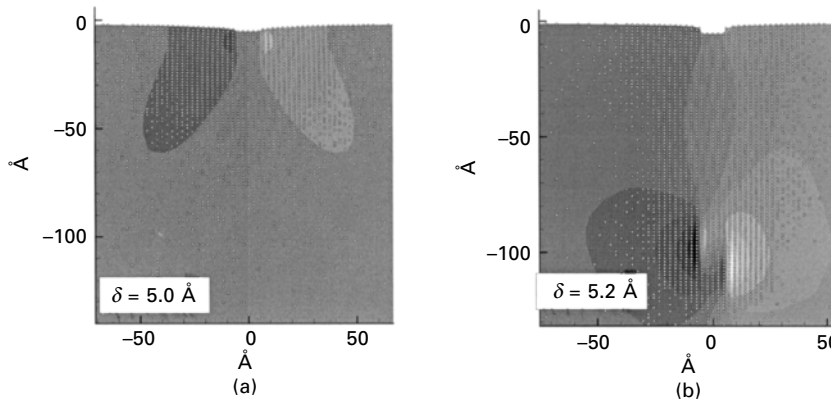
$$\max_{a,b;k} \left| \lambda_k^a - \lambda_k^b \right| < \varepsilon \qquad [7.12]$$

where λ_k^a is the kth eigenvalue of the right stretch tensor $\mathbf{U}_a = \sqrt{\mathbf{F}_a^{\mathrm{T}} \mathbf{F}_a}$ in element a, $k = 1...3$, and the indices a and b run over all elements within r_{nl} of a given repatom. The repatom will be kept local if this inequality is satisfied, and made non-local otherwise. In practice, the tolerance ε is determined empirically. A value of 0.1 has been used in a number of tests and found to give good results. The effect of this criterion is clusters of non-local atoms in regions of rapidly varying deformation, forming the 'atomistic' regions, Ω_{A}, in the QC simulation.

Figure 7.7 shows the load–displacement curve for the indentation example. At an indenter depth of 5.0 Å, we see the load drop corresponding to the nucleation of defects, and in Fig. 7.8 we see the atomistic region just before and just after this load drop. The contours behind the atoms show the out-of-plane displacements, revealing the width of the stacking fault between the two partials. It is clear how the atomistic region has grown to follow the dislocations as they move into the continuum, a process that will continue as the dislocations travel deeper into the crystal.



7.7 Load–displacement curve for the QC indentation example.



7.8 QC indentation result (a) just before and (b) just after the nucleation of dislocations from the indenter corners.

7.3 Kinematically identifying dislocations – the CADD method

7.3.1 The CADD formulation

An alternative approach to the problem of mobile dislocations is to recognize that once a dislocation has nucleated and moved away from the critical region of interest, the details of the dislocation core structure are of secondary importance. The main effect that needs to be captured with respect to these dislocations is their long-ranged elastic interactions with the atomistic region and the externally applied boundary conditions.

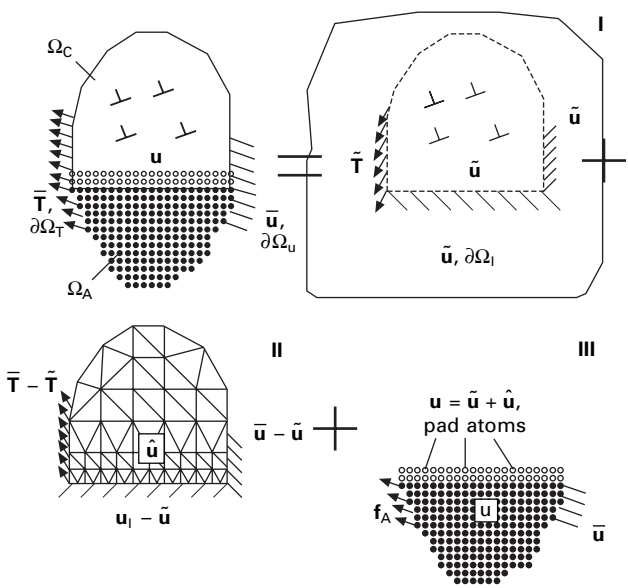
Continuum models of dislocation mechanics which treat dislocations as line defects in an elastic material exist in many forms, typified by the work of Van der Giessen and Needleman (1995), Zbib *et al*. (1998), Kubin and Canova (1992), Fivel and Canova (1999), Cleveringa *et al*. (1997), Weygand *et al*. (2002) and others and here we refer to these models collectively as DD methods. In these methods, the long-ranged interactions between the dislocations are modeled to the accuracy of linear elasticity, which tends to be suitable as long as the defects are at least a few nanometers apart. Smaller-scale phenomena and short-range dislocation interactions are described by a set of constitutive rules. The rules dictate, for instance, the conditions required for dislocation nucleation, the minimum dislocation separation for annihilation, the strength of dislocation junctions, the dislocation mobility, and dislocation/obstacle interactions. Suitable values for the parameters in these rules can be obtained, in principle, through detailed atomistic simulations.

These DD methods are continuum-based models that can be used to treat general boundary conditions and complex model geometries. As such, they can be used to model the continuum side of a coupled multiscale approach

with very little modification to the multiscale methodology. However, the difficulty comes when dislocations want to move between the two regions, either from the atomistic region to the continuum or vice versa. With a suitably robust algorithm to effect this passing of dislocations between the regions, the problem of mobile dislocations can be treated with an extremely small atomistic region and no need for automatic mesh adaption.

The CADD model was developed with exactly this goal in mind, and we discuss it here. Our emphasis is on the feature of CADD which makes it suitable for treating mobile dislocations: the algorithm for detecting and passing defects as they move across the atomistic/continuum interface. First, we present a brief overview of the CADD approach, the details of which can be found in Shilkrot *et al*. (2002). We focus our attention on the zero-temperature, static implementation of CADD, with quasi-static loading achieved by incrementally increasing boundary loads or displacements on previous equilibrium solutions. A dynamic, finite temperature CADD model has been successfully implemented (Shiari *et al*., 2005) and used to study nano-indentation.

In Fig. 7.9, we see the essential details of the CADD model. In contrast to Fig. 7.1, the continuum region Ω_C now contains N_d discrete dislocations. The solution we seek consists of the atom positions $\mathbf{x}_A$, dislocation positions $\mathbf{d}^i$ and continuum displacements $\mathbf{u}$, stresses σ, and strains ε.



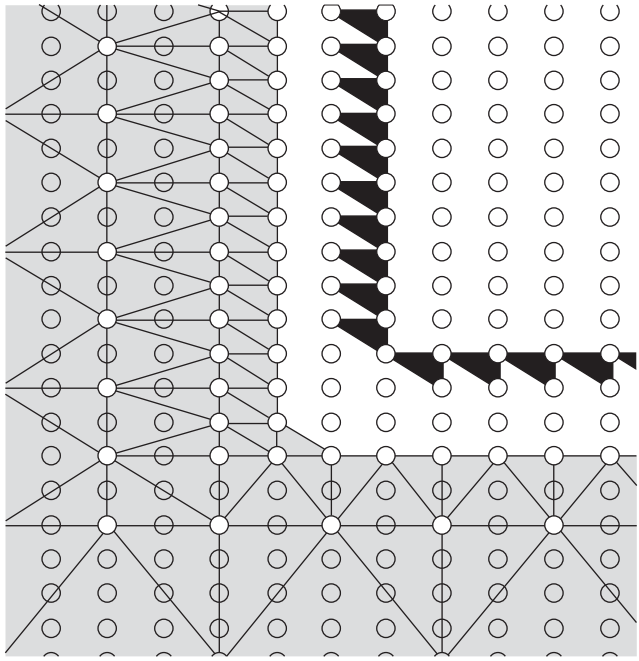
7.9 The augmented boundary value problem of CADD (which includes discrete dislocations in the continuum) is solved by superposition of two linear solutions (I and II) with the non-linear atomistic solution (III). Adapted from Shilkrot *et al.* (2002).

The continuum region is linear elastic, and requires the anisotropic elastic moduli, computed *a priori* from the atomistic model for an appropriately oriented crystal of the material. As we shall see below, the algorithm for passing dislocations to the continuum also requires a list of all dislocation Burgers vectors permitted to exist in the continuum, which is again determined from knowledge of the crystal structure.

The continuum problem is divided into two complementary problems (Fig. 7.9) and solved using the method of van der Giessen and Needleman (1995). Problem I consists of dislocations in an infinite elastic continuum and is solved by superposition of the analytical elastic fields due to the individual dislocations at positions $\mathbf{d}^i$, yielding a total field denoted as the ~ field. Problem I generates tractions $\tilde{\mathbf{T}}$ along $\partial\Omega_T$ and displacements $\tilde{\mathbf{u}}$ and $\tilde{\mathbf{u}}_I$ along $\partial\Omega_u$ and $\partial\Omega_I$ that differ from the values of $\overline{\mathbf{T}}$ and $\overline{\mathbf{u}}$ prescribed by the boundary conditions and the $\mathbf{u}_I$ imposed by the atomistic region. Problem II consists of a linear elastic continuum with no dislocations but subject to 'corrective' tractions $\hat{\mathbf{T}} = \overline{\mathbf{T}} - \tilde{\mathbf{T}}$ on $\partial\Omega_T$ and displacements $\hat{\mathbf{u}} = \overline{\mathbf{u}} - \tilde{\mathbf{u}}$ on $\partial\Omega_u$ and $\hat{\mathbf{u}} = \mathbf{u}_I - \tilde{\mathbf{u}}_I$ on $\partial\Omega_I$. All discontinuities and singularities of the dislocations are contained in the ~ fields of Problem I, so the fields of Problem II, denoted as ^ fields, are smooth and obtainable numerically by the finite-element method. The total fields in the continuum are the superposition of the fields from Problems I and II: $\mathbf{u} = \tilde{\mathbf{u}} + \hat{\mathbf{u}}$ $\boldsymbol{\sigma} = \tilde{\boldsymbol{\sigma}} + \hat{\boldsymbol{\sigma}}$ and $\boldsymbol{\epsilon} = \tilde{\boldsymbol{\epsilon}} + \hat{\boldsymbol{\epsilon}}$. For a given displacement along the atomistic/continuum interface $\partial\Omega_I$ (which is known at any instant from the atoms defining this interface) and a set of dislocation positions, this solution is directly obtained from the inverted stiffness matrix of the finite-element model and a simple summation since all parts of the continuum are linear. From this solution, the stress anywhere in the continuum can be found and used to obtain the Peach–Koehler force on each discrete dislocation. The displacement of any atom in the continuum region can be also be obtained, if needed, using the reference crystal structure, the field $\tilde{\mathbf{u}}$, and the interpolated field $\hat{\mathbf{u}}$.

Forces on the atoms in Problem III are computed based on a physical interpretation of atomic forces in such a way that spurious forces at the atomistic/continuum interface are eliminated. First, the positions of the *pad atoms* (shown as unfilled circles in Fig. 7.9) are found from interpolation of the continuum displacement fields. This pad is at least twice the cutoff radius of the interatomic potentials to ensure that atoms on $\partial\Omega_I$ do not 'see' the free surface created by the cut. Forces on the atoms in Ω_A and on $\partial\Omega_I$ are then the forces as obtained from a purely atomistic description of the material including the pad atoms. A close-up of the interface is shown in Fig. 7.10.

With the forces on the atoms and dislocations obtained as described above, the system can then be iteratively driven to equilibrium by an algorithm that finds the solution for which these forces are zero.



7.10 The atomistic/continuum interface of CADD. Light grey elements are the continuum, white region is the atomistic region. The dark elements are the detection band and the light grey circles in the continuum region are the pad atoms.

7.3.2 Detecting dislocations

The CADD model is deliberately designed so that the atomistic region can be very small, and it is therefore expected that dislocations generated within the atomistic region will move into the continuum region. Likewise, it is possible that dislocations may move in the other direction, after being generated from sources in the continuum or due to a reversal of the driving stresses as the deformation evolves. As a dislocation in the atomistic region approaches the interface, however, spurious forces will be generated: the elastic continuum is not able to properly accommodate the shear deformation associated with the Burgers vector of the dislocation. It resists this shear by a fully elastic response, whereas in the true atomistic case the non-linear response leads to softening under this shear, which permits continued glide of the dislocation.

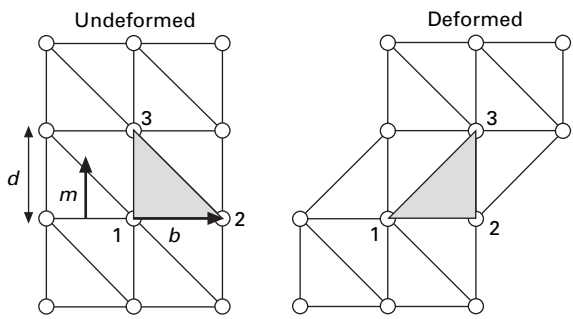
In order for the dislocation to pass through the interface we must be able to detect its approach and then artificially move across the interface, converting it to a discrete dislocation. This requires not only detection of the dislocation's presence, but an unambiguous identification of its Burgers vector and slip plane, so that once in the continuum its further motion will be properly characterized. At the same time, the atomistic core structure must be eliminated

from the atomistic region. In this section, we focus on this feature of the CADD model: how can we suitably detect and pass dislocations across the atomistic/continuum interface?

Although the continuum region is linear elastic, a dislocation in the atomistic region may fortunately move quite close to the atomistic/continuum interface before there are significant spurious effects from the interface. This can be inferred from the size of the region around a dislocation core in which there are significant deviations from linear response. As was shown by Miller and Phillips (1996), this core region is only a few Burgers vectors wide[2]. The compactness of the core region makes it possible leave the detection of dislocations until they have come to within a few atomic layers of the interface[3].

To detect dislocations approaching the interface, we define a thin 'detection band' of finite elements in the atomistic region some short distance from the interface, as shown by the darkly shaded elements in Fig. 7.10. As the atoms move during the deformation of this region, the elements in the detection band deform, and standard finite-element shape functions permit the efficient calculation of a deformation gradient, **F**, for each detection element.

Let us first consider the effect of a single dislocation on one of these elements in an initially perfect crystal. This is illustrated by the shaded element in Fig. 7.11, where a dislocation has passed through the center of the crystal and sheared the top with respect to the bottom. In the limit where the dislocation has traveled far from the element of interest, the effect is to move



7.11 The effect of slip on elements drawn between atomic planes.

[2]For dislocations which dissociate into partials, there is of course a larger region of non-linearity due to the stacking fault between the partials. In that case, we mean that the non-linearity of the dislocation dies off quickly as one moves away from either partial, and not into the stacking fault.

[3]Recent work by Dewald and Curtin (2005) has shown that there are other important and sometimes significant errors due to the discreteness of the lattice and ambiguities associated with the precise location of the dislocation core. These errors need to be dealt with carefully to preserve the 'seamlessness' of the interface.

atom 3 exactly one Burgers vector along the slip direction with respect to atoms 1 and 2, creating a uniform pure shear deformation in the element. For a generally oriented slip plane and Burgers vector in that plane, this leads to a Lagrangian strain matrix of

$$\mathbf{E}^{\mathrm{p}} = \frac{(\mathbf{b} \otimes \mathbf{m})_{\mathrm{sym}}}{d} + \frac{(\mathbf{m} \otimes \mathbf{b})(\mathbf{b} \otimes \mathbf{m})}{2d^2} \qquad [7.13]$$

where $\mathbf{b}$ is the Burgers vector, $\mathbf{m}$ is the slip plane normal, sym implies the symmetric part of the matrix and $\otimes$ implies the tensor product. For the example shown in Fig. 7.11, we have $\mathbf{b}_1 = (d, 0)$, $\mathbf{m}_1 = (0, 1)$ and the strain matrix becomes

$$\mathbf{E}_1^{\mathrm{p}} = \begin{bmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} \qquad [7.14]$$

By contrast, if the slip had occurred along the vertical slip plane passing through the shaded element we would have $\mathbf{b}_2 = (0, d)$ and $\mathbf{m}_2 = (-1, 0)$, leading to the strain matrix

$$\mathbf{E}_2^{\mathrm{p}} = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & 0 \end{bmatrix} \qquad [7.15]$$

This strain matrix is a unique signature of the particular slip system defined by the combination of $\mathbf{b}$ and $\mathbf{m}$. For problems of interest here, we know the crystal structure and orientation, as well as the set of expected dislocations in that particular crystal. For example, in an fcc crystal there are 24 possible slip planes (four slip planes, three Shockley partials on each, each of which can be positive or negative). If we are given a particular strain matrix in an element of the detection band, it is a relatively quick and straightforward calculation to compare it to this small library of 'special' strain matrices associated with dislocation motion.

As a dislocation passes through an element in the detection band, it will be deformed due to three contributions: the displacements associated with any dislocations that have previously passed through that element, the displacements associated with the new dislocation, and an additional elastic deformation associated with the loading and geometry of a particular problem. In general, we can safely assume that this last elastic contribution will be small, with components on the order of 1% or less (compared to components of $\mathbf{E}^{\mathrm{p}}$ like those shown above, which are on the order of 50 %), and these can be ignored for our purposes. Recall that the detection band is immediately adjacent to the continuum region that is assumed to be linearly elastic, and

so if this assumption of a small elastic part is violated then in fact the entire model is called into question. Thus we assume that there are only two important contributions to the displacements of the detection band: that due to previously passed dislocations and that due to the new defects.

Any displacement due to previous dislocations is already known and stored in the $\tilde{\mathbf{u}}$ field defined in the decomposed problem I of Fig. 7.9. Subtracting the $\tilde{\mathbf{u}}$ displacements due to the continuum dislocations, we have the relevant intermediate configuration $\hat{\mathbf{x}} = \mathbf{X} + \hat{\mathbf{u}}$ where $\hat{\mathbf{u}} = \mathbf{u} - \tilde{\mathbf{u}}$. The deformation gradient $\mathbf{F}$ used to identify new slip is then defined as

$$\mathbf{F} = \frac{\partial \hat{\mathbf{x}}}{\partial \mathbf{X}} = \mathbf{I} + \frac{\partial \hat{\mathbf{u}}}{\partial \mathbf{X}} \qquad [7.16]$$

where within each detection band element $\hat{\mathbf{u}}$ is interpolated using equation 7.7. From $\boldsymbol{F}$ we can obtain the strain as

$$\mathbf{E} = \frac{1}{2}\left[\mathbf{F}^{\mathrm{T}}\mathbf{F} - \mathbf{I}\right] \qquad [7.17]$$

We can now detect dislocations as follows. We first build the library of ideal strain matrices, $\mathbf{E}_i^{\mathrm{p}}$ for i representing all the possible slip systems in our crystal. For convenience, an additional 'dislocation' is defined with Burgers vector $\mathbf{b} = \mathbf{0}$, corresponding to no slip in the element and $\mathbf{E}_i^{\mathrm{p}} = \mathbf{0}$. During a simulation, with each update of the atomic positions, we compute the actual strain in each detection band element, and identify the slip in each element as being due to the dislocation which minimizes the L_2 norm of the difference between the actual elemental strain and the dislocation plastic strain:

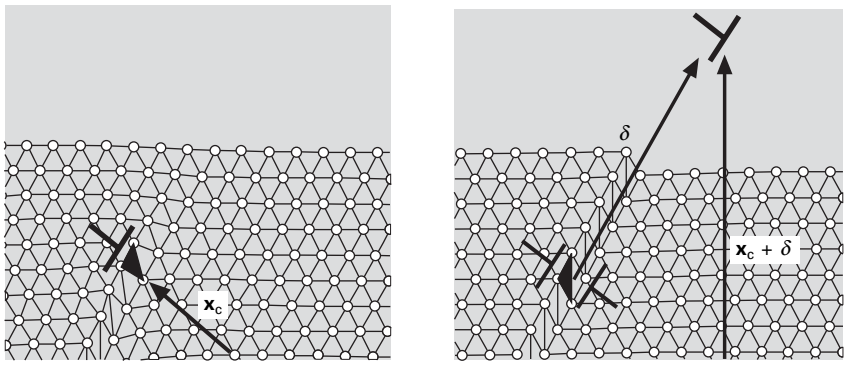
$$L_{2,i} = \sqrt{(\mathbf{E} - \mathbf{E}_i^{\mathrm{p}}) : (\mathbf{E} - \mathbf{E}_i^{\mathrm{p}})} \qquad [7.18]$$

If the dislocation minimizing this norm has $\mathbf{b} = \mathbf{0}$ then the element is undergoing only elastic deformation and no real dislocation is detected. Otherwise, it is assumed that the core of the detected dislocation is at the centroid of the element under consideration.

At this point, the location of the dislocation core, its Burgers vector and its slip plane are all known. What remains is to 'convert' this dislocation to a discrete dislocation and pass it across the interface.

7.3.3 Passing dislocations

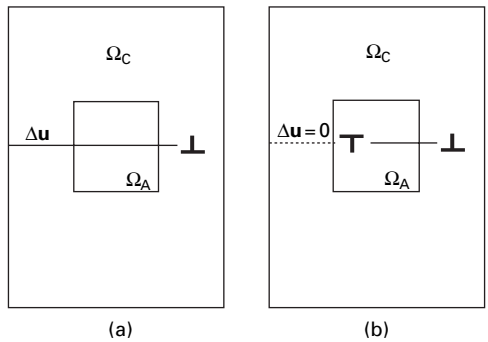
'Passing' a dislocation involves appropriately modifying the displacements of all atoms and nodes in the problem so that the core of the atomistic dislocation is eliminated while retaining the slip displacements. This is achieved by superimposing the continuum elastic displacements of a dislocation dipole onto those of the atoms and nodes, as illustrated in Fig. 7.12. The original



7.12 The dislocation detected in the highlighted detection element in (a) is 'passed' across the atomistic/continuum interface by superimposing the displacement field of the dipole shown in (b). Elements are shown between the atoms in the atomistic region to highlight the effect of the displacement field. Adapted from Shilkrot *et al.* (2004).

dislocation in Fig. 7.12(a) has been identified (location, Burgers vector, slip plane) by the detection band and the superimposed dipole configuration is shown schematically in Fig. 7.12(b). This superposition annihilates the atomistic core at the original position $\mathbf{x}_c$ and replaces it with a continuum dislocation at the position $\mathbf{x}_c + \delta$, which is in the continuum region and is added to the array of continuum dislocations. In practice, it is possible to choose the vector δ to be on the order of a few lattice spacings so that the effect of the superimposed dipole is short-ranged and confined to a few atoms surrounding the original core position.

Note that the displacement fields of an elastic continuum dislocation are not unique, since they require an arbitrary choice of the branch cut across which there is a displacement jump of one Burgers vector. In the atomistic region, however, the branch cut has a physical meaning; it corresponds to the plane along which the dislocation is gliding. As such, this passing procedure requires careful treatment of the continuum fields to ensure that the branch cut is correctly aligned with the slip plane. A second important consideration arises in multiply-connected continuum regions (an atomistic region embedded completely inside a 'hole' in the continuum, such as the problem schematically illustrated in Fig. 7.13). In this case, the continuum displacement field of the new dislocation that has been inserted will extend out the other side of the atomistic region and lead to incorrect results. This is easily remedied by including a discrete dislocation of opposite sign inside the atomistic region, which effectively negates the slip discontinuity along the branch cut of the real defect in the continuum. This 'image' defect is not actually present in the atomistic region, and only enters through the superposition of the dislocation

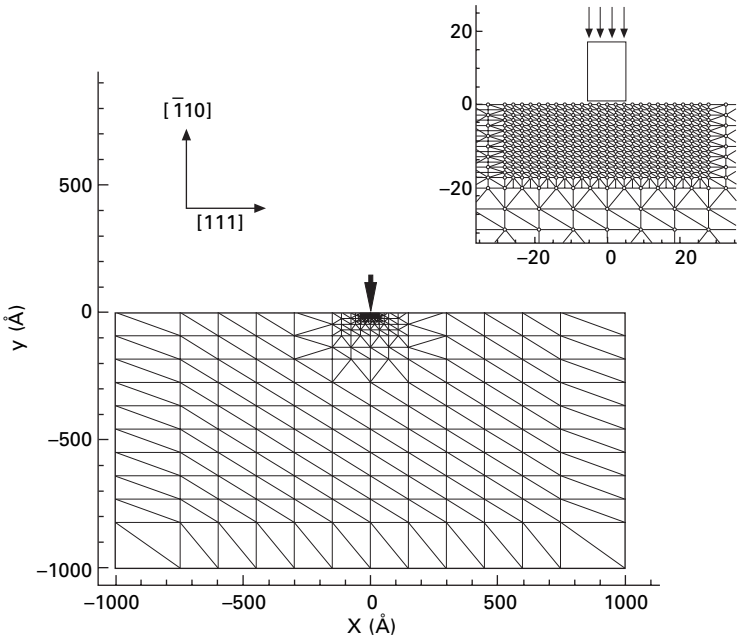


7.13 Schematic illustration of a problem where an atomistic region is embedded in a continuum region. This dislocation in (a) has been nucleated inside the atomistic region, Ω_A, and passed out through the right side. Once it is in the continuum region, its displacement field $\tilde{\mathbf{u}}$ carries a displacement jump along its branch cut all the way to the left edge of the model. In (b), an image dislocation is added inside the 'hole' in the continuum region with the same branch cut as the first dislocation, so that the displacement jumps from the two defects cancel to the left of Ω_A.

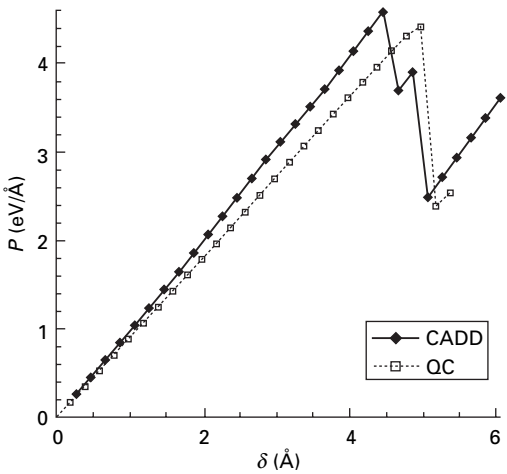
fields leading up to the solution, $\tilde{\mathbf{u}}$, of the complementary problem depicted in Fig. 7.9. This image dislocation *does* have a physical interpretation, however, as one cannot generate a dislocation inside a bulk material without also generating one of opposite sign. Inside the atomistic region, some atomic-scale process has driven the nucleation of the defect and left behind, for example, a surface step or grain boundary dislocation. The continuum needn't know these atomistic details but does require that an image dislocation reside in the atomistic part of the problem. It is possible to show that the inclusion of image dislocations outside the continuum domain of interest in this way still leads to the correct equilibrium solution by the superposition method described in Fig. 7.9.

7.3.4 Illustrative example

To illustrate the CADD method, we have run the same example as we used to illustrate the QC method previously. In this case, the initial mesh is as shown in Fig. 7.14, which models the same fcc Al crystal as the QC simulation, but with a smaller atomistic region as shown. The same quasi-static loading and boundary conditions are used, leading to the load–displacement curve shown in Fig. 7.15. There are a number of differences between the CADD and QC results shown. The most notable is that the CADD result shows two load drops instead of one. This is because in the CADD simulation, the dislocations nucleate first on one side of the indenter, then on the other, whereas the QC simulation saw both nucleated simultaneously. Since nucleation



7.14 The initial mesh for the illustrative indentation example using CADD.

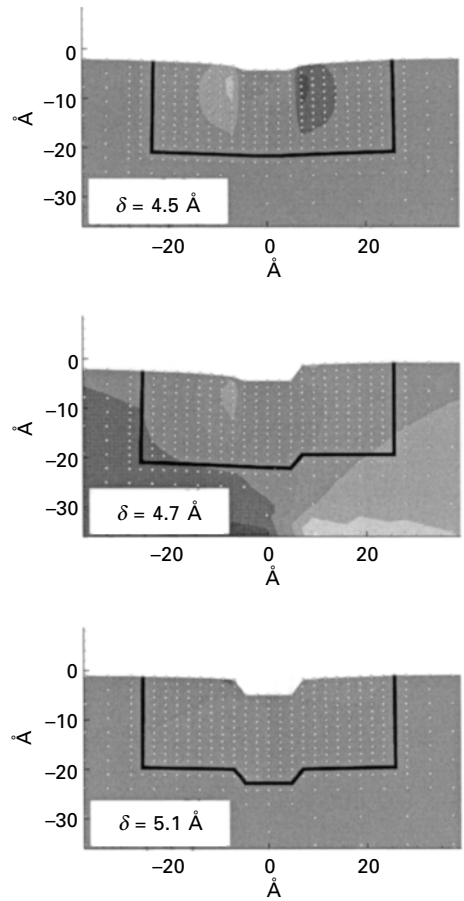


7.15 Load–displacement curve for the CADD indentation example. Previous QC result shown for comparison.

events are manifestations of instabilities in the material, iterative algorithms to solve such problems can be very sensitive to slight numerical noise. As such, the timing and order of nucleation events can be difficult to capture exactly. The second obvious difference between the two curves is their slight

divergence prior to the first load drop, which comes from the fact that QC uses a non-linear elastic continuum and CADD is purely linear.

The region under the indenter is shown just before the defects nucleate and then just after the first and second load drops in Fig. 7.16. The dark line shows the atomistic/continuum interface, and we can clearly see the deformation induced as dislocations cross the interface in figures (b) and (c). The dislocations finally come to rest near the bottom of the crystal, somewhat further down than predicted by QC. This discrepancy arises because the CADD simulation is done with an assumption of zero Peierls stress in the continuum. The QC, on the other hand, has the Peierls stress of the real crystal since the dislocations always reside in the evolving atomistic region.



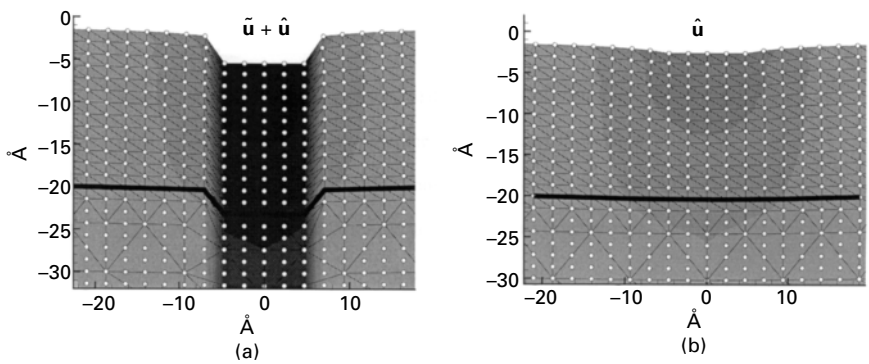
7.16 Deformation beneath the indenter during the CADD simulation. Contours are of the out-of-plane displacement and reveal the incipient dislocations at the indenter corners.

Looking at Fig. 7.16(b) and (c), it appears that the deformation of the linear elastic finite-element region just below the atomistic/continuum interface is rather severe. In fact, it is important to remember that this is really just a convenient visualization. As discussed in Section 7.3.1, the deformation is broken into $\hat{\mathbf{u}} + \tilde{\mathbf{u}}$, the slip discontinuity due to the dislocation is part of $\tilde{\mathbf{u}}$, and only $\hat{\mathbf{u}}$ is experienced by the finite elements. In Fig. 7.17, the full displacement (shown in (a)) is compared with only $\hat{\mathbf{u}}$ [shown in (b)) after subtracting off $\tilde{\mathbf{u}}$ due to the dislocations residing deep in the crystal. Clearly, the strain due to $\hat{\mathbf{u}}$ in the continuum region (below the heavy black line in the figure) is reasonable for a linear elastic domain.

7.4 Challenges and future directions

The QC and CADD methods have been used to study a large number of problems, including nano-indentation (Phillips *et al*., 1999; Tadmor *et al*., 1999a,b; Picu, 2000; Shenoy *et al*., 2000; Smith *et al*., 2000, 2001; Knap and Ortiz, 2001, 2003; Miller *et al*., 2003; Fago *et al*., 2004; Shiari *et al*., 2005), fracture (Miller *et al*., 1998; Pillai and Miller, 2001; Hai and Tadmor, 2003), dislocation interactions (Rodney and Phillips, 1999; Hardikar *et al*., 2001; Mortensen *et al*., 2002) and polarization switching in ferroelectrics (Tadmor *et al*., 2002). Currently, work at extending and improving the methods continues on several fronts.

The QC has been extended to a fully 3D approach (Rodney and Phillips, 1999; Knap and Ortiz, 2001, 2003; Fago *et al*., 2004), however, the issue of mesh adaption during the simulation becomes extremely challenging in this case. Mesh adaption requires frequent and robust Delaunay triangulation of the entire problem, a slow and difficult algorithm in 3D for the highly irregular meshes that arise (imagine, if you will, the 3D equivalent of Fig. 7.6). One



7.17 (a) Total displacement $\tilde{\mathbf{u}} + \hat{\mathbf{u}}$ and (b) $\hat{\mathbf{u}}$ only. The heavy black line shows the atomistic/continuum interface. Contours show relative magnitudes of the vertical (Y) component of displacement.

current area of research aimed at addressing this issue is the use of particle-based 'meshless' methods in place of the finite elements currently used in QC. Kucherov *et al*. (2006) have developed a meshless continuum formulation specifically for use in a multiscale approach and have developed a meshless version of the QC (Kucherev *et al*., 2006; Tadmor *et al*., 2006). The main advantage to this approach is that the meshless formulation does not require a structured tessellation of elements but instead uses shape functions based on a local collection of neighbors called the 'star' of each repatom. With some restrictions that ensure reasonable accuracy, these stars can have a more or less arbitrary size and shape and can be modified locally and independently of neighboring stars. As such, adding repatoms in a region during adaption can be done without expensive re-meshing. In fact, because the method is meshless, no expensive tessellation of the body (Delaunay, Voronoi, etc.) ever needs to be computed.

A 3D implementation of CADD is even more challenging, and is currently not available. The framework of CADD described here, including the solution approach and the detection algorithm, are in principle fully generalizable to 3D. However, the challenge which has not yet been overcome is how to treat the problem of dislocation loops. In 2D, dislocations are effectively point defects, and so passing a defect from the atomistic to the continuum region is a simple matter of moving it from one point to another. However, a 3D CADD model would have dislocation loops nucleate and grow from within the atomistic region. Ultimately, these loops would intersect the atomistic/continuum boundary at several points. Accurately modeling a dislocation loop that is partially atomistic and partially in the continuum, not to mention the book-keeping associated with such an entity, is a significant challenge indeed.

For all multiscale methods, a critical area of recent interest is the question of how to accurately treat dynamics and finite temperature. Much progress has been made, including the coarse-grained molecular dynamics method (Rudd and Broughton, 1998, 2005), and finite temperature extensions of both the CADD method (Qu *et al*., 2005; Shiari *et al*., 2005) and the QC method (Shenoy *et al*., 1999, Dupuy *et al*., 2005). As well, a number of authors have developed so called 'absorbing boundary conditions' (Cai *et al*., 2000; E and Huang, 2001; Park *et al*., 2005a) aimed at the specific problem of spurious wave reflections from the atomistic/continuum interface during dynamic simulations. Specific to the question of modeling mobile dislocations at finite temperature along the lines discussed here, there are a few issues which need to be considered further.

The CADD method depends on a constant monitoring of the strains in the detection band elements and their comparison to the slip library. At finite temperature, this becomes a more difficult task since thermal fluctuations are continually introducing strains in the detection band that can often be

locally and instantaneously very large. The detection algorithm of CADD works reasonably well at low temperatures (less than about 50 K) with almost no change. For these temperatures, the thermal fluctuations are small enough that they only occasionally will falsely trigger the algorithm to detect a dislocation. These false triggers are screened out by simply making sure that a trigger corresponds to a correlated motion along a slip plane, and not a random fluctuation. Specifically, after each trigger an additional step is implemented by making the detection band a few atomic layers thick and confirming that slip occurs in all neighboring elements on a slip plane. At higher temperatures, however, these false triggers happen so frequently as to severely bog down the detection algorithm. One possible approach that is currently being explored is to base the detection algorithm on a time-averaged strain in the detection element rather than an instantaneous strain. The key is to empirically determine a suitable averaging window; it must be long enough to smooth out thermal fluctuations but short enough that it doesn't allow a dislocation to reach the interface undetected.

A similar challenge is faced by QC when we try to implement mesh adaption at finite temperature, and the current finite temperature QC model (Dupuy *et al*., 2005) does not include adaption for this reason. Like the detection algorithm of CADD, the non-locality and adaption criteria described in Section 7.2.3 are based on an instantaneous measure of deformation. However, thermal fluctuations will routinely trigger these criteria in regions where adaption is not desired. As such, current research is aimed at a suitable time-averaged measure of the deformation to allow efficient adaption without false triggers due to thermal fluctuations.

7.5 References

Abraham F.F., Broughton J.Q., Bernstein N. and Kaxiras E., Spanning the length scales in dynamic simulation, *Comput. Phys.* **12**: 538 (1998a).

Abraham F.F., Schneider D., Land B., Lifka D., Skovira J., Gerner J. and Rosenkrantz M., Instability dynamics in three-dimensional fracture: an atomistic simulation, *J. Mech. Phys. Sol.* **45**(9), 1461–1471 (1998b).

Abraham F.F., Walkup R., Gao H.J., Duchaineau M., De la Rubia T.D. and Seager M., Simulating materials failure by using up to one billion atoms and the world's fastest computer: brittle fracture, *Proc. Nat. Acad. Sci. U.S.A.*, **99**(9), 5777–5782, (2002).

Cai W., de Koning M., Bulatov V.V. and Yip S., Minimizing boundary reflections in coupled-domain simulations, *Phys. Rev. Lett.* **85**(15), 3213–3216 (2000).

Cleveringa H.H.M., van der Giessen E. and Needleman A., Comparison of discrete dislocation and continuum plasticity predictions for a composite material, *Acta Mater.* **45**(8), 3163–3179 (1997).

Curtin W.A. and Miller R.E., Atomistic/continuum coupling methods in multi-scale materials modeling, *Model. Simul. Mater. Sci. Eng.* **11**(3), R33–R68 (2003).

Datta D.K., Catalin Picu R. and Shephard M.S., Composite grid atomistic continuum method: an adaptive approach to bridge continuum with atomistic analysis, *Int. J. Multiscale Comput. Eng*. **2**(3), 71–90 (2004).

Dewald M. and Curtin W.A., private communication, (2005).
Dupuy L.M., Tadmor E.B., Miller R.E. and Phillips R., Finite temperature quasicontinuum: molecular dynamics without all the atoms, *Phys. Rev. Lett.* **95**, 060202 (2005).
E Weinan and Huang Zhongyi, Matching conditions in atomistic-continuum modeling of materials, *Phys. Rev. Lett.* **87**(13), 135501 (2001).
Ercolessi F. and Adams J.B., Interatomic potentials from first-principles calculations – the force-matching method, *Europhys. Lett.* **26**, 583–588 (1994).
Ericksen J.L., in *Phase Transformations and Material Instabilities in Solids*, ed. M. Gurtin, New York: Academic Press, 1984, 61–77.
Fago M., Hayes R.L., Carter E.A. and Ortiz M., Density-functional-theory-based local quasicontinuum method: prediction of dislocation nucleation, *Phys. Rev. B* **70**(10), 100102 (2004).
Finnis M., *Interatomic Forces in Condensed Matter*, Oxford: Oxford University Press, 2003.
Fivel M.C. and Canova G.R., Developing rigorous boundary conditions to simulations of discrete dislocation dynamics, *Mod. Simul. Mater. Sci. Eng.* **7**(5), 753–768 (1999).
Gerberich W.W., Nelson J.C., Lilleodden E.T., Anderson P. and Wyrobek J.T., Indentation induced dislocation nucleation: the initial yield point, *Acta Mater.* **44**(9) 3585–3598 (1996).
Gerberich W.W., Tymiak N.I., Grunlan J.C., Horstemeyer M.F. and Baskes M.I., Interpretations of indentation size effects, *J. Appl. Mech.* **69**, 433–442 (2002).
Hai S. and Tadmor E.B., Deformation twinning at aluminum crack tips, *Acta Mater.*, **51**(1), 117–131 (2003).
Hardikar K., Shenoy V. and Phillips R., Reconciliation of atomic-level and continuum notions concerning the interaction of dislocations and obstacles, *J. Mech. Phys. Sol.* **49**, 1951–1967 (2001).
Kelchner C.L., Plimpton S.J. and Hamilton J.C., Dislocation nucleation and defect structure during surface indentation, *Phys. Rev. B* **58**(17), 11085–11088 (1998).
Knap J. and Ortiz M., An analysis of the quasicontinuum method, *J. Mech. Phys. Sol.* **49**, 1899–1923 (2001).
Knap J. and Ortiz M., Effect of indenter-radius size on au(001) nanoindentation, *Phys. Rev. Lett.* **90**, 226102 (2003).
Kubin L.P. and Canova G., The modelling of dislocation patterns, *Scripta Met.* **27**, 957–962 (1992).
Kucherov L., Tadmor E.B. and Miller R.E., Umbrella spherical integration: a stable meshless method for nonlinear solids, *Int. J. Numer. Meth. Eng.* in press (2006).
Li J., Van Vliet K.J., Zhu T., Yip S. and Suresh S., Atomistic mechanisms governing elastic limit and incipient plasticity in crystals, *Nature* **418**, 307–310 (2002).
Lilleodden E.T., Zimmerman J.A., Foiles S.M. and Nix W.D., An experimental and computational study of the elastic-plastic transition in thin films, *Mat. Res. Soc. Symp. Proc.* **673**, P1.3.1–P1.3.6 (2001).
Lilleoden E.T., Zimmerman J.A., Foiles S.M. and Nix W.D., Atomistic simulations of elastic deformation and dislocation nucleation during nanoindentation, *J. Mech. Phys. Solids.* **51**, 901–920 (2003).
Liu W.K., Karpov E.G., Zhang S. and Park H.S., An introduction to computational nano-mechanics and materials, *Comput. Meth. Appl. Mech. Eng.* **193**, 1529–1578 (2004).
Miller R. and Phillips R., Critical analysis of local constitutive models for slip and decohesion, *Phil. Mag. A* **73**(4), 803–827 (1996).
Miller R., Tadmor E.B., Phillips R. and Ortiz M., Quasicontinuum simulation of fracture at the atomic scale, *Model. Simul. Mater. Sci. Eng.* **6**, 607–638 (1998).

Miller R.E. and Tadmor E.B., The quasicontinuum method: overview, applications and current directions, *J. Comput-Aided Mater. Des.* **9**(3), 203–23, (2002).

Miller R.E., Shilkrot L.E. and Curtin W.A., A coupled atomistic and discrete dislocation plasticity simulation of nano-indentation into single crystal thin films, *Acta Mater.* **52**(2), 271–284 (2003).

Mortensen J.J., Schiøtz J. and Jacobsen K.W., The quasicontinuum method revisited, *Challenges Mol. Simul.*, **4**, 119–136 (2002).

Olmsted D.L., Hardikar K.Y. and Phillips R., Lattice resistance and Peierls stress in finite size atomistic dislocation simulations, *Model. Simul. Mater. Sci. Eng.* **9**, 215–247 (2001).

Park H.S., Karpov E.G. and Liu W.K., Non-reflecting boundary conditions for atomistic, continuum and coupled atomistic/continuum simulations, *Int. J. Numer. Methods Eng.* **64**(2), 237–259, (2005a).

Park H.S., Karpov E.G., Klein P.A. and Liu W.K., Three-dimensional bridging scale analysis of dynamic fracture, *J. Comput. Phys.* **207**, 588–609, (2005c).

Park H.S., Karpov E.G., Liu W.K. and Klein P.A., The bridging scale for two-dimensional atomistic/continuum coupling, *Phil. Mag. A*. **85**(1), 79–113 (2005b).

Phillips R., Rodney D., Shenoy V., Tadmor E.B. and Ortiz M., Hierarchical models of plasticity: Dislocation nucleation and interaction, *Model. Simul. Mater. Sci. Eng.* **7**, 769–780 (1999).

Picu P.C., Atomistic-continuum simulation of nano-indentation in molybdenum, *J. Comput.-Aided Mater. Des.* **7**, 77–87 (2000).

Pillai A.R. and Miller R.E., Crack behaviour at bi-crystal interfaces: a mixed atomistic and continuum approach, *Mater. Res. Soc. Symp. Proc.* **653**, Z2.9.1–Z2.9.7 (2001).

Qu S., Shastry V., Curtin W.A. and Miller R.E., A finite temperature, dynamic, coupled atomistic/discrete dislocation method, *Model. Simul. Mater. Sci. Eng.* **13**(7), 1101–1118 (2005).

Rodney D. and Phillips R., Structure and strength of dislocation junctions: an atomic level analysis, *Phys. Rev. Lett.* **82**(8), 1704–1707 (1999).

Rodriguez de la Fuente O., Zimmerman J.A., Gonzalex M.A., de la Figuera J., Hamilton J.C., Wu Pai Woei and Rojo J. M., Dislocation emission around nanoindentations on a (001) fcc metal surface studied by scanning tunneling microscopy and atomistic simulations, *Phys. Rev. Lett.* **88**(3), 036101 (2002).

Rudd R.E. and Broughton J.Q., Coarse-grained molecular dynamics and the atomic limit of finite elements, *Phys. Rev. B* **58**(10), R5893–R5896 (1998).

Rudd R.E. and Broughton J.Q., Coarse-grained molecular dynamics: nonlinear finite elements and finite temperatures, *Phys. Rev. B* **72**, 144104 (2005).

Rudd R.E. and Broughton J.Q., Concurrent coupling of length scales in solid state systems, *Phys. Status Solidi B* **217**, 251–291 (2000).

Saraev D. and Miller R.E., Atomic-scale simulations of nanoindentation-induced plasticity in copper crystals with nanometer-sized nickel coatings, *Acta Mater.* **54**(1), 33–45 (2006).

Saraev D. and Miller R.E., Atomistic simulation of nanoindentation into copper multilayers, *Model. Simul. Mater. Sci. Eng.* **13**, 1089–1099 (2005).

Schiotz J. and Jackobsen K.W., A maximum in the strength of nanocrystalline copper, *Science* **301**, 1357–1359 (2003).

Schiotz J., Di Tolla F.D. and Jackobsen K.W., Softening of nanocrystalline metals at very small grain sizes, *Nature* **391**, 561–563 (1998).

Shenoy V., Shenoy V. and Phillips R., Finite temperature quasicontinuum methods, *Mater. Res. Soc. Symp. Proc.* **538**, 465–471 (1999).

Shenoy V.B., Miller R., Tadmor E.B., Phillips R. and Ortiz M., Quasicontinuum models of interfacial structure and deformation, *Phys. Rev. Lett.* **80**(4), 742–745 (1998a).

Shenoy V.B., Miller R., Tadmor E.B., Rodney D., Phillips R. and Ortiz M., An adaptive methodology for atomic scale mechanics: the quasicontinuum method, *J. Mech. Phys. Sol.* **47**, 611–642 (1998b).

Shenoy V.B., Phillips R. and Tadmor E.B., Nucleation of dislocations beneath a plane strain indenter, *J. Mech. Phys. Solids*, **48**, 649–673 (2000).

Shiari B., Miller R. and Curtin W.A., Coupled atomistic/discrete dislocation simulations of nanoindentation at finite temperature, *Trans ASME*, *ASME J. Eng. Mater. Technol.* **127**(4), 358–368 (2005).

Shilkrot L.E., Miller R.E. and Curtin W.A., Coupled atomistic and discrete dislocation plasticity, *Phys. Rev. Lett.* **89**(2), 025501 (2002).

Shilkrot L.E., Miller R.E. and Curtin W.A., Multiscale plasticity modeling: coupled atomistic and discrete dislocation mechanics, *J. Mech. Phys. Solids* **52**(4), 755–787 (2004).

Smith G.S., Tadmor E.B. and Kaxiras E., Multiscale simulation of loading and electrical resistance in silicon nanoindentation, *Phys. Rev. Lett.* **84**(6), 1260–1263 (2000).

Smith G.S., Tadmor E.B., Bernstein N. and Kaxiras E., Multiscale simulations of silicon nanoindentation, *Acta Mater.* **49**(19), 4089–4101 (2001).

Tadmor E.B. and Miller R.E., The theory and implementation of the quasicontinuum method, in *Handbook of Materials Modeling, Volume I (Methods and Models)*, ed. S. Yip, New York: Springer Science and Business Media, Chapter 2.13, (2006).

Tadmor E.B., Miller R., Phillips R. and Ortiz M., Nanoindentation and incipient plasticity, *J. Mater. Res.* **14**(6), 2233–2250 (1999).

Tadmor E.B., Miller R.E. and Kucherov L., Smoothed atom mechanics: a mesh-free reformulation of the quasicontinuum method in preparation (2006).

Tadmor E.B., Ortiz M. and Phillips R., Quasicontinuum analysis of defects in solids, *Phil. Mag. A* **73**(6), 1529–1563 (1996).

Tadmor E.B., Smith G.S., Bernstein N. and Kaxiras E., Mixed finite element and atomistic formulation for complex crystals, *Phys. Rev. B* **59**(1), 235–245 (1999b).

Tadmor E.B., Waghmare U.V., Smith G.S. and Kaxiras E., Polarization switching in $PbTiO_3$: An *ab initio* finite element simulation, *Acta Mater.*, **50**, 2989–3002 (2002).

Valiev R., Nanomaterial advantage, *Nature* **419**, 887–889 (2002).

Van Swygenhoven H., Derlet P.M. and Froseth A.G., Stacking fault energies and slip in nanocrystalline metals, *Nat. Mater.* **3**, 399–403 (2004).

Van Swygenhoven H., Grain boundaries and dislocations, *Science* **296**, 66–67 (2002).

Van Vliet K.J., Li J., Zhu T., Yip S. and Suresh S., Quantifying the early stages of plasticity through nanoscale experiments and simulations, *Phys. Rev. B* **67**, 104105 (2003).

van der Giessen E. and Needleman A., Discrete dislocation plasticity: a simple planar model, *Model. Simul. Mater. Sci. Eng.* **3**, 689–735 (1995).

Wagner G.J. and Liu W.K., Coupling of atomistic and continuum simulations using a bridging scale decomposition, *J. Comput. Phys.* **190**, 249–274, (2003).

Weygand D., Friedman L.H., van der Giessen E. and Needleman A., Aspects of boundary-value problem solutions with three-dimensional dislocation dynamics, *Model. Simul. Mater. Sci. Eng.* **10**(4), 437–468 (2002).

Yamakov V., Wolf D., Phillpot S.R., Mukherjee A.K. and Gleiter H., Deformation-mechanism map for nanocrystalline metals by molecular-dynamics simulation, *Nat. Mater.* **3**, 43–47 (2004).

Yamakov V., Wolf D., Phillpot S.R., Mukherjee A.K. and Gleiter H., Dislocation processes in the deformation of nanocrystalline aluminium by molecular-dynamics simulation, *Nat. Mater.* **1**: 1–4 (2002).

Yashar P.C. and Sproul W.D., Nanometer scale multilayered hard coatings, *Vacuum* **55**, 179–190 (1999).

Zbib H.M., Rhee M. and Hirth J.P., On plastic deformation and the dynamics of 3d dislocations, *Int. J. Mech. Sci.* **40**, 113–127 (1998).

Zienkiewicz O.C. and Zhu J.Z., A simple error estimator and adaptive procedure for practical engineering analysis, *Int. J. Num. Meth. Eng.* **24**, 337–357 (1987).

Zimmerman J.A., Kelchner C.L., Klein P.A., Hamilton J.C. and Foiles S.M., Surface step effects on nanoindentation, *Phys. Rev. Lett.* **87**(16), 165507 (2001).

8

Multiscale modelling of carbon nanostructures

T. Y. NG, S. H. YEAK and Y. X. REN, Nanyang Technological University, Singapore; K. M. LIEW, City University of Hong Kong, Hong Kong

8.1 Introduction to carbon nanotube dynamics

Carbon nanotubes (CNTs) were first discovered by Iijima[1] in 1991 and that has since stimulated very broad and intense research into the syntheses and theoretical analyses of CNTs and their applications. This is due to excellent mechanical properties such as small size, low density, high stiffness, high strength, and excellent electronic properties. Yakobson *et al.*[2] studied CNT behavior under axial compression using the Tersoff–Brenner potential. Their simulations showed that at large deformations, an abrupt release of energy is accompanied by a reversible switch into a different morphological pattern. Yakobson *et al.*[3] also studied CNT behavior under high rate tension using Tersoff–Brenner's reactive empirical bond-order (REBO) potential (see Brenner *et al.*[4]). Their research revealed the ability of CNTs to undergo large elastic deformation when subjected to axial tension.

The above-mentioned investigations were carried out using empirical molecular dynamics. However, as pointed out by Abraham *et al.*[5,6,7], a more refined description is necessary, especially in the bond-breaking area. In other words, potentials due to electron–electron, electron–ion as well as ion–ion interactions should be considered. In this respect, the tight-binding (TB) method has the advantage of being quantum mechanical, taking into account the kinetic energies of both the ions and electrons, as well as the electron–electron, electron–ion, and ion–ion interactions.

In this chapter, a multiscale model is developed by introducing near and far regions in order to achieve a seamless coupling between molecular dynamic (MD) and TB. We will examine the elastic and plastic properties of single-walled carbon nanotubes (SWCNTs) under axial compression and tension. This is carried using MD as well as a multiscale technique where a handshaking region between MD and TB is described and implemented. The interaction forces between the carbon atoms are calculated based on the second-generation REBO potential, TB-derived forces as well as long-range Lennard-Jones potential. A smooth cutoff Lennard-Jones with switch function is also proposed

and investigated. The viability of the presently developed handshaking region between MD and TB in CNTs under axial compression and tension will be subjected to validation studies. The detection of sideways buckling due to the asymmetrical axial compression will be examined in detail. This sideways buckling phenomenon is observed when using both pure MD and MD/TB multiscale models.

8.2 Overlap TB/MD multiscale model

8.2.1 Theoretical development

In this theoretical framework, the short-range interaction force between atoms is modeled using the second-generation REBO potential of Brenner *et al.*[4] In addition, the van der Waals potential (Mao *et al.*[8] and Sinnott *et al.*[9]) is applied. The potential sum is thus

$$E = \sum_i \sum_{j>i} (E_{ij}^{\text{REBO}} + E_{ij}^{\text{vdw}}) \qquad [8.1]$$

$$E_{ij}^{\text{REBO}} = [V_{\text{R}}(r_{ij}) + b_{ij} V_{\text{A}}(r_{ij})] \qquad [8.2]$$

where V_{R} and V_{A} are pair-additive interactions that represent all interatomic repulsion and attraction from valence electrons, and b_{ij} is the reactive empirical bond order between atoms.

For SWCNTs, the Lennard-Jones 12–6 potential [10] is used

$$E_{ij}^{\text{vdw}} = 4\varepsilon \left[\left(\frac{\sigma}{r_{ij}} \right)^{12} - \left(\frac{\sigma}{r_{ij}} \right)^{6} \right] \qquad [8.3]$$

As pointed out by Mao *et al.*[8], it is advisable to incorporate the van der Waals potential only if the short-distance potential becomes zero. This is to prevent an artificial reaction barrier from forming due to the steep repulsive wall of the Lennard-Jones 12–6 potential in the short range. If E_{vdw} were activated in the short range as well, the sum effect of ($E_{\text{REBO}} + E_{\text{vdw}}$) will be erroneously high due to the additional accounting, thus forming the artificial force barrier that restricts unbonded atoms from undergoing chemical reaction. Theoretically, the above barrier will produce non-smooth potential energy profiles and inter-atomic forces, leading to a non-smooth optimization problem during relaxation. Since the non-smooth optimization algorithm is very time-consuming, it is proposed here to modify the above van der Waals potential such that it behaves smoothly around the cutoff distance, D_{max}. The current work employs the conjugate-gradient method and truncated-Newton via Lanczos minimization. Both methods are based on smooth objective minimization using the gradient of the objective function.

The original equation can be modified by including switch function $S(r)$ as shown below

$$E_{\text{vdw}}^{\text{s}} = E_{ij}^{\text{vdw}} \cdot S(r) \tag{8.4}$$

which must satisfy the following conditions

$$\left.\begin{aligned} E_{\text{vdw}}^{\text{s}}(D_{\max}) &= 0 \\ E_{\text{vdw}}^{\text{s}}(D_{\max}+d) &= E_{ij}^{\text{vdw}}(D_{\max}+d) \\ \frac{\partial}{\partial r} E_{\text{vdw}}^{\text{s}}\Big|_{D_{\max}} &= 0 \\ \frac{\partial}{\partial r} E_{\text{vdw}}^{\text{s}}\Big|_{D_{\max}+d} &= \frac{\partial}{\partial r} E_{ij}^{\text{vdw}}\Big|_{D_{\max}+d} \end{aligned}\right\} \tag{8.5}$$

where $D_{\max}$ is the distance at which the REBO potential becomes zero, and d is the neighbor distance that causes the switch function to vary from zero to one. The switch function is cubic polynomial and can be calculated by solving equation [8.5] simultaneously. Finally, we derive a smooth cutoff van der Waals potential as

$$E_{\text{vdw}} = \begin{cases} 0 & r < D_{\max} \\ E_{ij}^{\text{vdw}} \cdot S(r) & D_{\max} \le r \le D_{\max} + d \\ E_{ij}^{\text{vdw}} & r > D_{\max} + d \end{cases} \tag{8.6}$$

For the quantum mechanical approach, semi-empirical TB is employed.[11,12] The electronic structure of the simulated CNTs is calculated by a TB Hamiltonian so that the quantum mechanical many-body nature of the interatomic forces is naturally taken into account.

The Hamiltonian of a system of ion cores and valence electrons can be written in the adiabatic approximation as

$$H_{\text{tot}} = T_i + T_{\text{e}} + U_{\text{ee}} + U_{\text{ei}} + U_{\text{ii}} \tag{8.7}$$

where T_{i}, T_{e}, U_{ee}, U_{ei}, and U_{ii} are respectively the kinetic energies of the ions and electrons, and the potential energies due to electron–electron, electron–ion, and ion–ion interactions. Within the adiabatic one-electron assumption, the many-body electron Hamiltonian can be reduced to that of one electron moving in the average field due to the other valence electrons and ions. The reduced one-electron Hamiltonian H and its nth eigenfunction $|\Psi_n\rangle$ can therefore be written as

$$H|\Psi_n\rangle = \varepsilon_n |\Psi_n\rangle \tag{8.8}$$

where ε_n is the energy of the nth single-particle state. The wavefunctions $|\Psi_n\rangle$ can be approximated by linear combination

$$|\Psi_n\rangle = \sum_{l\alpha} c_{l\alpha}^n |\phi_{l\alpha}\rangle \quad [8.9]$$

where l is the quantum number index and α labels the ions. The computational load will increase significantly if the basis set, $|\phi_{l\alpha}\rangle$ is not orthogonal. However, it is possible to obtain a new orthogonal basis set $\{\varphi_{l\alpha}\}$ through the Löwdin transform

$$|\Psi_n\rangle = \sum_{l\alpha} b_{l\alpha}^n |\varphi_{l\alpha}\rangle \quad [8.10]$$

$$Hb^n = \varepsilon_n b^n \quad [8.11]$$

where $b^n = \{b_{l\alpha}^n\}$.

Finally, the Schrödinger equation for the single-particle states becomes

$$\sum_{l'\beta} (\langle\varphi_{l'\beta}|H|\varphi_{l\alpha}\rangle - \varepsilon_n \delta_{ll'}\delta_{\alpha\beta})\, c_{l'\beta}^n = 0 \quad [8.12]$$

In the general approach, the matrix elements in equation [9.12] are calculated after fitting a suitable database obtained either from experiments or by first principles calculation. Once the single-particle energies are known by solving equation [9.12], the total energy of ion cores and valence electrons can be written as

$$E_{\text{tot}} = \sum_n \varepsilon_n f(\varepsilon_n, T) + U_{\text{ii}} - U_{\text{ee}} = E_{\text{bs}} + U_{\text{rep}} \quad [8.13]$$

where $f(\varepsilon_n, T)$ or the band structure energy, E_{bs}, are the Fermi–Dirac distribution functions. The U_{ee} term corrects the double counting of the electron–electron interactions in the first term. The last two terms constitute the effective repulsive potential, $U_{\text{rep}} = U_{\text{ii}} - U_{\text{ee}}$. This repulsive potential can be expressed as a sum of two-body potentials as

$$U_{\text{rep}} = U_{\text{ii}} - U_{\text{ee}} = \sum_{\alpha,\beta>\alpha} \Phi(r_{\alpha\beta}) \quad [8.14]$$

where $r_{\alpha\beta}$ is the distance between atoms located at α and β. This pairwise potential $\Phi(r_{\alpha\beta})$ between atoms at α and β can be solved using the transferable TB potential introduced by Xu *et al.*[13].

The forces $\vec{f}_\alpha (\alpha = 1, 2, \ldots, N_{at})$ that are required in the simulation can be calculated from the Hamiltonian H_{TBMD} as

$$H_{\text{TBMD}} = \sum_\alpha \frac{p_\alpha^2}{2m_\alpha} + \sum_n \varepsilon_n f(\varepsilon_n, T) + U_{\text{rep}} \quad [8.15]$$

and the forces are given as

$$\vec{f}_\alpha = -\sum_n \left\langle \Psi_n \left| \frac{\partial H}{\partial \bar{r}_\alpha} \right| \Psi_n \right\rangle f(\varepsilon_n, T) - \frac{\partial U_{\text{rep}}}{\partial \bar{r}_\alpha} \quad [8.16]$$

It is found that the second term in Eq. [8.16] is short ranged and can be solved analytically. The first term, however, is solved using the Hellmann–Feynman theorem

$$\sum_n \left\langle \Psi_n \left| \frac{\partial H}{\partial \bar{r}_\alpha} \right| \Psi_n \right\rangle f(\varepsilon_n, T)$$

$$= -2 \sum_n f(\varepsilon_n, T) \sum_{l\gamma} \sum_{l'\beta} c^n_{l'\beta} \frac{\partial H_{l'\beta, l\gamma}(r_{\beta\gamma})}{\partial \bar{r}_\alpha} c^n_{l\gamma} \quad [8.17]$$

where

$$H_{l'\beta, l\gamma}(r_{\beta\gamma}) = g(r_{\beta\gamma}) \langle \varphi_{l'\beta} \mid H \mid \varphi_{l\gamma} \rangle \mid_{r_{\beta\gamma} = r_0} \quad [8.18]$$

and $g(r_{\beta\gamma})$ is a scaling function. It is important to point out that the Hellmann–Feynman forces require the full set of eigen-system solutions and are therefore very computing time intensive.

For the subsequent optimization process, the truncated-Newton method will be effective if only a small number of inner iterations is sufficient to produce a converged step. The implementation of preconditioning will accelerate the convergence of the inner iterations.

The present truncated-Newton method will minimize a twice continuously-differentiable function, $f(x)$. The first-order optimality condition leads to

$$\nabla f(x) = 0 \quad [8.19]$$

$$\nabla^2 f(x_x) p_x = -\nabla f(x_x) \quad [8.20]$$

$$x_{k+1} = x_k + p_k \quad [8.21]$$

The truncated-Newton method is embedded with the following techniques:

- conjugate-gradient method to solve large system, equations [8.20] and [8.21], iteration is truncated before the exact solution is attained.
- line-search method:
  for $k = 0,1,\ldots$
    stop if stopping rule satisfied
    compute a search direction p_k
    determine an improved estimate, $x_{k+1} = x_k + \alpha_k p_k$.
- Trust-region method:
  for $k = 0,1,\ldots$
    stop if stopping rule satisfied

choose p_k so as to minimize $y_k(p) \approx f(x_k + p)$,
subject to $\| p \| \leq \Delta_k$,
compute x_{k+1} and Δ_{k+1} using p_k.

- Lanczos shift is used when Hessian matrix, $\nabla^2 f(x)$, is indefinite.

8.2.2 Concurrent TB/MD overlap coupling

Sherwood[14] demonstrated a successful multiscale approach in the form of a hybrid quantum mechanics/molecular mechanics (QM/MM) model. They introduce outer, boundary, and inner regions in order to achieve seamless coupling. Their subtractive schemes have the advantage of relative simplicity of implementation, with no requirement for validating the QM/MM interactions.

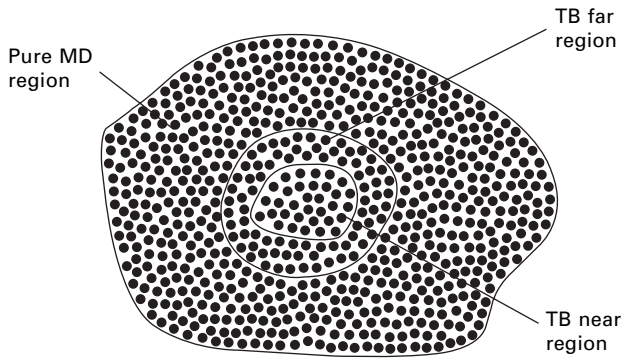
Rafii-Taber *et al.*[15] introduced a multiscale model for simulating brittle-crack propagation. Their model seamlessly couples the crack dynamics at the macroscales and nanoscales via an intermediate mesoscale continuum using a combination of finite elements (FE) and MD. Abraham *et al.*[5,6] studied the rapid brittle fracture of a silicon slab, flawed by a central microcrack and subject to uniaxial tension. They labeled the atomic length scale as the mesoscopic regime, but pointed out that treating bond breaking with an empirical MD potential may be questionable, and a quantum-mechanical treatment is highly desirable.

A multiscale method which successfully couples FE or mesh-free methodology with MD was introduced by Dong *et al.*[16]. A multiple-scale decomposition of the atomic displacements in terms of FE nodal displacements and MD displacements is first carried out. The total scale is then calculated based on the usual FE interpolation, MD displacements, and the projection of the MD displacements onto the FE basis. Dong *et al.*[17] also introduced a virtual atom cluster (VAC) in the coarse scale treatment, where the number of quadrature points used in the VAC is far less than the actual number of atoms. An 'isoparametric-like' mesh-free approximation was also formulated, and this approximation was found to be valid for interpolating a general class of low-dimensional nano-structured materials such as 2D graphite.

In the presently developed multiscale approach, interpolation scaling is not required. Both the MD and TB methods utilize the information available at the atomic locations. However, unlike the TB method, the MD method does not consider the contributions of valence electrons.

As shown in Fig. 8.1, the TB domain comprises the near and far regions. The far region constitutes a relatively small sub-domain and is used to achieve a seamless coupling between the TB (near region) and MD potentials.

The MD method is applied to both the pure MD and TB far regions. The resultant forces and motions of the atoms in these two regions are therefore determined by the MD method. The TB method is applied to both the TB near and far regions, but only the motions of the atoms in the near region are



8.1 Terminology for regions used in the TB/MD overlap multiscale model.

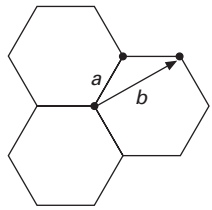
determined by the TB method. The model is therefore not affected by the periodic boundary conditions of the TB analysis as the motions of the atoms in the TB far region are based on MD calculations. For the TB analysis of the TB near and far regions, the forces acting on the atoms at the boundary of the TB far region are derived from the MD analysis. The average width of the far region is 2.6 Å since it is the standard cutoff radius of carbon. With the introduction of this far region, the concept of a 'silogen atom' is not required in the present model. Silogen is a terminology used by Abraham *et al.*[5,6] to describe TB terminating silicon atoms which bond like silicon but are monovalent like hydrogen.

8.2.3 Nearest distance and density of point

The nearest mutual distance between two carbon atoms will determine whether their bond is single, double, or if it is non-existent (broken). Thus, it is useful to calculate the nearest distance among carbon atoms. The nearest distant is indicated as ndist(i, j). The ndist(i, j) table is derived based on the double partial sorting of the entire CNT subject to the constraint of the order $i < j$, where i and j are the indexes of the atoms. The constraint is required to avoid the double counting of each pair of atoms. Initially in its undeformed state, each atom in the CNT will have three nearest mutual distances, and two for those atoms at the boundary. However, during the simulation, as the CNT deforms, each atom will have at least one nearest mutual distance. From the standard graphene pentagon structure (see Fig. 8.2), we have

$$b = 2a \cdot \sin(60°) \approx 1.732a \qquad [8.22]$$

All the distances larger than or equal to the designated nearest distance cutoff value of $1.732a$ will be excluded.



8.2 Nearest distances *a* and *b*.

The density of point (*DOP*) is defined here and is calculated so as to detect the buckling, kink shape occurrences as well as bond breaking. All the distances larger than or equal to a designated cutoff distance should be excluded. For instance, if we set the cutoff radius as 2.6 Å, each carbon atom will initially have a *DOP* value of 9, indicating that it has nine neighboring carbons within the cutoff radius. However, when the SWCNTs start to buckle sideways, it is expected that at the regions of relative larger deformations, some carbon atoms will have *DOP* values higher than 9. This region will be automatically detected and designated as the near region.

8.3 Simulation results of carbon nanotubes under axial loading

8.3.1 Validation of simulation results and effects of switch function

In our validation, we apply a velocity of 20 m/s at both ends in order to obtain the stress–strain relationship. The atoms located at the ends of the CNT are moved according to this velocity at very small time steps of 1 femto-second (fs). In this MD simulation, the locations of the atoms are calculated according to Gear's fifth-order predictor–corrector algorithm. The whole tube was relaxed by truncated-Newton with Lanczos algorithm or conjugate-gradient method in order to minimize the energy of CNTs but without affecting the atoms at the two ends.

To validate the presently developed multiscale model, we carry out several comparisons with published data of existing methodologies. Of interest in these comparisons are the Young's modulus and strain energy.

In the microcanonical ensemble MD simulations of CNTs, strain is derived as

$$\varepsilon = |(L - L_0)|/L_0 \qquad [8.23]$$

where L_0 and L are the undeformed and deformed lengths of the CNT. The stress is calculated as

$$\sigma = F/S \qquad [8.24]$$

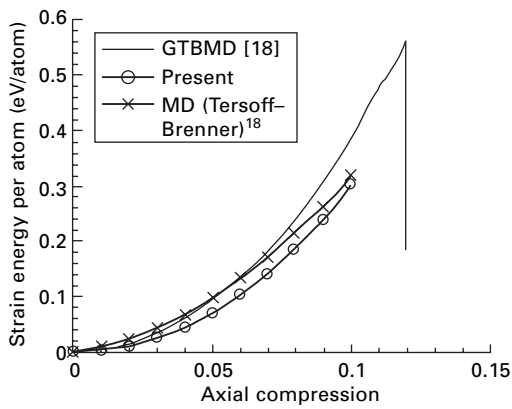
where F is axial force and the cross-sectional area is given by $S = \pi dh$, with d being the CNT diameter and h the thickness, which is usually taken to be 0.34 nm.

For the validation, we compare the strain energy of SWCNT (8,0) undergoing axial compression. This comparison is made against results obtained via the quantum generalized tight binding molecular dynamics (GTBMD) method, as well as MD results computed with Tersoff–Brenner potential reported by Srivastava[18] in Fig. 8.3. The reported results show that collapse occurs at the respective strains of 0.12 and 0.8–0.9 for GTBMD and MD (using Tersoff–Brenner potential), whereas the present MD (using second- generation REBO potential) yields a collapse strain of 0.1. The present results are thus in reasonable agreement and are actually more refined than those obtained using the Tersoff–Brenner potential.

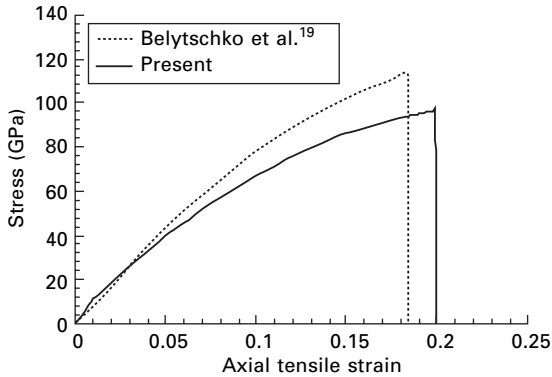
Validation of the present MD algorithm is also made for SWCNTs under axial tension. We examine the stress–strain relations of SWCNT (12,12) with length-to-diameter ratio, $L/D = 9.1$. The comparison is made with the MD (using modified Morse potential) results of Belytschko *et al.*[19], as shown in Fig. 8.4. The present results yield a slightly higher collapse strain at a lower stress value.

The third comparison carried out is for the Young's modulus and is made against the *ab initio* results of Kudin and Scuseria[20]. These results are tabulated in Table 8.1, and it can be observed that the two sets of results are comparable with an average discrepancy of about 5 %.

Next we investigate the effects of implementing the switch function into the Lennard-Jones potential. Obviously, the existence of this switch function will not affect the results for a SWCNT under axial tension, where all the atoms are moving away from each other. However, we would expect some



8.3 Comparison of strain energy curve of SWCNT (8,0) under axial compression.



8.4 Comparison of stress–strain curve of a SWCNT (12,12) under axial tension.

Table 8.1 Comparison of Young's modulus

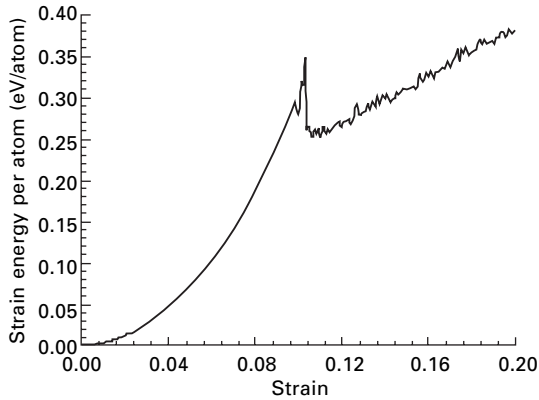
CNT geometry	Young's modulus ($\propto d^2E/d\varepsilon^2$)	
	Present	*ab initio*[20]
(4,4)	53.1	56.4
(7,0)	54.6	56.3
(7,7)	52.7	56.5
(12,0)	53.0	55.2

differences for the case of axial compression. From Figs 8.5 to 8.8, we observe that this is indeed the case.

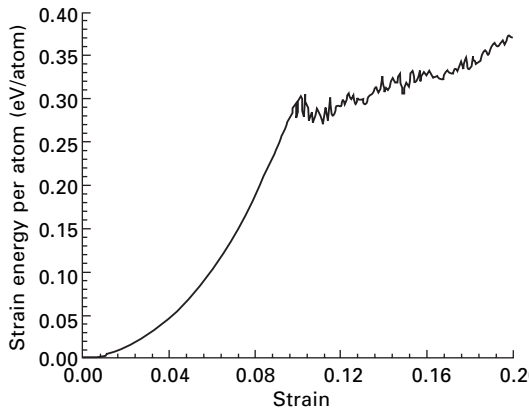
According to Figs 8.5 to 8.8, we find that both sets of results have the same collapse strain of 0.1. However, we also find that after the initial collapse, results via the smooth cutoff Lennard-Jones show lower strain energy as well as less fluctuation. It is also the same case for maximum distance ndist (i, j) where results using the smooth cutoff Lennard-Jones show less fluctuation in maximum distance. Obviously, this switch function plays an important role immediately after the initial collapse. This observation was first reported by the present authors in Yeak *et al.*[21].

8.3.2 Carbon nanotubes subject to axial tension

In order to compare the two presently developed MD and TB/MD schemes, we shall simulate the case of a SWCNT (7,7) with length-to-diameter ratio $L/D = 7.1$, and under axial tension. In this problem, the total number of atoms is 784, with 616 atoms located in the MD region when multiscaling first occurs. First we perform the pure MD simulation and examine the elastic and plastic behaviors of the SWCNT.



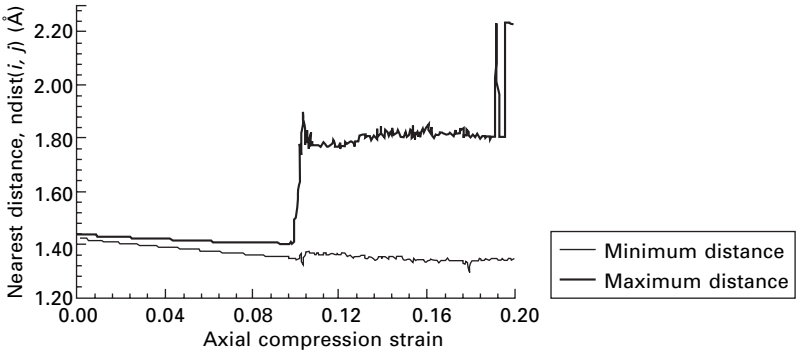
8.5 Strain energy using original Lennard-Jones for SWCNT (8,0) in axial compression.



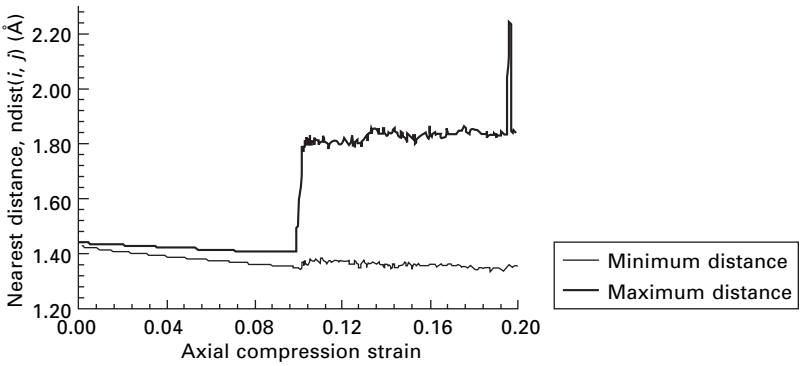
8.6 Strain energy using Lennard-Jones with switch function for SWCNT (8,0) in axial compression.

Upon observation of Figs 8.9 to 8.11, it is found that after the ultimate load point S, the ndist(i, j) tends to converge to two values, namely a minimum value of 1.3 Å and a maximum value of 1.75 Å. The collapse of the SWCNT occurs at the point where a sudden drop in strain energy is observed and at strain $\varepsilon = 0.235$. It is interesting to note that the variations of the maximum and minimum ndist(i, j) distances are linear for strain values lower than the ultimate load strain occurring at the point S in the stress–strain curve.

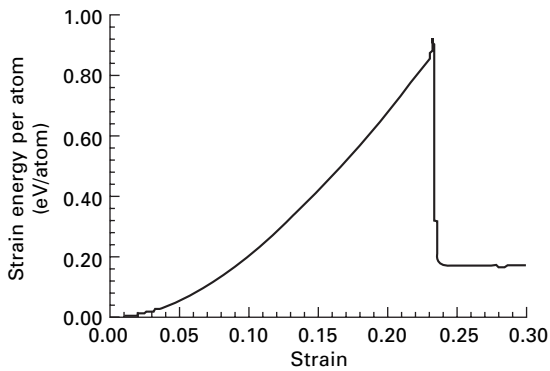
The straight line from the origin to the point P represents the linear elastic region. The points P and S are calculated using the least squares fitting of stress–strain data points. The fitting curve is based on $\sigma = A_P\varepsilon$ and $\sigma = A_S\varepsilon^2 + B_S\varepsilon$ respectively, where A_P, A_S, and B_S are the fitting coefficients.



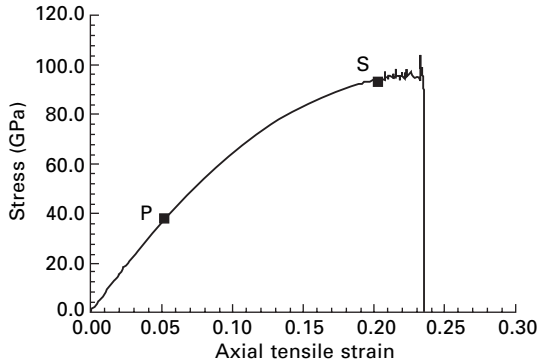
8.7 Nearest distance of SWCNT (8,0) in axial compression using original Lennard-Jones potential.



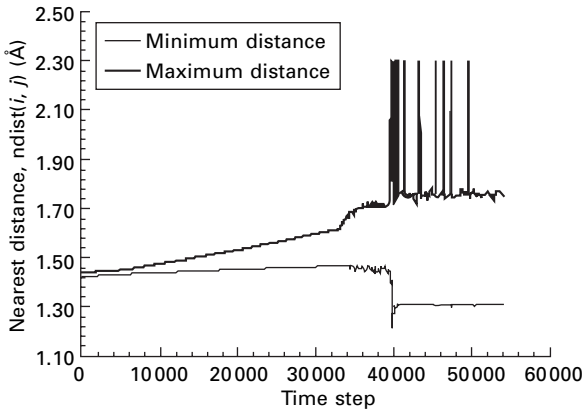
8.8 Nearest distance of SWCNT (8,0) in axial compression using Lennard-Jones with switch function.



8.9 Strain energy per atom using MD for SWCNT (7,7) under axial tension.



8.10 Stress–strain curve of SWCNT (7,7) under axial tension using MD.

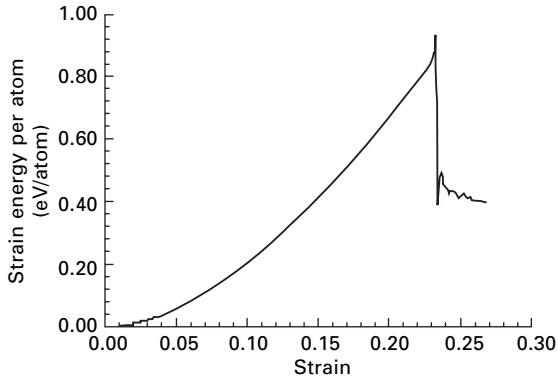


8.11 Nearest distance of SWCNT (7,7) under axial tension using MD.

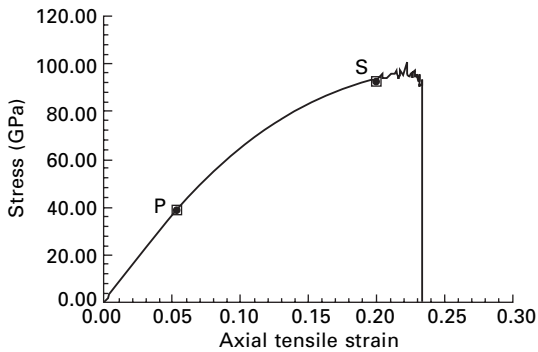
Using the same SWCNT under similar axial tension, we perform the corresponding multiscale model simulation. This multiscale model involves handshaking between MD and TB where 56 atoms are located in the far region and 112 atoms in the near region. Thus, the TB method is used to solve for 168 atoms, and this requires the solution of an eigensystem of dimension 672 (168×4).

For the multiscale modeling of the axial tension case, the TB near and far regions are implemented, after detecting from initial MD simulations, the maximum ndist(i, j) distance exceeding 1.71 Å. From the current comparison of Figs 8.9 to 8.13, it is observed that for both pure MD and multiscale simulations, the SWCNT collapses at strain $\varepsilon = 0.235$.

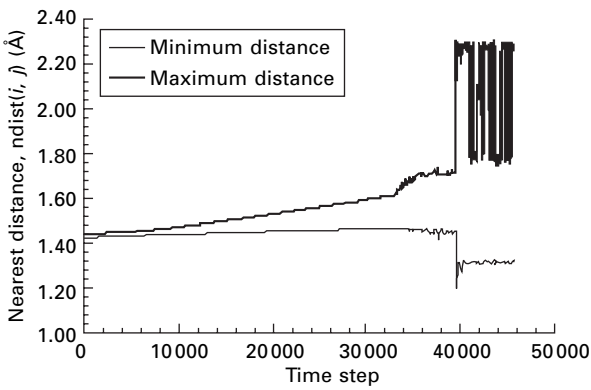
From Fig. 8.14, it is observed that after the ultimate load point S, the minimum and maximum ndist(i, j) distance converges to 1.31 Å and 1.78 Å respectively. The above results also show that for the multiscale method, the



8.12 Strain energy per atom using multiscale method for SWCNT (7,7) under axial tension.



8.13 Stress–strain curve of SWCNT (7,7) under axial tension using multiscale method.



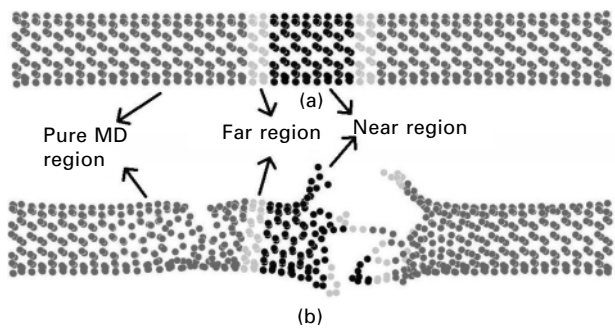
8.14 Nearest distance of SWCNT (7,7) under axial tension using multiscale method.

SWCNT collapses at higher strain energy per atom (0.935 eV/atom) compared with the pure MD case (0.918 eV/atom). Figure 8.15 depicts the collapse mode of this case study when implementing the multiscale model. As can be observed, the near and handshaking regions are applied only to the critical regions so as to significantly reduce the computational time.

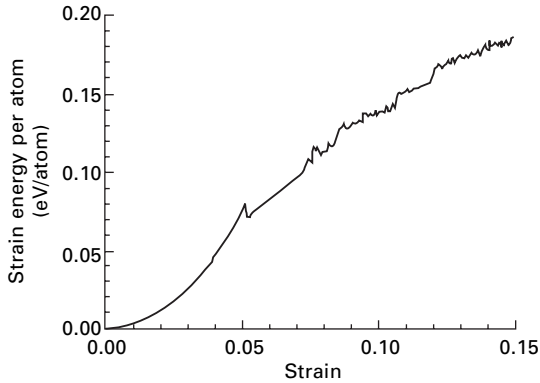
8.3.3 Carbon nanotubes subject to axial compression

We now move on to compare the numerical results of a SWCNT under axial compression using MD and multiscale TB/MD. In this case study, a SWCNT (7,7) with the length-to-diameter ratio $L/D = 7.4$ is used. The total number of carbon atoms is 812, with 644 atoms located in the MD region when multiscaling first occurs. For the first 800 fs of this simulation, we apply a constant velocity of 20 m/s to the left boundary, while the other boundary is subjected to a constant velocity of 20 m/s plus a slight linear velocity increase along the diameter. After 800 fs, all velocities are set to the constant value of 20 m/s. In the corresponding multiscale simulation, again we initially perform a full MD analysis. The full MD model will shift to multiscale model after detecting a *DOP* higher than 9. The present setting acts as the necessary trigger to induce the sideways buckling, and avoid the symmetrical modes.

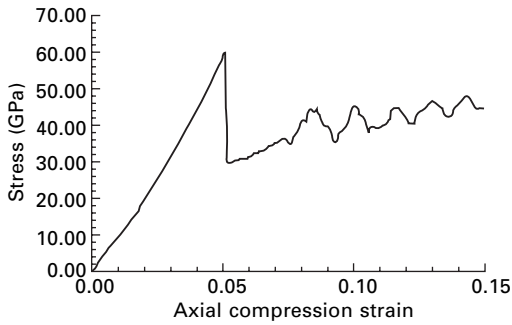
Basically, in the present work, the near and handshaking far TB regions are implemented after detecting a *DOP* exceeding 9. In this case study, this shift occurs at strain, $\varepsilon = 0.0518$. Figures 8.16 to 8.18 show the MD results for the variation strain energy per atom with strain, the stress–strain curve, and the maximum/minimum ndist(i, j) distances, respectively. The corresponding multiscale results are presented in Figs 8.19 to 8.21. We observe that the maximum ndist(i, j) distance behaves in a more stable manner for the multiscale model.



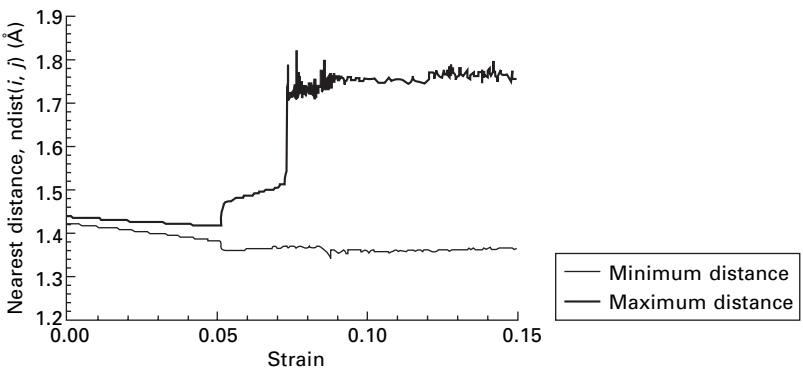
8.15 Multiscale model showing the pure MD, handshaking TB far, and TB near regions in SWCNT: (a) before collapse and (b) after collapse.



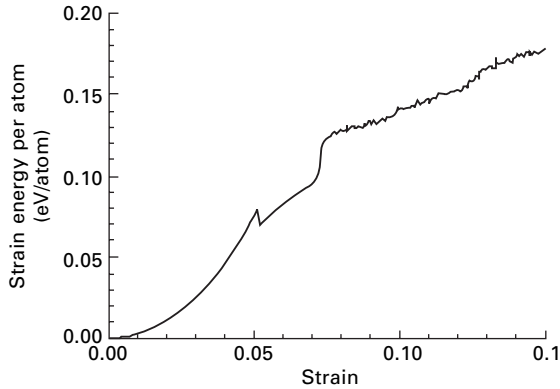
8.16 Strain energy per atom using MD for SWCNT (7,7) under axial compression.



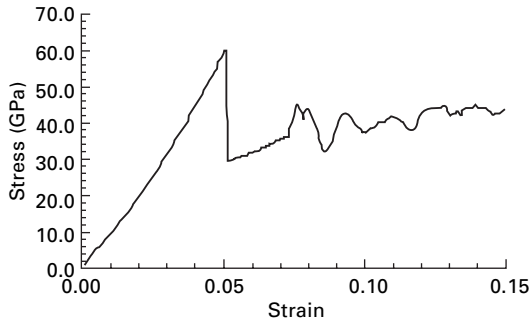
8.17 Stress–strain curve of SWCNT (7,7) under axial compression using MD.



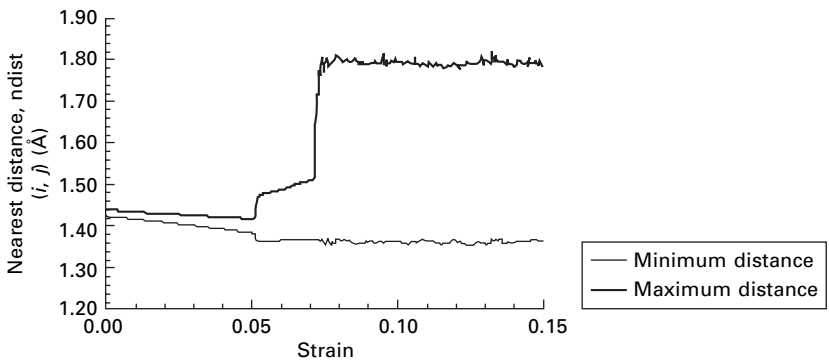
8.18 Nearest distance of SWCNT (7,7) under axial compression using MD.



8.19 Strain energy per atom using multiscale method for SWCNT (7,7) under axial compression.

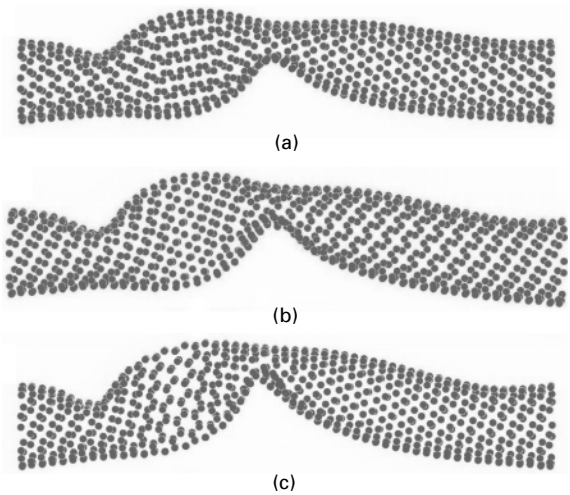


8.20 Stress–strain curve of SWCNT (7,7) under axial compression using multiscale method.

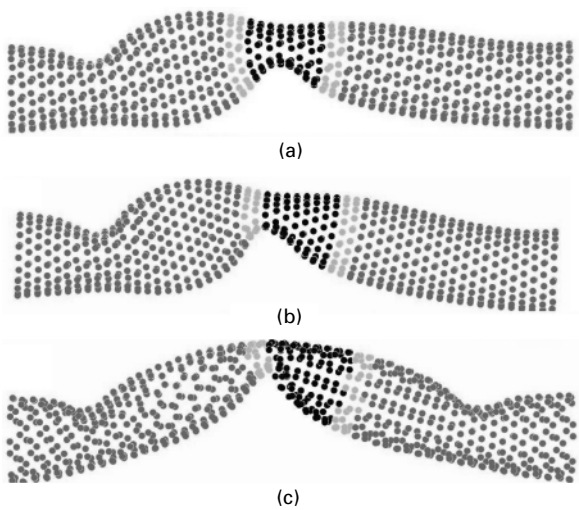


8.21 Nearest distance of SWCNT (7,7) under axial compression using multiscale method.

Figures 8.22 and 8.23 depict the sideways buckling modes when, respectively, implementing the MD and multiscale models. Comparing these two figures, we observe that both deformations are very similar at strain values below $\varepsilon = 0.072$. However, at strain values higher than $\varepsilon = 0.072$, the results from the multiscale model show more severe deformations.



8.22 MD results for sideways buckling at strain values of: (a) ε = 0.0569; (b) ε = 0.0683; and (c) ε = 0.078.



8.23 Multiscale results for sideways buckling at strain values of: (a) ε = 0.0569; (b) ε = 0.0683; and (c) ε = 0.078.

8.3.4 A note on minimization schemes

At present, we employ two minimization schemes, namely the conjugate-gradient method and the truncated-Newton via the Lanczos algorithm. Both schemes are applicable only to the smooth solution fields. Theoretically, the truncated-Newton scheme shows superior convergence properties if the starting coordinates are in the vicinity of the solution.

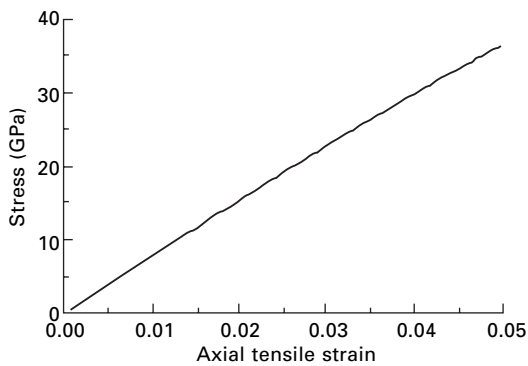
From the simulation results, we observe that both schemes produce quite similar results. However, in terms of computational time, we find that the truncated-Newton method is significantly more efficient.

As can be observed from Table 8.2, the truncated-Newton scheme is relatively more efficient over the entire simulation range. The conjugate-gradient method on the other hand, is very computationally intensive, especially in the bond breaking regime. It is also interesting to note that prior to reaching the yield strain, the conjugate-gradient method is somewhat faster. Finally, we present the smoothness of stress–strain curves when using both minimization schemes.

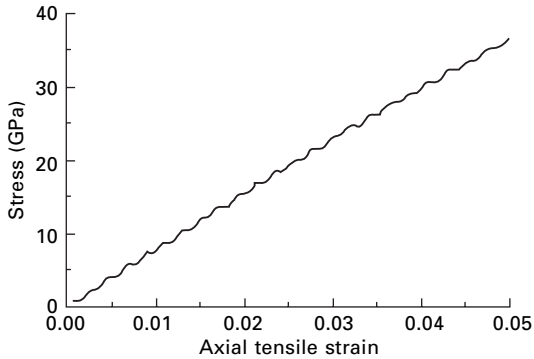
From Figs 8.24 and 8.25, we note that the conjugate-gradient method produces results which are somewhat oscillatory (see Fig. 8.25). Results from the truncated-Newton scheme are significantly smoother. We can thus

Table 8.2 Comparison of computational time

Strain	Time consumption (seconds, clock time)	
	Conjugate-gradient	Truncated-Newton[21]
0–0.118	398	463
0.118–0.236	5952	852
0.236–0.32	25 716	682



8.24 Smoothness of stress–strain curve when using truncated-Newton minimization.



8.25 Smoothness of stress–strain curve when using conjugate-gradient minimization.

conclude that the truncated-Newton via Lanczos minimization performs significantly better than the conjugate-gradient method, in terms of both the smoothness of the results and computational time.

8.4 Introduction to hydrogen interaction with carbon nanostructures

At present, it is well accepted that physisorption and chemisorption are the two distinct types of possible interaction between hydrogen and CNTs. Under most conditions, molecular physisorption is preferred, where the hydrogen molecules do not dissociate and the interactions between the CNT and the hydrogen molecules are usually described by weak van der Waals forces. This interaction is attributed to the exchange and correlation effects arising from the weak overlap between the electron densities of the nanotube and the closed-shell H_2 molecule[22].

Existing experimental data suggest that under certain conditions of energetic impact and/or high pressure, H_2 molecules can approach sufficiently close to the nanotube wall for possible dissociation, and atomic chemisorption thus becomes possible[23,24]. This also substantially weakens the C–C bonds[25]. Liu *et al.*[23] showed in their experiments that after treating CNTs with hydrogen gas under high pressure, there was residual H_2 during the desorption cycle that could be released only upon heating to temperatures above 400 K. They therefore suspected that these residuals may be related to chemical adsorption. A more recent experimental process by Ye *et al.*[24] using SWCNTs of high purity found a first-order phase transition under high H_2 pressure, similar to the hydride phase observed in metal-hydrogen systems.

For chemisorption mechanisms, several interesting schemes have been proposed recently. Nikolaev *et al.*[26] and Terrones *et al.*[27] put forth an H-

activated coalescence mechanism, in which the gas phase H atoms attack the side of neighboring nanotubes, breaking the C–C bonds and producing defective sites on adjacent nanotubes. This idea was generated based on the discovery that under atomic hydrogen atmosphere, single-walled armchair nanotubes annealed up to 1500 °C coalesce with neighboring tubes, resulting in larger nanotubes with twice and occasionally three times the diameter of the original ones. They suggested that once these adjacent defects are formed, the strong thermodynamic force, resulting from the released strain energy, drives the two neighboring smaller tubes to join together forming larger tubes[27]. Another interesting storage mechanism is that of a H atom flipping into the CNT, and this was proposed by Lee *et al.*[28]. However, they assumed that the chemisorption took place via a hypothetical electrochemical process in solutions, which are rather different from typical conditions reported for storage experiments.

A frequently used theoretical technique used by many researchers to investigate chemisorption parameters is to keep the SWCNT fixed while changing the distances and/or angles of the H atoms with reference to the SWCNT. For example, Arellano *et al.*[29] kept the SWCNT geometry fixed while employing a plan-wave basis set, obtaining binding energies of about –1 eV. Veronica *et al.*[30] kept the SWCNT structure fixed as well, and allowed the H atom to approach from the center of the nanotube (inner wall) to several exohedral positions in the radial direction (outer wall). However, this scheme is usually unable to provide very satisfactory results. This is because when reaction occurs, the nature of the two bonding atoms determines the chemical reaction, but they will be modified by the effects of the rest of the molecules as well. These effects cannot be reflected by this method.

Another method commonly seen in many reported theoretical works is the use of a finite cluster cut from the actual surface, with the distances between the atoms forming the cluster kept at their bulk values[31,32]. One fundamental problem for the cluster modeling of a surface is the convergence of the adsorption energy and other properties with respect to cluster size. Generally, a very large cluster is required. However, if more than a few molecules are explicitly described by quantum calculation, computational costs rapidly become prohibitive. The multiscale ONIOM method of Morokuma and coworkers[33–37] can improve this situation, allowing the division of a large molecular system into several layers and treating them using different theoretical methods. The advantage of using the ONIOM method is that it imposes the geometrical constraints of the tube during the geometry optimization. To date, this multiscale framework has been mostly applied to large organic molecules and organometallic complexes.

In this chapter, we calculate the energies of fully optimized hydrogenated nanotubes using two-layer ONIOM2 integrated schemes, which divide the entire system into two levels. The small section, essential for studying the properties of the chemisorption, is treated using quantum mechanical density-

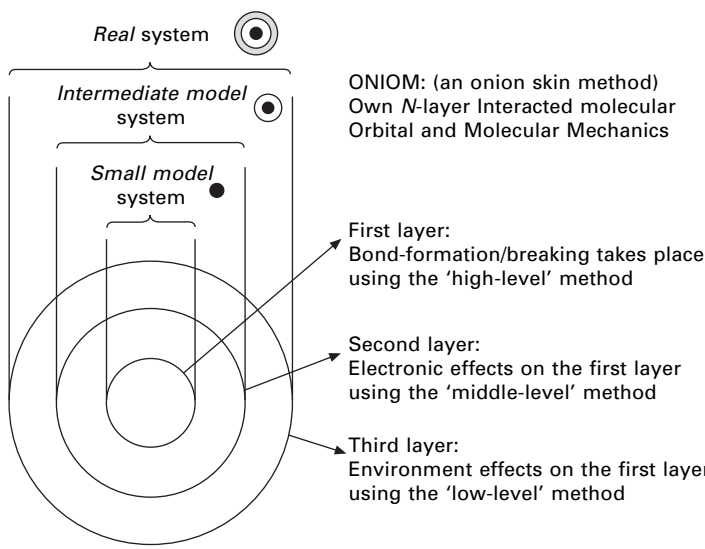
function theory (DFT). The remaining layer constrains the general geometry and is described by a molecular mechanics method, and the Universal Force Field (UFF) molecular mechanics model is chosen here. Based on this methodology, a systematic study of the interaction of H atoms with the sidewall of CNTs is carried out where the basic and integral features of the interaction are outlined.

8.5 Hybrid calculations with multiscale ONIOM scheme

8.5.1 Theoretical development

The multiscale (Own N-layer Interacted Molecular Orbital and Molecular Mechanics (ONIOM) scheme, a hybrid method developed by Morokuma and coworkers[33–35], allows the partitioning of a chemical system into layers, and different parts of the layers are treated by different computational levels, and combined to produce a consistent energy expression, see Fig. 8.26. The objective of this scheme is to perform a high-level calculation only on a small part of the system while considering the effects caused by the remainder at lower levels of theory, with the result being of similar accuracy to a high-level calculation on the full system, obtained at a reasonable computational cost.

However, it must be noted that the system is not randomly divided into different parts, but rather depends on the purpose each region serves and the



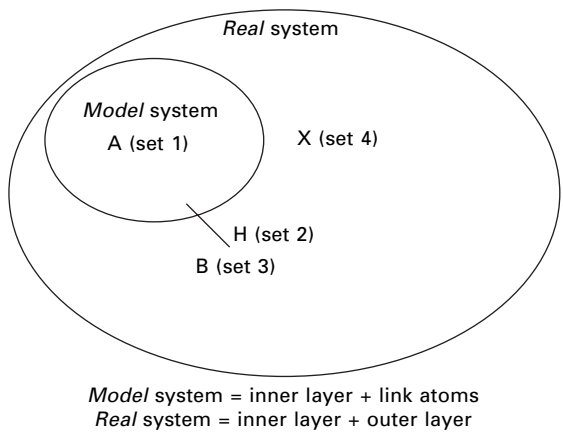
8.26 The onion skin-like layers and models of a three-layer partitioning scheme[36].

level of interest in particular regions. Due to the division of the system, link atoms like hydrogen and fluorine are often used to replace the broken bonds. A typical case is when one is interested in the accurate description of a particular region of a large organic molecule or a macromolecule, covalent bonds have to be cut in order to generate the inner model system. This leaves dangling bonds at the border of the inner layer, which have to be saturated in order to avoid a chemically unrealistic model. Hydrogen atoms are often used as the link atoms, and this is one of the critical features of treating the link atoms in the ONIOM method.

8.5.2 ONIOM link atoms

When the layers are not covalently bound, the model system is identical to the high-level layer. In that case, the ONIOM formulation does not require link atoms. However, when covalent bonds do exist, the resulting dangling bonds should be saturated with link atoms, which are chosen so that they best mimic the substituent. The link atoms for bonds spanning two regions are generated from the bond information, where the additional parameters on the model and real systems are derived to describe the broken bonds, including scale factors for placement of the link atom and the type of link atom. Usually the default link atom is hydrogen.

In Fig. 8.27, we provide more details on the treatment of the link atoms by taking the two-layer ONIOM2 model as an example. In this model, the atoms are divided into four groups of sets 1, 2, 3, and 4. The atoms present in both the model system and the real system are called set 1 atoms and their coordinates are donated by $\mathbf{R}_{\text{model}}$, for simplicity this is written as $\mathbf{R}_1$. The set



8.27 Schematic layout for the link atoms in a two-layer ONIOM2 scheme.

2 atoms are the artificially introduced link atoms, such as hydrogen atoms. They only occur in the model system and their coordinates are described by $\mathbf{R}_{\mathrm{LAH}}$, and simplified as $\mathbf{R}_2$, with 'LAH' denoting the link atom hosts, and they represent the atoms replaced by the link atoms in the model system. In the real system, they are replaced by the atoms described by $\mathbf{R}_{\mathrm{link-real}}$, denoted by $\mathbf{R}_3$. Atoms that belong to the outer layer and are not substituted by link atoms are called set 4 atoms with coordinates $\mathbf{R}_{\mathrm{real}}$, and simplified as $\mathbf{R}_4$. The geometry of the real system is thus described by $\mathbf{R}_1$, $\mathbf{R}_3$ and $\mathbf{R}_4$, and they are independent coordinates for calculating the ONIOM energy:

$$\begin{aligned} E_{\mathrm{ONIOM}} &= E_{\mathrm{ONIOM}}(\mathbf{R}_{\mathrm{model}}, \mathbf{R}_{\mathrm{link-real}}, \mathbf{R}_{\mathrm{real}}) \\ &= E_{\mathrm{ONIOM}}(\mathbf{R}_1, \mathbf{R}_3, \mathbf{R}_4) \end{aligned} \quad [8.25]$$

For the model system, it is described by $\mathbf{R}_{\mathrm{model}}$ and the link atom $\mathbf{R}_{\mathrm{LAH}}$, which is defined as a function of $\mathbf{R}_{\mathrm{model}}$ and $\mathbf{R}_{\mathrm{link-real}}$, and the explicit functional form of $\mathbf{R}_{\mathrm{LAH}}$ dependency can be written as:

$$\mathbf{R}_{\mathrm{LAH}} = f(\mathbf{R}_{\mathrm{model}}, \mathbf{R}_{\mathrm{link-real}}) \quad [8.26]$$

or

$$\mathbf{R}_2 = f(\mathbf{R}, \mathbf{R}_3) \quad [8.27]$$

Considering the fact that the link atoms are introduced to mimic the corresponding covalent bonds of the real system, they should follow the movement of the atoms they replace. A coupling scheme is introduced here to handle this. If atom A belongs to set 1 (model system) and atom B to set 3 (atoms in the real system but not substituted by the artificially introduced atoms), the link atom H in set 2 (artificially introduced link atoms in the model system) is placed onto the bond axis A–B. In terms of the internal coordinates, the same bond angles and dihedral angles are chosen for both the set 2 and set 3 atoms. Thus, link atoms are placed along the bond vector connecting the first to the second atom according to the equation

$$\mathbf{R}_{\mathrm{link}} = (1 - g)\mathbf{R}_1 + g \times \mathbf{R}_2 \quad [8.28]$$

or

$$\mathbf{R}_{\mathrm{link}} = (1 - g)\mathbf{R}_{\mathrm{model}} + g \times \mathbf{R}_{\mathrm{LAH}} \quad [8.29]$$

where g is a scale factor. If the scale factor is one, then the link atom is placed where the second atom was. More usually, the scale factor is less than one, in which case the link atom is placed between the original two atoms, but is always aligned along the bond vector. The scale factor should be chosen so that the link atom (usually hydrogen) is placed near its equilibrium bond length from the model atom. For example, when simulating the breaking of a single carbon–carbon single bond (typical length 1.528 Å for the sp^3 hybrid bond) using a hydrogen link atom, we will require a carbon-hydrogen

(link) bond length of about 1.084 Å, and the scale factor is therefore 1.084/1.528 ≈ 0.709[38].

In this section, we will use the two-layer ONIOM2 scheme to study the chemical reaction processes for hydrogen atoms and molecules on the exterior and interior of the CNT. The whole system is partitioned into the model and real systems described above, where the atoms involved in the chemical reaction are treated by a high-level calculation and the rest by a low-level method. Since the layers of high-level and low-level are covalently bonded, the resulting dangling bonds due to the partitioning of the system into two different layers will be saturated with link atoms, which are chosen so that they best mimic the substituent. As is the practice in the ONIOM method, the link atoms of hydrogen are used, and this yields good results when the C–C bond is broken. The C_{model}–H_{link} bonds are assigned the same angular and dihedral values as the C_{model}–C_{real} in the real system. The scaling factor is used for obtaining the bond length of C_{model}–H_{link}, which ensures that the number of degrees of freedom remains at 3N-6, so that any method for the investigation of potential energy surfaces available for conventional methods can be used for the ONIOM as well.

Theoretically, the ONIOM scheme allows a partition of the molecular system into two, three, or even more layers, and 'distributes' computational methods of various levels among these layers. The subdomain of greatest interest forms the innermost layer that is described by the highest level theory. Subsequent layers or subdomains are treated using progressively computationally cheaper lower-level approaches. The ONIOM method approximates the energy of a molecular structure subdivided into n layers as

$$\begin{aligned} E_{\text{ONIOM}}(n(\text{Level}(1)&: \text{Level}(2) : \ldots : \text{Level}\ (n))) \\ &= E(\text{Level}(1), \text{Layer}(1)) + \sum_{i=2}^{n} \left[E\left(\text{Level}\ (i), \bigcup_{j=2}^{n} \text{layer}(j) \right) \right] \\ &\quad - \sum_{i=2}^{n} \left[E\left(\text{Level}\ (i), \bigcup_{j=1}^{i-1} \text{Layer}\ (j) \right) \right] \end{aligned} \qquad [8.30]$$

where Level(i) is the level of theory used for layer i, starting from the highest Level(1) and ending with the lowest Level(n), while $\bigcup_{j=1}^{i} \text{Layer}\ (j)$ denotes the part of molecular structure consisting of layers 1 through i, where the entire subdomain is given by $\bigcup_{j=1}^{n} \text{Layer}\ (j)$. The geometries of the subsequent subdomains, Layer(1), Layer(1) ∪ Layer(2), Layer(1) ∪ Layer(2) ∪ Layer(3), etc., represent unmodified cutouts of the entire molecular structure with one exception, namely, bonds connecting atoms belonging to subsequent layers are saturated by adding link atoms.

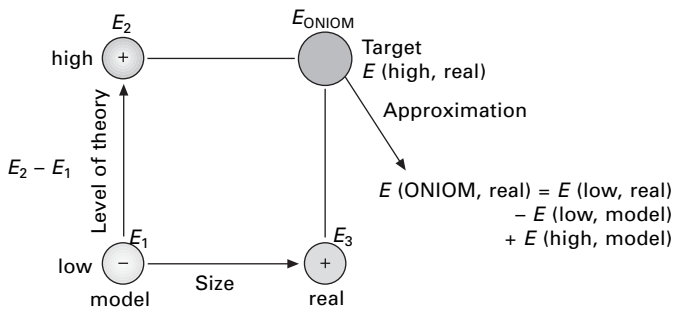
The idea of partitioning a molecular system into two or more parts or layers, where the interesting or difficult part of the system (the inner layer) is treated by a more refined level of theory and the rest of the system by a lower and therefore computationally less demanding method, is not new. Literature on hybrid QM/MM (quantum mechanical–molecular mechanics) methods is moderately extensive. However, a simple QM/MM simulation without a smoothing scheme suffers a surge in the total energy from time to time due to the solvent exchange, and such abrupt changes in the total energy occur during particle exchanges between the QM and MM subsystems. A smoothing scheme such as that in the ONIOM scheme is thus essentially due to the large differences in background potential energy calculated by different levels of theory.

8.5.3 ONIOM energy definition

In the two-layered ONIOM2 method, see Fig. 8.28, the total energy of the system is obtained from independent calculations

$$E_{\text{ONIOM2}} = E_3 + E_2 - E_1 \qquad [8.31]$$

where E_3 denotes energy calculated by the lower level method, while $(E_2 - E_1)$ denotes the part of the system in which the energy is calculated using both the high- and low-level methods. The energy of the real molecular system, calculated at the low level (E_3), is corrected using the energy difference ($\Delta E_{\text{high–low}} = E_2 - E_1$) between high ($E_2$) and low ($E_1$) levels of calculation on the inner layer, also called the 'model system', see Fig. 8.28. Such a definition in principle also holds for the first and second derivatives of the energy, and both geometrical optimization and frequency computation can be performed. Equation [8.31] can also be seen as a system effect of $(E_3 - E_1)$ being added to a high-level computation on the model system (E_2). Such a rationale of the equation led Morokuma and coworkers[39] to define a test to gauge the



8.28 The two-layer ONIOM method[36].

applicability of the coupling scheme where the 'S value test' was performed on $(E_3 - E_1)$.

Here we are dealing with a two-layer QM/MM ONIOM combination, for which the total energy of the system is obtained from three independent calculations

$$E^{\text{ONIOM}} = E^{\text{OM}}_{\text{model}} + E^{\text{MM}}_{\text{real}} - E^{\text{MM}}_{\text{model}} \qquad [8.32]$$

where 'real' denotes the full system, which only needs to be calculated at the MM level, and 'model' denotes the part of the system that needs to be calculated at both the QM and MM levels.

8.5.4 Geometry optimization

In the construction of the ONIOM model system, atoms that belong to the high-level layer have the same coordinates as the corresponding atoms in the real system. Even during geometry optimizations, these coordinates remain identical to one another. When no bond exists between the two layers, the first derivative of the energy with respect to the geometry coordinates can be obtained according to

$$\frac{\partial E^{\text{ONIOM}}}{\partial \mathbf{q}} = \frac{\partial E^{\text{high}}_{\text{model}}}{\partial \mathbf{q}} + \frac{\partial E^{\text{low}}_{\text{real}}}{\partial \mathbf{q}} - \frac{\partial E^{\text{low}}_{\text{model}}}{\partial \mathbf{q}} \qquad [8.33]$$

where $\mathbf{q}$ stands for the coordinates. The assumption for the link atoms, that they are connected to the high-level layer with the same angular and dihedral values as the link atom hosts (the atoms replaced by the link atoms in the model system), is also used in the geometry optimization. However, in most cases, there exist bonds between the low-level layer atoms and those of the high-level layer, and simply using equation [8.33] for geometry optimization in this case will lead to the loss of one degree-of-freedom for each link between the high- and low-level layers. This causes problems with dynamic or frequency calculations. Thus, bond distances between the high-level layer and the link atoms have to be obtained by scaling the corresponding distances between the high-level layer and the LAH atoms using equation [8.28] or [8.29]. The geometrical derivatives of the ONIOM energy can now be obtained in a similar fashion to the energy, where the correct number of degrees-of-freedom ensures that the potential energy surface is properly defined, and all gradients and higher derivatives are available[38]. In this case, the link atoms are present, and the Jacobian $\mathbf{J}$ must be used to convert the coordinate system for the model system to that for the real system. The gradient can thus be written as

$$\frac{\partial E^{\text{ONIOM2}}}{\partial \mathbf{q}} = \frac{\partial E^{\text{high}}_{\text{model}}}{\partial \mathbf{q}} \cdot \mathbf{J} + \frac{\partial E^{\text{low}}_{\text{real}}}{\partial \mathbf{q}} - \frac{\partial E^{\text{low}}_{\text{model}}}{\partial \mathbf{q}} \cdot \mathbf{J} \qquad [8.34]$$

or in simpler representation

$$\nabla E^{\mathrm{ONIOM2}} = \nabla E^{\mathrm{high}}_{\mathrm{model}} \cdot \mathbf{J} + \nabla E^{\mathrm{low}}_{\mathrm{real}} - \nabla E^{\mathrm{low}}_{\mathrm{model}} \cdot \mathbf{J} \quad [8.35]$$

where $\mathbf{J}$ is the Jacobian, which projects the forces on all the set 2 link $\mathbf{R}_{\mathrm{LAH}}$ atoms onto the set 1 $\mathbf{R}_{\mathrm{model}}$ and set 3 $\mathbf{R}_{\mathrm{link-real}}$ atoms. Adopting equation [8.28] or [8.29], the differentiation of the vector components of one set 2 atom with respect to the components of $\mathbf{R}_{\mathrm{model}}$ and $\mathbf{R}_{\mathrm{link-real}}$ yields a simple Jacobian

$$\frac{\partial \mathbf{R}_{\mathrm{LAH},a}}{\partial \mathbf{R}_{\mathrm{link-real},b}} = g\delta_{a,b} \quad [8.36]$$

$$\frac{\partial \mathbf{R}_{\mathrm{LAH},a}}{\partial \mathbf{R}_{\mathrm{real},b}} = (1 - g)\delta_{a,b} \quad [8.37]$$

where the indices a and b denote the Cartesian components x, y and z, and δ is the Kronecker delta.

We apply equation [8.34] to the hydrogen storage problem in Section 8.6, with QM methodology for the model system and an MM method for the real system. Thus equation [8.34] can be written here as

$$\frac{\partial E^{\mathrm{ONIOM}}}{\partial \mathbf{q}} = \frac{\partial E^{\mathrm{QM}}_{\mathrm{model}}}{\partial \mathbf{q}} \cdot \mathbf{J} + \frac{\partial E^{\mathrm{MM}}_{\mathrm{real}}}{\partial \mathbf{q}} - \frac{\partial E^{\mathrm{MM}}_{\mathrm{model}}}{\partial \mathbf{q}} \cdot \mathbf{J} \quad [8.38]$$

From equation [8.32] it follows that the interactions between the QM and MM layers are included at the MM level, via the $E^{\mathrm{MM}}_{\mathrm{real}}$ term. This type of embedding is referred to as mechanical embedding. In fact, in the ONIOM scheme, the interaction between any two layers is always included at the lower of the two associated levels.

The higher-order derivatives can be uniquely defined in a similar fashion and will be discussed in the next section. Any method for the investigation of potential energy surfaces based on conventional techniques can therefore be applied with the ONIOM method.

8.5.5 Second derivatives of ONIOM

For the frequency problem calculation, the second derivatives of the ONIOM total energy E^{ONIOM} with respect to the nuclear coordinates, the Hessian matrix $\boldsymbol{H}^{\mathrm{ONIOM}}$, which is $\nabla^2 E^{\mathrm{ONIOM}}$, can be obtained. The force constant matrix of the model system at low level $\boldsymbol{H}^{\mathrm{low}}$ and at high level $\boldsymbol{H}^{\mathrm{high}}$ have to be transformed by applying the Jacobian $\mathbf{J}$ and its transpose $\mathbf{J}^{\mathrm{T}}$, according to

$$\begin{aligned} \boldsymbol{H}^{\mathrm{ONIOM2}} &= \nabla^2 E^{\mathrm{ONIOM2}} \\ &= \mathbf{J}^{\mathrm{T}} \cdot \nabla^2 E^{\mathrm{high}}_{\mathrm{model}} \cdot \mathbf{J} + \nabla^2 E^{\mathrm{low}}_{\mathrm{real}} - \mathbf{J}^{\mathrm{T}} \cdot \nabla^2 E^{\mathrm{low}}_{\mathrm{model}} \cdot \mathbf{J} \end{aligned} \quad [8.39]$$

Using the same method, the gradient and second derivative expressions can be derived for three- or *n*-layer ONIOM partitioning.

In most computations, it is common practice to include an empirical scale factor (usually ~ 0.89) to the frequencies, which leads to good agreement with experimental data for a wide range of systems[40]. The reason is that the normal vibrational frequencies computed with molecular methods are well known to be over-estimated due to the incomplete consideration of electron correlation and the harmonic approximation. In the ONIOM scheme, different methods are used for different parts of a system, and different scale factors should therefore be used for the Hessians at different levels. As the ONIOM combines the Hessian matrices at different theoretical levels, the scaling process has to be performed for the Hessian matrix at each level

$$\begin{aligned} \boldsymbol{H}^{\mathrm{ONIOM2}} &= \nabla^2 E^{\mathrm{ONIOM2}} \\ &= c_2^2 \cdot \mathbf{J}^{\mathrm{T}} \cdot \nabla^2 E^{\mathrm{high}}_{\mathrm{model}} \cdot \mathbf{J} + c_3^2 \cdot \nabla^2 E^{\mathrm{low}}_{\mathrm{real}} - c_1^2 \cdot \mathbf{J}^{\mathrm{T}} \cdot \nabla^2 E^{\mathrm{low}}_{\mathrm{model}} \cdot \mathbf{J} \end{aligned} \quad [8.40]$$

where c_1, c_2, and c_3 represent the scale factors for different sets, see Figure 8.27. Since the frequency corresponds to the square-root of the Hessian, the square of the frequency scale factor has to be used for Hessian scaling.

Other derivatives such as the density, dipole moment, and polarization tensor can be defined as follows:

For the two-layered ONIOM, the density can be obtained as

$$\rho_{\mathrm{ONIOM2}} = \rho^{\mathrm{low}}_{\mathrm{real}} - \rho^{\mathrm{low}}_{\mathrm{model}} + \rho^{\mathrm{high}}_{\mathrm{model}} \quad [8.41]$$

When calculating the nuclear magnetic resonance (NMR) chemical shifts, the elements of a 3×3 shielding tensor for a nucleus can be written as[37]

$$\sigma_{ab} = \frac{\partial^2 E}{\partial B_b \partial \mu_a} \quad [8.42]$$

where μ and B are the nuclear magnetic moment and external magnetic field, respectively, and the indices a and b denote the Cartesian components x, y, and z. When isotropic shielding is required, the integrated nuclear shielding constant can be calculated with an expression analogous to the ONIOM energy expression

$$\sigma^{\mathrm{ONIOM2}}_{\mathrm{iso}} = \sigma^{\mathrm{low,real}}_{\mathrm{iso}} - \sigma^{\mathrm{low,model}}_{\mathrm{iso}} + \sigma^{\mathrm{high,model}}_{\mathrm{iso}} \quad [8.43]$$

and the ONIOM2 dipole moment, which is related to the electronic field $\mathbf{F}$[41], is defined as

$$\mu = \frac{\partial E_{\mathrm{ONIOM2}}}{\partial \mathbf{F}} = \frac{\partial E^{\mathrm{low}}_{\mathrm{real}}}{\partial \mathbf{F}} - \frac{\partial E^{\mathrm{low}}_{\mathrm{model}}}{\partial \mathbf{F}} + \frac{\partial E^{\mathrm{high}}_{\mathrm{model}}}{\partial \mathbf{F}} \quad [8.44]$$

The polarization tensor for ONIOM2 is

$$\alpha = \frac{\partial^2 E_{\text{ONIOM2}}}{\partial \mathbf{F}_a \partial \mathbf{F}_b} = \frac{\partial^2 E^{\text{low}}_{\text{real}}}{\partial \mathbf{F}_a \partial \mathbf{F}_b} - \frac{\partial^2 E^{\text{low}}_{\text{model}}}{\partial \mathbf{F}_a \partial \mathbf{F}_b} + \frac{\partial^2 E^{\text{high}}_{\text{model}}}{\partial \mathbf{F}_a \partial \mathbf{F}_b} \quad [8.45]$$

The third-order hyperpolarizability tensor for ONIOM2 is

$$\beta = \frac{\partial^3 E_{\text{ONIOM2}}}{\partial \mathbf{F}_a \partial \mathbf{F}_b \partial \mathbf{F}_c} = \frac{\partial^3 E^{\text{low}}_{\text{real}}}{\partial \mathbf{F}_a \partial \mathbf{F}_b \partial \mathbf{F}_c} - \frac{\partial^3 E^{\text{low}}_{\text{model}}}{\partial \mathbf{F}_a \partial \mathbf{F}_b \partial \mathbf{F}_c} + \frac{\partial^3 E^{\text{high}}_{\text{model}}}{\partial \mathbf{F}_a \partial \mathbf{F}_b \partial \mathbf{F}_c} \quad [8.46]$$

while the infrared intensity is

$$I_{\text{IR}} = \frac{\partial^2 E^{\text{low}}_{\text{real}}}{\partial \mathbf{R} \partial \mathbf{F}} - \frac{\partial^3 E^{\text{low}}_{\text{model}}}{\partial \mathbf{R} \partial \mathbf{F}} \cdot \mathbf{J} + \frac{\partial^2 E^{\text{high}}_{\text{model}}}{\partial \mathbf{R} \partial \mathbf{F}} \cdot \mathbf{J} \quad [8.47]$$

and the Raman intensity is

$$I_{\text{Raman}} = \frac{\partial^3 E^{\text{low}}_{\text{real}}}{\partial \mathbf{R} \partial \mathbf{F}_a \partial \mathbf{F}_b} - \frac{\partial^3 E^{\text{low}}_{\text{model}}}{\partial \mathbf{R} \partial \mathbf{F}_a \partial \mathbf{F}_b} \cdot \mathbf{J} + \frac{\partial^3 E^{\text{high}}_{\text{model}}}{\partial \mathbf{R} \partial \mathbf{F}_a \partial \mathbf{F}_b} \cdot \mathbf{J} \quad [8.48]$$

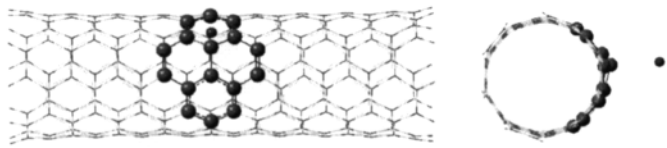
8.6 Chemosorption of hydrogen atoms onto carbon nanotubes

Although CNTs are very stable structures, hydrogen may under certain circumstances still penetrate the CNT wall. Atomic state hydrogen also penetrates more easily than hydrogen molecules[42]. It is interesting to note that there is no energy barrier for hydrogen atoms which have been chemisorbed at the outside the tube wall, regardless of initial placement. Here, the optimized structure, activation energy of the reaction, and the transition states (TS) are calculated using the ONIOM2 scheme. The ONIOM2 simulation shows that three TS are required for the entire reaction. The transition state species is verified by frequency analysis. The small model system is treated by DFT with the hybrid functional B3LYP (Becke's three-parameter hybrid method[43–48] with the exchange functional of Lee *et al.*[49]) in conjunction with the 6.31G basis set. The UFF[50] molecular mechanics is used as the low-level treatment.

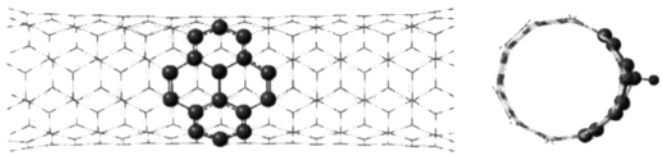
8.6.1 Initial weakening of C–C bond and chemisorption of one H atom onto a CNT

We will first examine the effects of hydrogen atoms on the geometry structures and the energy change of the CNT when the H atom approaches the sidewall of the CNT. Calculations for the adsorption of an H atom by a (5,5) SWCNT and the reaction path scheme of the H atom to the tube are carried out. Following optimization of a free (5,5) SWCNT structure, giving an average

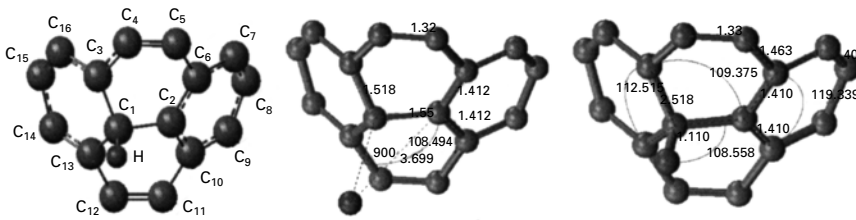
C–C bond length of 1.42 Å, a hydrogen atom is placed at different sites exterior to the nanotube, and the geometry optimization and energy calculation of the system is carried out using the ONIOM2 method. As shown in Fig. 8.29(b), after geometry optimization, the H atom was adsorbed at the outer surface of the tube, and the preferred site for the adsorption is above a carbon atom (perpendicular to the wall surface). This observation is in accordance with those reported by Arellano *et al.*[25,29], Bauschlicher *et al.*[51,52], Lee and Lee[53], and Lee *et al.*[54]. In the simulation, the atoms which are directly involved in bond breaking and formation are highlighted and labeled in the manner shown in Fig. 8.29(c). An inspection of the TS structure of Fig. 8.29(a, d) reveals that the distance between the H atom and C_1 is 2.90 Å and the angle $\angle HC_1C_2$ is 108.49°. The distance between neighbouring atoms C_1 and C_2 increases from 1.42 Å for the pure nanotube to 1.553 Å. For the relaxed system shown in Fig. 8.29(b, e), the calculated H–C_1 bond length is 1.11 Å, which is close to the value of 1.10 Å for CH_4 molecules. The chemisorbed H atom results in an increase of the bond length of its neighbour atoms, where the C_1–C_2 carbon bond is further elongated to 1.555 Å which is close to the typical sp^3 C–C bond length of 1.545 Å, while that of the C_1–C_3 and C_1–C_{13} bonds increases to 1.52 Å. The related overlap population has



(a) Transition state I



(b) Optimized structure of the stable intermediate I



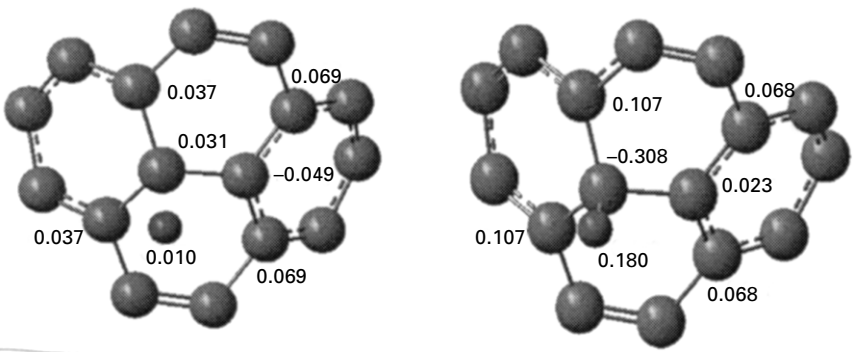
(c) Label of the atoms (d) Transition state parameters (e) Relaxed system parameters

8.29 Sideview along the SWCNT axis of the optimized structure of transition state and the optimized structures for the stable intermediate. The fragments represent the small model system (interatomic distances in Å, bond angles in degrees).

changed as well. From the small model of the system, we note that the overlap population between C_1–C_2 decreases from 0.40 to 0.24, and for the other two neighboring sets of carbon atoms C_1–C_3 and C_1–C_{13}, it has dropped from 0.41 to 0.30, indicating that the chemisorbed H atom weakens the carbon bonds in the vicinity of the chemisorbed site.

The H–C_1 overlap population is found to be 0.35, which is higher than those of the nearest carbon neighbors C_1–C_2 which is around 0.3. Nevertheless the mutual distances for C_2–C_6 and C_2–C_{10} remain almost the same as in a pure CNT at about 1.41 Å. The corresponding distances for C_4–C_5 and C_{11}–C_{12}, on the other hand, diminish to 1.38 Å, while the overlap population between them increased significantly to 0.53. Bond angles $\angle C_3C_1C_2$ and $\angle C_{13}C_1C_2$ are reduced from 120° to 109.4°, which are in good agreement with the typical bond angles found in diamond of 109.47°. For atoms relatively distant from the H atom, their angles remain as 120°, which is the typical bond angle for pure CNTs. Figure 8.30 shows the associated Mulliken charges for hydrogen and carbon atoms at the adsorption site.

The interaction energy of the H atom with the exterior wall of the (5,5) SWCNT is summarized in Table 8.3, where the reaction energy includes the



8.30 Mulliken charges for hydrogen and carbon atoms at the adsorption sites.

Table 8.3 Relative energy of reactants, TS, and reaction products for the reaction between H atom and SWNT, calculated with ONIOM2(B3lYP/6-31G:UFF) approach

System	Method	SCF energy (Hartree)
Reactant: H atom	B3lYP/6-31G:UFF	–0.5003
Reactant: CNT	B3lYP/6-31G:UFF	–615.5241
Product: CNT-H	B3lYP/6-31G:UFF	–616.0784
CNT-H transition state (TS)	B3lYP/6-31G:UFF	–615.9874
Forward reaction activation energy (eV)		1.01
Reverse reaction activation energy (eV)		2.32
Energy of reaction (eV)		–1.31

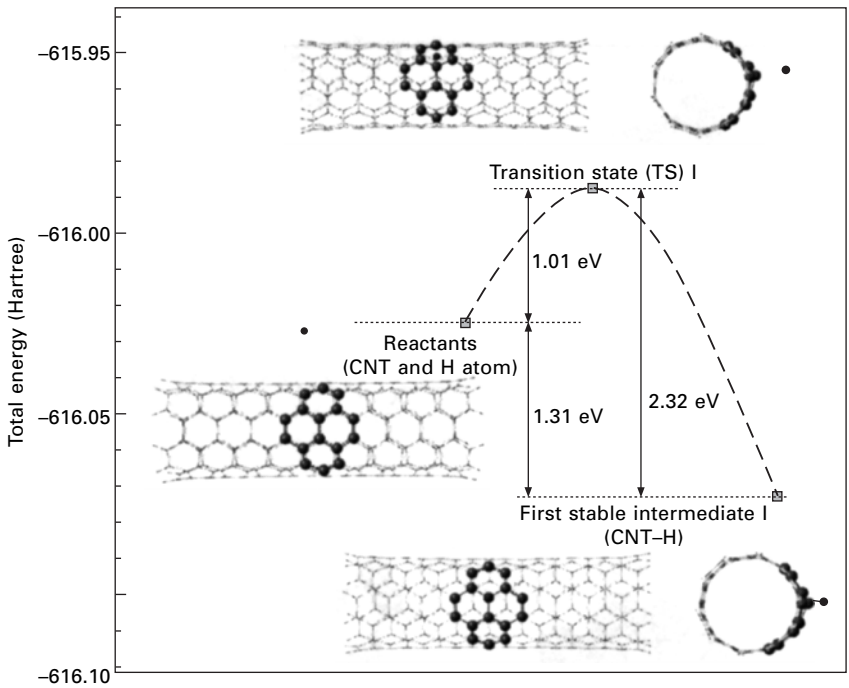
calculated energy for the H and CNT reactants, TS, and the products of H–CNT. For the single H atom, the chemisorption adsorption energy is calculated as the difference between the energy of the H–CNT and the isolated energies of the H atom and pure CNT[30]

$$E_{\text{chemisorption}} = E(CNT - H) - E(CNT) - E(H) \quad [8.49]$$

where $E(CNT - H)$ denotes the energy of the CNT and the H atom, $E(CNT)$ the energy of the bare CNT, and $E(H)$ the energy of the single H atom. According to this definition, a stable system is supposed to have a negative binding energy. The corresponding chemisorption energy for the reactants, TS structure, and the stable intermediate I are plotted in Fig. 8.31.

From the calculation, it is found that the adsorption process on the sidewall of the tube is exothermic, with the H–C bonding energy being –1.31 eV as indicated in Fig. 8.31. Thus, it is evident that the reactants achieve an energetically favorable state during the chemical reaction process.

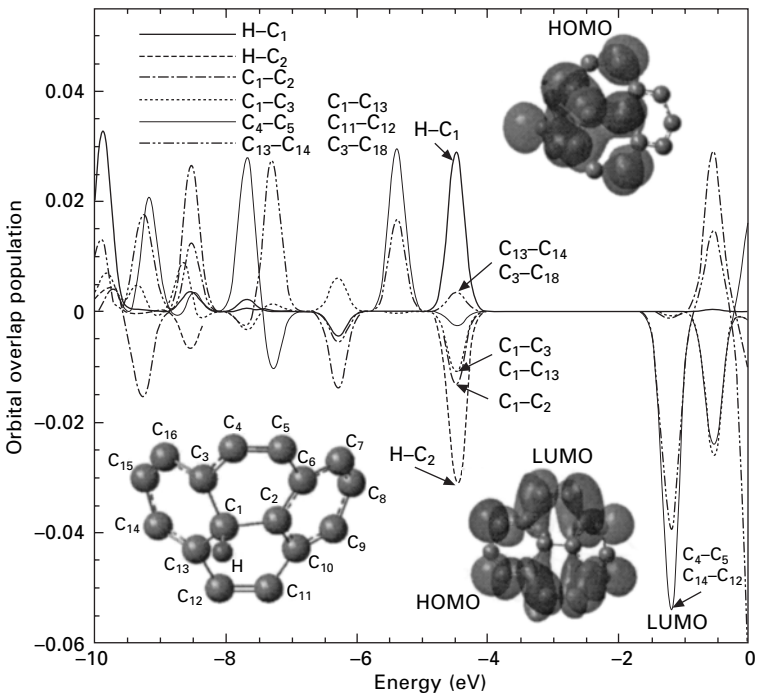
To obtain more insight into the change of electronic structure associated with this H chemisorption process, the orbital overlap population[55] is calculated for the stable intermediate state I. The orbital overlap population curve indicates that the H–C_1 corresponding to the highest occupied molecular orbital (HOMO)



8.31 Schematic pathway for the chemisorption of one H atom on a (5,5) SWCNT.

are bonding, while those of H–C_2 and C_1–C_2 are antibonding. On the other hand, the bonds of H–C_1, H–C_2, and C_1–C_2 at the lowest unoccupied molecular orbital (LUMO) energy level are non-bonding, while those of C_4–C_5, C_{11}–C_{12} exhibit strong antibonding. These observations are presented in Fig. 8.32. In addition, for one H atom chemisorbed, we obtained E_{HOMO} as –4.45 eV, and E_{LUMO} as –1.20 eV, and E_{gap} is thus 3.25 eV.

It can be seen that when one H atom approaches the sidewall of a pure CNT, it will eventually be chemisorbed on the tube, resulting in the weakening of the sp^2 bonds between its two nearest carbon atoms, C_1 and C_2. In addition, there is large charge transfer occurring from the H atom to C_1, filling up the p valence orbital of C_1. In Fig. 8.32, the HOMO depicts C_1 exhibiting typical sp^3 hybrid bonds with the nearby carbon atoms, and a covalent σ bond forms between H and C_1. During the formation of the H–C bond, the Mulliken analysis of bond populations, see Fig. 8.30, shows 0.339 electrons transferred from the neighbouring atoms to C_1. It is also noted that 0.17 electrons are from the H atom, after the H atom has been adsorbed onto the sidewall of the



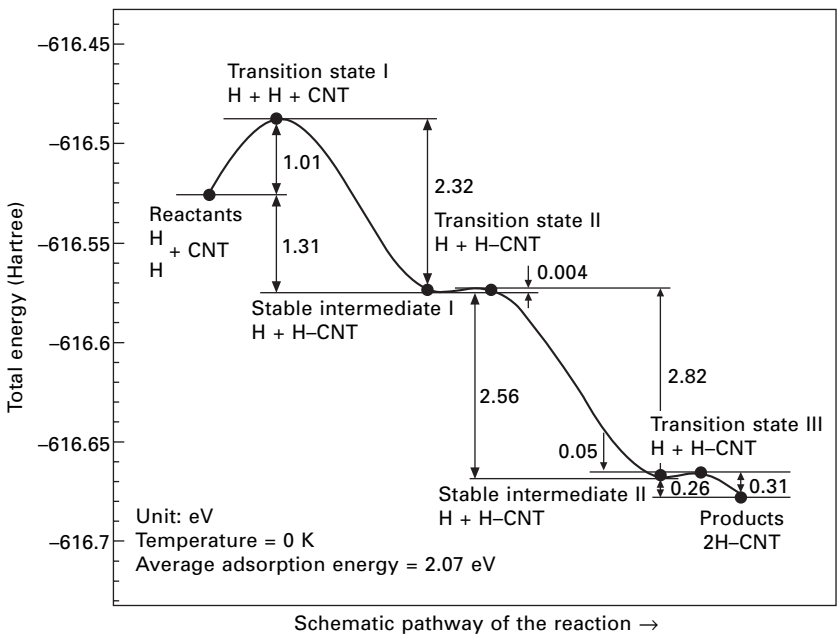
8.32 Orbital overlap population contributed from H, C_1, C_2, C_3, C_4, C_5 and C_{13} which are involved in the chemisorption process for the intermediate state I. Inset: LUMO and HOMO of the small model system.

CNT, and this is consistent with earlier reported calculations for fully hydrogenated nanotubes[56].

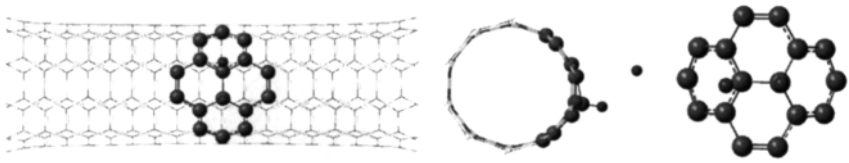
8.6.2 ONIOM calculations of chemisorption of two hydrogen atoms onto a CNT

Here we examine the reaction of a second hydrogen atom chemisorbed on the sidewall of the (5,5) SWCNT. Once again, we carry out simulations based on the ONIOM2 method for the case of two hydrogen atoms being chemisorbed, and these simulations start with the stable intermediate structure I that was obtained in the preceding computation. The second hydrogen atom is placed close to the tube wall, which can be physically achieved under high-pressure conditions, or by injecting hydrogen atoms with high kinetic energies into the reaction cell.

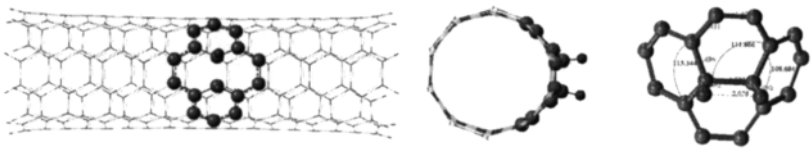
Similar to the results in the previous calculation for the first H atom, the results here show that the adsorption of the second H atom on the CNT wall results in a distortion of the CNT structure as well. The energy parameter variation is shown in Fig. 8.33 while the corresponding configurations are shown in Fig. 8.34. The distance between C_1 and C_2 is 1.73 Å for stable intermediate state II and elongates to 2.50 Å in the final products. Another



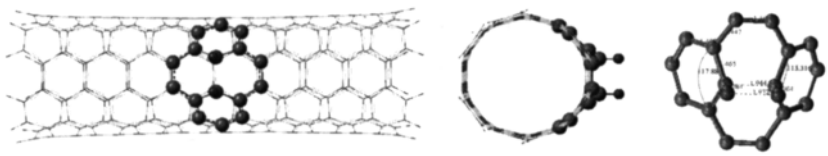
8.33 Schematic pathway for the chemisorption of two H atoms on a (5,5) SWCNT. Three transition states are found during the reaction.



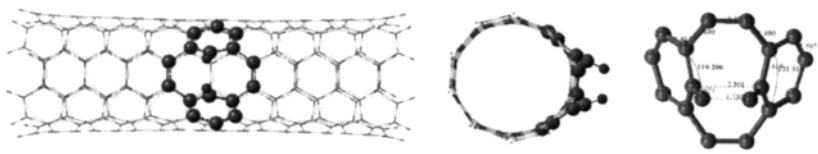
(a) Transition state II



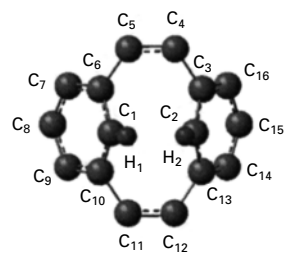
(b) Stable intermediate II



(c) Transition state III



(d) Optimized structure of products



(e) Labels of the atoms

8.34 Optimized structure for the transition states and the final products.

important parameter is the distance between the two hydrogen atoms, which generally decreases in the manner 2.78 Å → 2.08 Å → 1.97 Å → 1.74 Å from TS II to the final product of 2H–CNT.

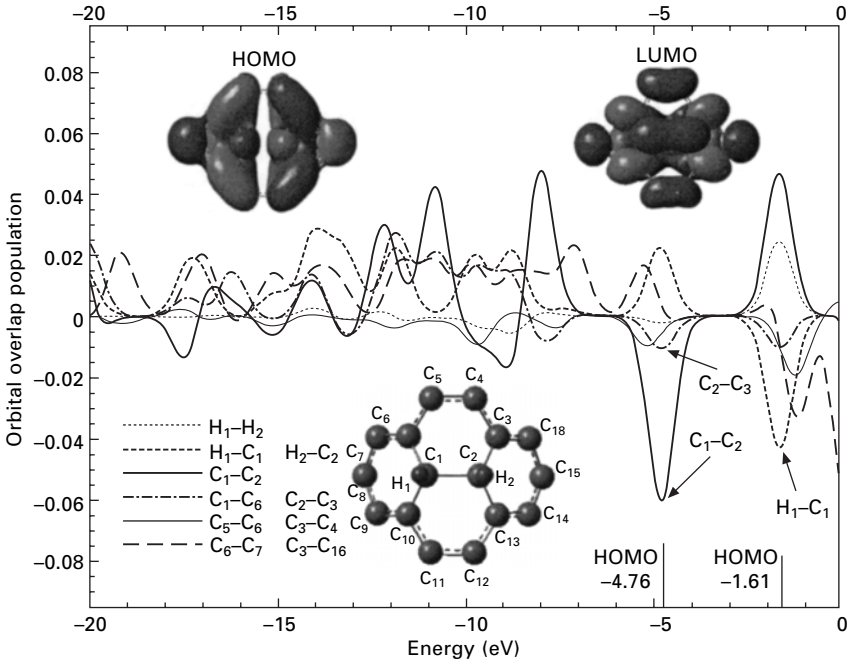
From the simulation results, it is noted that the second H atom prefers to be chemisorbed with C_2, which is located at the same layer as C_1. The chemical adsorption energy for the second H–C bond is about –2.82 eV, which is a dramatic increase in magnitude compared to the corresponding value for the single H case of –1.31 eV. The average binding energy for

forming the C–H bonds is about –2.07 eV for the entire process. This energy value has also been previously estimated by Gülseren *et al.*[57] and Han *et al.*[58]. The present results indicate that the adsorption of H atoms on the outside wall of the (5,5) SWCNT is preferred with carbon atoms which are perpendicular to the tube axis, and this adsorption is exothermic and stable. These conclusions are consistent with the observations of Gülseren *et al.*[57], Han and Lee[58], and Yang and Yang[32].

From Fig. 8.5, we note that the reaction will follow the process of the H atoms first approaching sufficiently near the surface of the CNT, which could happen under high pressure or high kinetic energy injection, and substantially weakening the C–C π bond of the adjacent tube surface while forming the first H–C σ bond (stable intermediate I). Next the weakening of the C–C bond results in an unpaired electron in the π orbital of the carbon atom, thus initiating another H atom to form the second H–C σ bond with the carbon atom (stable intermediate II). This second C–H bond further weakens the C–C bond and elongates the bond length. The attraction between the hydrogen atoms reduces the mutual H–H distance and increases the C–C bond until it eventually breaks and the 2H–CNT products are formed.

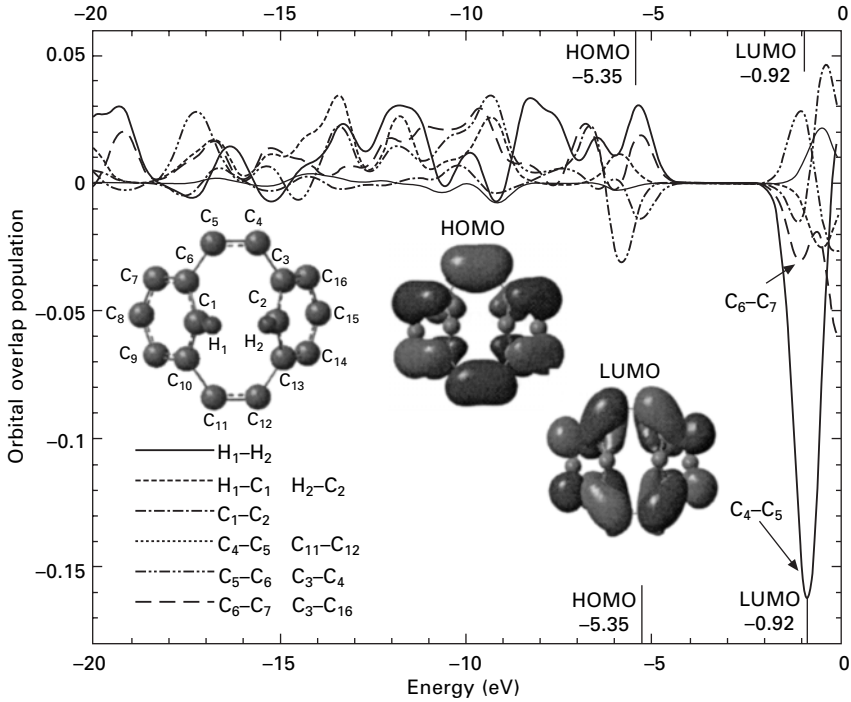
It is obvious from the results that the adsorption energy for forming the C–H bonds is different in the three stages. For the first H atom, it is found that 1.31 eV is released during the adsorption process, while for the second H atom forming the C–H bond with the adjacent C–C bond still intact, 2.56 eV of energy is emitted. When the adjacent C–C bond is broken, 0.26 eV is further released. It turns out that the energy values for the two stable intermediates and the final products are lower than the energies of the reactants, indicating that during the chemisorption, the hydrogen ions will be exothermically adsorbed onto the sidewall of the CNT.

The orbital overlap population plots and the HOMO and LUMO orbitals for the small model clusters of the stable intermediate state II are studied and shown in Fig. 8.35. The energies of the HOMO and LUMO are –4.76 eV and –1.61 eV respectively. The HOMO–LUMO gap in the stable intermediate state II has dropped to 3.15 eV. The orbital overlap population curve for the C_1–C_2 bond shows very strong antibonding at the HOMO level and strong bonding at the LUMO level, which is consistent with the results of the surfaces obtained for the HOMO and LUMO orbitals. The properties of the C_2–C_3, C_1–C_6, C_1–C_{10}, and C_2–C_{13} bonds at the HOMO and LUMO level are qualitatively similar to that of the C_1–C_2 bond, but showing relatively weak antibonding at HOMO level and weak bonding at the LUMO level. The situation for the H_1–C_1 and H_2–C_2 and H_1–H_2 bonds is the converse of that observed for the above C–C bonds. Bonds between hydrogen and carbon indicate bonding at HOMO, and non-constructive overlap between the two H atoms gives rise to non-bonding at the HOMO level, but shows antibonding at the LUMO level.



8.35 Orbital overlap population of the small model system for stable intermediate state II.

A similar analysis is also carried out for the relaxed atomic structure of the products, where two H atoms have been adsorbed onto the tube wall and the electron clouds indicate that the C_1–C_2 bond is severed. The equilibrium H_1–C_1, H_2–C_2, C_1–C_2 lengths for the final products are 1.07 Å, 1.07 Å, 2.50 Å, respectively, see Fig. 8.34(d). The present result for the H–C bond length is consistent with the outcome of Yang and Yang[32] where the reported distance between C and H is 1.09 Å for two hydrogen atoms bonded to two adjacent carbon atoms on a basal plane of graphite. The H_1 and H_2 atoms show positive charges of 0.151, while the neighboring C_1 and C_2 carbon atoms show negative charges of –0.216. The overlap population of H_1–C_1, H_2–C_2, and C_1–C_2 and H_1–H_2 are 0.364, 0.364, 0.007 and 0 respectively, indicating that H_1 and C_1 have formed a covalent bond, and the same can be concluded for the H_2 and C_2 atoms. The overlap population between H_1 and H_2, and between C_1 and C_2, is almost zero, inferring the non-binding between C_1 and C_2, and between H_1 and H_2. This result is consistent with the fact that the C–C bonds in smaller radii nanotubes are under relatively greater strain, especially those perpendicular to the tube axis[59,60]. These C–C bonds are therefore more susceptible to being weakened or broken to form H–C bonds when H atoms are appropriately introduced. Gang *et al.*[61] predicted that two



8.36 Orbital overlap population between hydrogen and carbon atoms in the small model system of the final product.

rows of H atoms chemisorbed on selective sites exterior to the smaller armchair nanotubes will sever the nearest-neighbor C–C bonds of the tube. To verify the above results, the orbital overlap populations are also calculated for the products and illustrated in Fig. 8.36. For the optimized structure of the small model, we obtained E_{HOMO} as –5.35 eV, E_{LUMO} as –0.92 eV, and E_{gap} as 4.43 eV, which is greater than both the stable intermediate states discussed earlier. The orbital overlap population curve indicates that the H_1–C_1, H_2–C_2 bonds corresponding to the HOMO level are bonding, while those of H_1–C_2 and C_1–C_2 are non-bonding. On the other hand, the LUMO level coincides with the C_4–C_5 and C_{11}–C_{12} bonds exhibiting strong antibonding.

The simulation results illustrate the exothermic nature of the reaction when H atoms are adsorbed onto the outside of the CNT wall. However, as dissociating one hydrogen molecule into two hydrogen atoms requires about 4.57 eV (105.269, kcal/mol)[58,62] which is higher than the exothermic energy of 4.13 eV for two H atoms being chemisorbed onto the CNT, the entire reaction may still be required to absorb a certain amount of energy if we include the dissociation of the hydrogen molecule into hydrogen atoms.

8.7 References

1. Iijima S., *Nature* (London) **354**, 56 (1991).
2. Yakobson B.I., Brabec C.J., and Bernhold J., *Phys. Rev. Lett.* **76**, 2511 (1996).
3. Yakobson B.I., Campbell M.P., Brabec C.J. and Bernholc J., *Comput. Mater. Sci*. **8**, 341 (1997).
4. Brenner D.W., Shenderova O.A., Harrison J.A., Stuart S.J. and Sinnott S.B., *J. Phys. Condens. Matter* **14**, 783 (2002).
5. Abraham F.F., Broughton J.Q., Bernstein N. and Kaxiras E., *Europhys. Lett.* **44**, 783 (1998).
6. Abraham F.F., Bernstein N., Broughton J.Q. and Hess D., *MRS Bull.* **25**, 27 (2000).
7. Broughton J.Q., Abraham F.F., Bernstein N. and Kaxiras E., *Phys. Rev. B* **60**, 2391 (1999).
8. Mao Z., Garg A. and Sinnott S.B., *Nanotechnology* **10**, 273 (1999).
9. Sinnott S.B., Shenderova O.A., White C.T. and Brenner D.W., *Carbon*, **36**, 1 (1998).
10. Lennard-Jones L.E., *Proc. Roy. Soc. London, Ser. A* **106**, 441 (1924).
11. Colombo L., *Comput. Mater. Sci.* **12**, 278 (1998).
12. Colombo L. and Rosati M., *Comput. Phys. Commun*. **128**, 108 (2000).
13. Xu C.H., Wang C.Z., Chan C.T. and Ho K.M., *J. Phys. Condens. Matter*. **4**, 6047 (1992).
14. Sherwood P., Hybrid quantum mechanics/molecular mechanics approaches, in *Modern Methods and Algorithms of Quantum Chemistry, Proeedings*, 2nd Edn, Ed. J. Grotendorst, NIC Series, Vol. 3, Jülich: John von Neumann Institute for Computing, 2000, 285–305.
15. Rafii-Taber H., Hua L. and Cross M., *J. Phys. Condens. Matter.* **10**, 2375 (1998).
16. Dong Q., Wagner G.J. and Liu W.K., *Appl. Mech. Rev*. **55**, 495 (2002).
17. Dong Q., Wagner G.J. and Liu W.K., *Comput. Methods Appl. Mech. Eng*. **193**, 1603 (2004).
18. Srivastava D., Menon M. and Cho K., *Phys. Rev. Lett*. **83**, 2973 (1999).
19. Belytschko T., Xiao S.P., Schatz G.C. and Ruoff R.S., *Phys. Rev. B* **65**, 235430 (2002).
20. Kudin K.N. and Scuseria G.E., *Phys. Rev. B* **64**, 235406 (2001).
21. Yeak S.H., Ng T.Y. and Liew K.M., *Phys. Rev. B* **72**, 165401 (2005).
22. Roy G. and Yung K.S., *J. Chem. Phys.* **56**, 3122 (1972).
23. Liu C., Fan Y.Y., Liu M., Cong H.T., Cheng H.M. and Dresselhaus M.S., *Science* **286**, 1127 (1999).
24. Ye Y., Ahn C.C., Witham C., Fultz B., Liu J., Rinzler A.G., Colbert D., Smith K.A. and Smalley R.E., *Appl. Phys. Lett*. **74**, 2307 (1999).
25. Alonso J., Arellano J., Molina L., Rubio A. and Lopez M., *IEEE Trans. Nanotechnol.* **3**, 304 (2004).
26. Nikolaev P., Thess A., Rinzler A.G., Colbert D.T. and Smalley R.E., *Chem. Phys. Lett*. **266**, 422 (1997).
27. Terrones M., Terrones H., Banhart F., Charlier J.C. and Ajayan P.M., *Science* **288**, 1226 (2000).
28. Lee A.M., An K.H., Kim W.S., Lee Y.H., Park Y.S., Seifert G. and Frauenheim T., *Synth. Metals* **121**, 1189 (2001).
29. Arellano J.S., Molina L.M., Rubio A., Lopez M.J. and Alonso A., *J. Chem. Phys.* **117**, 2281 (2002).
30. Veronica B., Jochen H. and Gustavo E.S., *J. Chem. Phys.* **120**, 7169 (2004).

31. Jeloaica L. and Sidis V., *Chem. Phys. Lett.* **300**, 157 (1999).
32. Yang F.H. and Yang R.T., *Carbon* **40**, 437 (2002).
33. Kerdcharoen T. and Morokuma K., *Chem Phys. Lett*. **355**, 257 (2002).
34. Svensson M., Humbel S., Froese R., Matsubara T., Sieber S. and Morokuma K., *J. Phys. Chem*. **100**, 19357 (1996).
35. Vreven T. and Morokuma K., *J. Chem. Phys.* **113**, 2969 (2000).
36. Vreven T., Morokuma K., Farkas O., Schlegel H.B. and Frisch M.J., *J. Comput. Chem.* **24**, 760 (2003).
37. Karadakov P.B. and Morokuma K., *Chem. Phys. Lett.* **317**, 589 (2000).
38. Dapprich S., Komaromi I., Byun K.S., Morokuma K. and Frisch M.J., *J. Mol. Struc. (Theochem)* **462**, 1 (1999).
39. Maseras F. and Morokuma K., *J. Comput. Chem.* **16**, 1170 (1995).
40. Bauschlicher, Jr, C.W. and Harry P., *J. Chem. Phys.* **103**, 1788 (1995).
41. Davies D.W., *The Theory of the Electric and Magnetic Properties of Molecules*, New York: Wiley, 1967.
42. Han S. and Lee H., *Carbon* **42**, 2169 (2004).
43. Becke A.D., *J. Chem. Phys.* **98**, 1372 (1993).
44. Becke A.D., *J. Chem. Phys.* **96**, 2155 (1992).
45. Becke A.D., *J. Chem. Phys.* **97**, 9173 (1992).
46. Becke A.D., *J. Chem. Phys.* **98**, 5648 (1993).
47. Becke A.D., *J. Chem. Phys.* **104**, 1040 (1996).
48. Becke A.D., *J. Chem. Phys.* **107**, 8554 (1997).
49. Lee C., Yang W. and Parr R.G., *Phys. Rev. B* **37**, 785 (1988).
50. Rappe A.K., Casewit C.J., Colwell K.S., Goddard W.A. and Skiff I.W.M., *J. Am. Chem. Soc*. **114**, 10024 (1992).
51. Bauschlicher, Jr, C.W., *Nano Lett.* **1**, 223 (2001).
52. Bauschlicher, Jr, C.W. and So C.R., *Nano Lett*. **2**, 337 (2002).
53. Lee S. and Lee Y., *Appl. Phys. Lett*. **76**, 2877 (2000).
54. Lee E.C., Kim Y.S., Jin Y.G. and Chang K.J., *Phys. Rev. B* **66**, 073415(1) (2002).
55. O'Boyle N.M. and Vos J.G., GaussSum 0.9, Dublin City University, http://gausssum.sourceforge.net.
56. Yildirim T., Gülseren O. and Ciraci S., *Phys. Rev. B* **64**, 075404 (2001).
57. Gülseren O., Yildirim T. and Ciraci S., *Phys. Rev. Lett.* **87**, 116802 (2001).
58. Han S.S. and Lee H.M., *Carbon* **42**, 2169 (2004).
59. Blasé X., Benedict L.X., Shirley E.L. and Louie S.G., *Phys. Rev. Lett.* **72**, 1878 (1994).
60. Sánchez-Portal D., Artacho E., Soler J.M., Rubio A. and Ordejón P., *Phys. Rev. B* **59**, 12678 (1999).
61. Gang L., Henry S. and Nicholas K., *Phys. Rev. B* **68**, 205416 (2003).
62. Lide D.R., *CRC Handbook of Chemistry and Physics*, Boca Raton, FL: CRC Press, 1995.

9

Multiscale modelling of structural materials

D. PORTER, QinetiQ, UK

9.1 Introduction

Perhaps the greatest practical challenge to materials modelling is the prediction of failure in structural materials. Since most of the specific modelling techniques are covered elsewhere in this book, this chapter looks at structural materials from the perspective of an applied modeller in an 'industrial' context. Here, modelling is required to make reliable quantitative predictions of a wide range of different materials and combinations of materials in applications that span dimensions from nanometers to tens of metres and times from microseconds to tens of years. Since lives are at risk in many of these engineering applications, absolute accuracy of predictions is moderated by conservative design tolerances; ironically, experience suggests that where reliability is important, a broad physical description is often better than a very precise empirical relation, which may break down under slightly different conditions. Moreover, the modelling has to be effective and efficient in terms of both cost and time.

Griffith first asked the question about why there are such large variations in the strength of apparently similar materials.[1] This question is arguably the start of multiscale modelling of materials. Writing this chapter in the Griffith Building of a research organisation descended from the Royal Aircraft Establishment in which Griffith worked, the same question is arguably just as valid now as it was in 1920, albeit from a position of greater understanding about how strong a material might be and why it usually falls far short of that ideal performance. It seems appropriate, therefore, to focus here upon modelling the engineering properties of materials from the atomic scale up to the point of predicting these properties as a constitutive equation input to continuum-level techniques such as finite element analysis (FEA).

Although modelling aims to use best available theory, modelling structural materials often exposes severe limitations in 'fundamental' theory, such that empiricism is the only practical approach; this is the case in many structural properties of metals, for example. For this reason, the multiscale approach of

combining the contribution of physical effects over different scales of dimension and time to model the structural properties of materials is considered here from the pragmatic and quite limited perspective of an applied modeller; a model that 'works' is better than an elegant but non-validated theory.

The chapter starts with a brief introduction and review of general aspects of structural materials and multiscale modelling, and then looks specifically at different material family types and how modelling at different scales can be used to predict structural properties, including limitations on current methods. Finally, future requirements and trends for modelling are discussed. Where possible, practical examples are used to illustrate how modelling tools have been used in structural material problems.

9.2 Structural materials

9.2.1 Properties and scales

Structural materials are defined mainly by virtue of their properties of rigidity, strength, and toughness, and the effect of environmental variables such as temperature and humidity on those properties. As applications become ever more demanding, the combination with other properties such as density and dimensional stability gains in importance, and the complexity of new materials becomes so great that empirical development becomes prohibitively difficult.[2] Moreover, cost is a key issue.

The three main structural material families are described here very loosely by the terms metal, polymer, and ceramics; fibre-reinforced composites are not discussed here, but are considered in detail in a sister publication.[3] Since the largest body of research and development has historically been on metals, it is ironic that the greatest problems for quantitative predictive modelling remain in that area; namely the dominant effect of dislocations and grain boundaries upon the strength of metals. For this reason, metals are considered first, since they challenge the capability of current modelling methods, and are the subject of the most research attention. Polymers are then discussed from a deliberately provocative standpoint as an inherently 'simpler' structural material, which might provide novel insights that can be applied across other materials. Finally, other materials are considered under the catch-all term 'ceramics', which is also taken to include key structural materials such as glasses, concrete, and minerals.

The next step is to consider the spatial scales of dimension that control structural properties. At the limits of scale are classical continuum methods and molecular modelling, which are complementary tools that allow some basic mechanical properties such as elastic modulus to be related directly to chemical structure. However, the failure of structural materials is determined by defects and dislocations at a wide range of intermediate scales, so scales

are considered here very broadly in terms of their physical roles. The atomic/molecular scale determines the intrinsic properties and upper limiting values of all structural properties. The nanometer scale of inhomogeneous materials controls the synergy between adjacent materials and structures. Dislocations and 'defects' at the micron to millimetre scale limit the practical strength of materials. Finally, the bulk continuum scale above millimetres is the domain of classical engineering simulations, which use constitutive equations that embody effects at all the lower scales. In this chapter, the aim is to discuss in general terms how the three lower scales combine and contribute to the continuum level.

In engineering terms, scales of time are of equal importance to spatial scales. The three main engineering time scales are taken here to be those of shock events from 10^{-6} to 10^{-3} s, the quasi-static loading case of 10^{-3} s to hours, and the creep scale of hours to tens of years. Clearly, there is a synergy between the time and spatial domains in terms of mechanisms that control properties, but each spatial dimension can contribute to any time scale.

9.2.2 Motivation

Before considering detailed property profiles and modelling strategy in a problem, the first step in modelling structural materials problems is to scale the property bounds of general material families and match material types to an application. The method of Ashby is a valuable practical tool, based upon an enormous experimental dataset, which develops the simple concept of property maps of one property against another into a functioning engineering property database, and also embodies the need for a balance of two or more, often conflicting, properties.[4]

Realistically, not all materials have been synthesised and characterised, and experimental data on those that have is often limited, inappropriate, or unavailable. The author sees the role of multiscale modelling as helping the engineer to find the right material chemical composition with the right morphology to satisfy their specific property requirements. The material properties then need to be supplied in the right form and to a useful degree of accuracy for the engineer to use. Another important role of modelling is to provide a reasonable equation of state or constitutive equation for urgent problems in the field; again determined by engineering requirements.

A useful discussion of future needs for modelling in structural materials in the context of mechanics is presented in a recent DOE report.[2] Significant reviews of materials modelling also provide a useful background for the motivation for modelling efforts at the national and international level.[5–7]

9.2.3 Approaches

The most straightforward approach to modelling structural materials that spans all the scales is sequential; where modelling results at a lower dimensional scale are used as input to simulations at the next higher scale, in a sequence from atoms to continuum engineering properties. Inversely, if a problem or effect is observed at one scale, look at least one level below that for a solution. This approach is ideal for the applied modeller as a 'pick-and-mix' strategy to use methods that are appropriate to each very different problem. This allows empirical tools to be mixed pragmatically with more 'fundamental' methods.

The alternative modelling approach is termed 'concurrent', in that all the different scales are modelled in parallel to include the synergy between them. The problem here is to couple the different physics and dimensions in the domains that are used to partition an overall simulation process. This approach is relatively new, and likely to remain in the research sector for some time.

In addition to the presentation of individual modelling tools and combinations of tools in this book, a number of valuable reviews (and the reference therein) of multiscale modelling of structural materials are recommended: an edition of the *MRS Bulletin* gives a clear general overview of materials research by multiscale computer simulations, which looks at both methods and specific applications;[8] two reviews by Ghoniem outline the methods used in multiscale modelling, with particularly useful emphasis on nanomechanics and micromechanics;[9,10] Lu and Kaxiras give an elegant overview of multiscale simulations, which compares the sequential and concurrent approaches to multiscale modelling, with interesting comments on the difficulty of bridging time scales from atomic vibrations to engineering events.[11]

The above references concentrate on metals. A useful overview of multiscale modelling of polymers is edited by Kotelyanskii and Theodorou.[12] The earlier handbook by Bicerano presents a more applied series of articles that both present the modelling methods and give examples to illustrate their application.[13]

9.3 Metals

To simplify the discussion on modelling the structural properties of metals on the basis of scales of dimension, properties are classified as being either intrinsic to the atomic structure of a material or due to the nucleation and growth of dislocations under load. Intrinsic properties are determined largely by energetics at the atomic scale, and dislocations are a consequence of effects around the micrometer scale.

9.3.1 Intrinsic properties

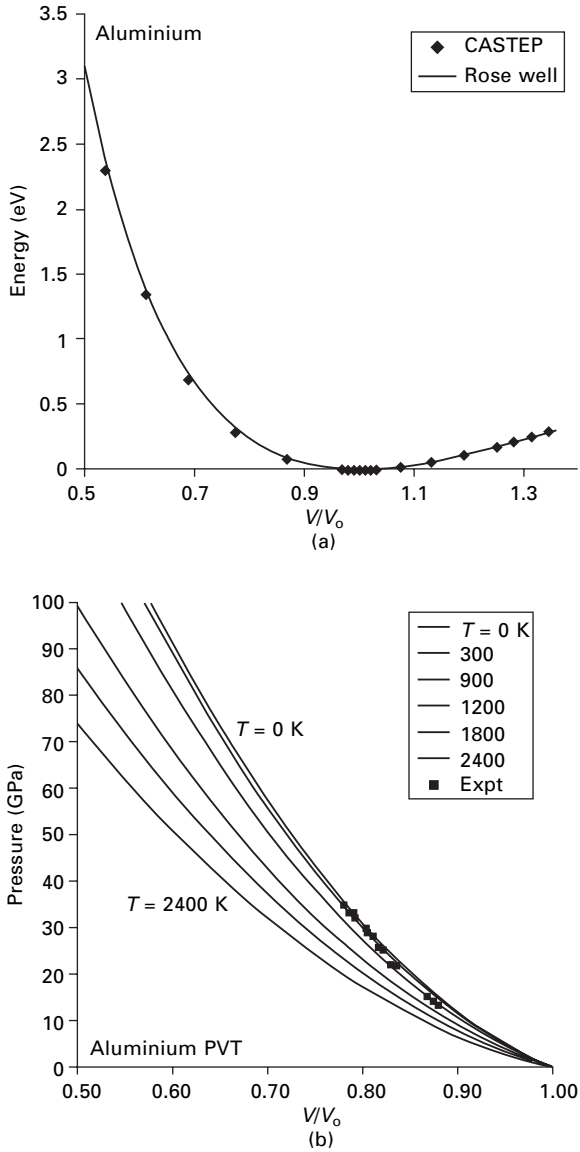
Intrinsic properties are taken here to mean the volumetric, thermal, and small-strain elastic properties of a solid material. These properties are determined directly by inter-atomic bonding, which is quantified by potential energy functions in the form of potential wells that have a minimum (negative) energy at the equilibrium position of atoms in a structure. Potential wells are calculated as the energy-dependence of some dimensional parameter (inter-atomic spacing or unit cell volume, for example), with constraints specified on the overall geometry changes that are allowed in a simulation; whether atom positions are allowed to relax in the perpendicular axes after imposing a uniaxial deformation, for example.

Potential wells can be calculated by quantum mechanics, molecular mechanics, or simply have an empirical form to fit experimental data, obtained usually from pressure–volume measurements. For metals, the author uses the plane-wave density functional method CASTEP,[14] since it is a reliable and flexible code with the essential capability for structural properties of cell minimisation; noting here that most applied modellers do not develop code and have limited time to test new or different packages. No method is perfect, but extensive in-house validation and an awareness of the consequences of the limits of accuracy to any specific composition or property simulation generally lead to an acceptable predictive capability. The speed of molecular mechanics is essential for large-scale modelling such as the simulation of dislocations.

Potential wells predicted by quantum methods are immediately applicable in the screening of the likely space group that any given alloy composition might preferentially adopt as its lowest energy state, and predicted values of density are usually reliable enough for engineering calculations. Possible phase transitions between space groups can be inferred if potential wells for different structures cross or the energy difference is of the order of the thermal energy in the system. For magnetic systems, such as the multiple transitions in iron, this requires that exchange interactions are included in the method.

Empirical potential functions are useful for validation purposes, but data availability limits their use. The ‘universal’ potential function of Rose works very well for metals, since it has the flexibility to fit the shape and depth of the well for most materials.[15] The analytical form of this potential function is very useful for calculating intrinsic properties, since expressions can be derived in terms of a limited number of parameters such as the well depth (cohesive energy) or the asymmetry of the well about the minimum.

To illustrate the form and use of potential wells, Fig. 9.1(a) compares the predictions of CASTEP with the empirical potential well for aluminium.[16] Figure 9.1(b) then uses the CASTEP potential well to predict the pressure–



9.1 (a) Comparison of CASTEP predictions with the empirical potential well for aluminium. (b) Predicted pressure–volume–temperature relations (line) compared with experimental data (points).[16]

volume–temperature (PVT) equation of state by differentiating energy with respect to volume to give pressure. The zero temperature well is used for reference, and pressure is simply scaled with the cohesive energy density as a function of temperature; this is the depth of the potential well (cohesive

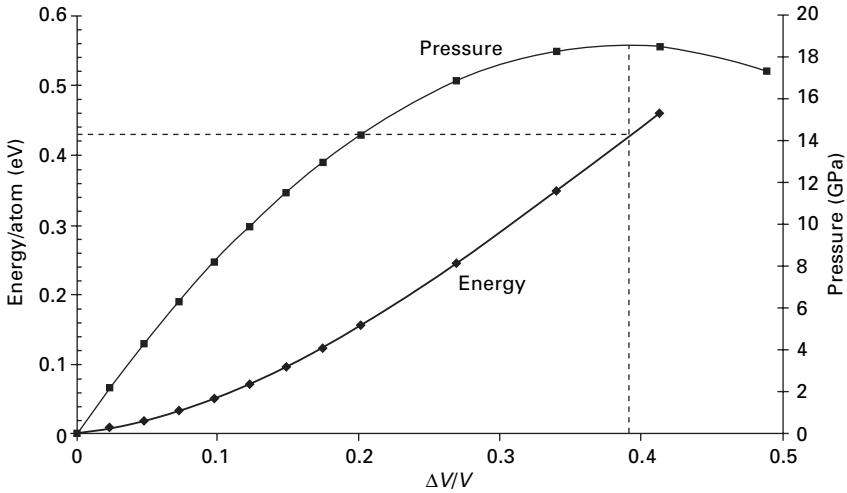
energy minus thermal energy) divided by the volume at temperature *T*. Points in Fig. 9.1b are data from plate impact experiments that were used to parameterise the empirical well. Such predictions of equations of state are valuable for dynamic FEA simulations of impact or explosive events, for example, since plate impact experiments are both expensive and time-consuming. The form of empirical potential wells can be helpful in compression at high pressure, since the pseudopotentials of some metals can be unreliable under these conditions and a good general fit of a Rose-type well can be a useful cross-reference check.

Pressure–volume relations from a potential well give the bulk elastic modulus, which can be used to estimate the Debye temperature, θ_D.[17] The Debye temperature then gives the heat capacity,[18] which is used to reduce the cohesive energy to scale the pressure predictions outlined above. Thermal expansion coefficients can be calculated from the asymmetry of a potential well, simply by averaging the volumetric changes at a series of energy (thereby temperature) points on either side of the well minimum.

A more general extension of the pressure–volume calculations to predict bulk elastic modulus is to predict the full stiffness matrix using a series of potential energy wells around the minimum energy equilibrium point with different deformation constraints. In practice, many metals have redundancy in the full stiffness matrix, which allows a small number of calculations to predict a sufficiently populated matrix for engineering calculations. At its simplest, bulk modulus and C_{11} calculations with volumetric and uniaxial deformations can estimate most of the elastic constants and Poisson's Ratio or Lamé parameters; to make life easier, this process is automated in the commercial CASTEP code. Mayer shows how *ab initio* calculations can be used to predict elastic constants and thermal expansion in quite complex Laves phases.[19] The experience of the author is that such predictions are again usually either within experimental error or close enough for acceptable engineering simulations, provided that all relevant effects such as exchange interactions are included.

Given that atomistic simulations can provide many intrinsic properties of metals, how far can this approach be taken towards predicting mechanical and failure properties under realistic conditions? Griffith showed that whiskers of glass with diameters of the order microns have strength far higher than that of bulk material.[1] Similarly, metal whiskers such as copper have shear strengths of about 3 GPa, relative to a typical measured yield stress an order of magnitude lower.[20] This 'ideal' elastic strength has been reproduced in *ab initio* simulations of shear stress as a function of engineering strain and is determined by an elastic instability point, at which one or more stresses go through a maximum.[21]

The elastic instability is illustrated in Fig. 9.2 for the volumetric strain of copper, with a maximum in pressure at a volumetric strain of about 0.4 and



9.2 Instability criterion in the volumetric potential function of copper.

an energy of about 0.4 eV/atom in a CASTEP simulation. The author has suggested at the simplest level that such an instability point can also be identified with the melting temperature, T_m, where a material changes from a solid to a liquid state via the Born criterion as the second derivative of energy in the potential well tends to zero.[22,23] This was also suggested by Guinea as a development of the properties that can be estimated from the 'universal' potential well, in the form of a good correlation between T_m and the depth of the potential well.[24] This issue is discussed in more detail by Wang, with a broader discussion of the different Born criteria for the combination of thermal and mechanical energy for instability criteria;[25] see below under dislocations.

Although the elastic instability is important mainly as an ideal reference strength, the thermal instability for melting may be useful in a more general manner as a link to predicting phase diagrams for metal alloys by thermodynamic methods such as ThermoCalc and CALPHAD.[26,27] Work here has shown that the general trends in the liquidus temperature of alloys over their composition range can be scaled quite well using the maximum in the differential of volumetric potential functions, although the absolute value of a transition temperature prediction must be taken with caution due to error bars. As a general comment, the use of predicted intrinsic properties in empirical methods is a very practical way to use atomistic simulations at higher scales, and tools such as JMatPro (a spin-off from the CALPHAD approach) can provide useful values of applied engineering properties.[28]

9.3.2 Failure through microscale 'defects' and dislocations

After the very positive discussion on predicting intrinsic properties at the atomic scale, we now need to take a realistic look at properties that actually determine the use of structural materials; when, how, and why do they fail? In practice, metals have values of strength (or stress to failure) that are significantly lower than the ideal values predicted from elastic instability criteria. Blame is invariably assigned to one of an enormous number of types of atomic scale defects and dislocations, usually separated by dimensions of the order of microns, which are either embedded in the material or can nucleate and grow spontaneously under load.

Historically, Griffith first suggested in 1921 that dimensions are an important factor in strength and suggested that microscopic cracks are responsible for the relative weakness of brittle solids. By 1934, dislocations had been identified in crystalline solids. The consequence of some major incidents such as metal failure in ships during World War II was a massive effort to understand the role of dislocations on metal properties. By the early 1950s, numbers of books reflect a confidence to collate the large body of experimental observations on dislocations and develop physical models to understand the role of dislocations on a semi-quantitative level.[29,30] By 1968, student texts such as that of Honeycombe are able to present a coherent picture of the plastic deformation of metals due mainly to dislocation growth.[31] Since then, more experimental work with ever-increasing levels of refinement and considerable resources allocated to numerical simulations of deformation in metals still leaves us not able truly to predict dislocation effects at an *ab initio* level.

The rather negative tone of that last comment is not meant as a slight to the excellent work of recent years to develop methods that will eventually allow defect and dislocation effects to be predicted reliably enough for use in engineering design and simulations. A brief commentary article in *Science* in 1998 perhaps reflects the point where computer simulation of failure in metals comes-of-age.[32] Chapters in this book give a flavour of recent work, and general literature and web searches using names such as M. Ortiz, E. Kaxiras, N.M. Ghoniem, and S. Yip yield a considerable body of excellent overviews of and contributions to developments in microscale modelling.

Current state-of-the-art modelling of dislocation effects is probably the Livermore work on ParaDis, which uses molecular and dislocation dynamic simulations on the (currently) most powerful computer in the word to predict stress–strain curves in metals such as iron and aluminium up to strain of about 1 %.[33] The key role of these large-scale simulations is to help us to understand critical features of dislocation nucleation and growth, which can then be used in more practical modelling tools at a more realistic economic scale; the main concern of the author is that the computational routines that

control the simulations are not just 'quick fixes' that give apparently correct results, rather than serious underpinning physics. Eventually, the author hopes that insights from these simulations will contribute to our understanding sufficiently to allow models to be developed that use *ab initio* local simulations of intrinsic properties as input to predict 'bulk' properties that incorporate dislocations. For example, simplification of dislocation modelling to generic solutions from detailed modelling of the volume immediately surrounding a dislocation as discrete 'zones' within a relatively simple elastic continuum medium using a 'concurrent' approach mentioned in Section 9.2.3.

Clearly, most readers will not be able to use supercomputer simulations or develop new microscale simulation techniques. However, there are a number of key issues at a more modest level that warrant more extensive discussion and are relatively straightforward developments of concepts from the student textbook level that have been facilitated by calculations at the atomic scale. Two such important and interrelated issues are dislocation nucleation and grain size effects.

The search for criteria for the nucleation of dislocations generally follows two possible routes. The first is a simple Peierls-type stress for initiation of a local slip-like event that is much lower than the ideal failure stress, but this route is not popular. The main route is to assume that the ideal elastic instability criterion is the driver and find ways to combine all the thermodynamic energy contributions to achieve a local shear instability. Taking the work of Yip and coworkers as an example, the starting premise is that the Born criterion of modulus tending to zero is a condition for initiating elastic instability; they showed that some combinations of modulus and hydrostatic loading could be used to specify instability conditions.[25] Five years later, this basic idea emerges as an energy-based local elastic stability criterion, termed the Λ-criterion, which is used in an elegant combination of molecular dynamics and FEA computer simulations, nano-indentation experiments and experimental bubble-raft models to provide atomistic insights into the early stages of plasticity.[34] Although attractive in its elegance, the specific nano-indentation simulations are also quite limited, in that the nanometer scale of events around dislocation initiation under the tip of an indenter does not qualify as a 'bulk' process, in the same way as whisker experiments do not reflect bulk deformation events. This healthy scepticism is expressed well in a recent article by Miller and Acharya, who simultaneously suggest that a nucleation criterion based on the magnitude of local stress components is incorrect and present a criterion motivated from continuum thermodynamic considerations, based on stress-gradients.[35]

Stress gradients were used by Cottrell in an early formulation of the Hall–Petch relation for the effect of grain size, d, on the yield stress, σ_y, in a metal[36]

$$\sigma_y = \sigma_i + \frac{k_y}{d^{1/2}} \qquad [9.1]$$

where σ_i is the yield stress of a single crystal (due to dislocation nucleation, perhaps) and k_y accounts for unpinning of dislocations. In practice, the Hall–Petch relation is a loose empirical guide to yield stress in granular materials, and cannot yet be predicted for any given metal or alloy. Increased interest in nanocrystalline materials has focused attention on the way the Hall–Petch relation might break down at very small grain sizes;[37] termed the inverse Hall–Petch effect.

A remarkably simple composite model for the Hall–Petch effects down to nanograined materials has been suggested by Jiang and Weng.[38] Here, crystalline grains are embedded in a disordered matrix of grain boundaries with a fixed thickness of about 1 nm around each grain; at larger micrometer grain sizes, deformation is dominated by the grains; below a critical 'equicohesive' grain size, deformation is controlled by the plasticity of the grain boundaries. The main problem with the model is that some of the input parameters (such as yield stress due to dislocations in the components) cannot be predicted. Grain boundaries, in turn, change dislocation dynamics, which demonstrates the vicious circle of problems for modelling at this scale.

9.3.3 Simplified bridging of scales

Modelling structural materials is not just about detailed modelling with specific approaches, it is also about a very general understanding of the interrelation between effects at different scales of dimension. We might be able to put more than 10^{10} atoms in a virtual box and simulate a dislocation at intervals of one micron using molecular dynamics, but could we not use the advances in atomistic modelling of local intrinsic properties within a very simple scaling approach to achieve the same effective result, and simultaneously learn more about the key physics involved in inter-scale processes?

At its heart, micromechanics looks at how mechanical energy density in a material generates voids or dislocations that can develop into large-scale cracks and cause failure. It is interesting to look at why the microscale is associated with 'defects' by considering at a very simple level the 'defect' surface area that can be generated spontaneously by the application of a stress, σ. Consider a cube of side length d inside a perfect crystal with a modulus, M, and an energy per unit area of defect or free surface, Γ. The elastic energy density inside the cube volume can generate a 'defect' when

$$\sigma = \sqrt{\frac{12\Gamma M}{d}} \qquad [9.2]$$

Taking the example of a copper-like metal with $M \approx 130$ GPa and

$\Gamma \approx 1$ J/m^2 as generic parameters, a yield stress of 250 MPa (the value of σ_i in the Hall–Petch relation in the composite model[38]) corresponds to a defect spacing of $d \approx 20$ μm. The ideal shear strength of copper is suggested to be about 2.5 GPa with a shear modulus of about 30 GPa, so the minimum defect spacing corresponds to d ≈ 57 nm. Unsurprisingly, this spacing corresponds very well with the maximum flow stress and grain size in both experimental[37] and modelling simulations.[39]

More generally, equation [9.2] raises a problem, since defects can be generated at any stress and will continue to be generated/grow at that stress with increasing strain to convert mechanical work into either free surface (brittle crack growth) or shear dislocations (plastic flow), depending upon the lowest energy mechanism for Γ. Why then do 'defects' start to form at a specific stress in a perfect crystal lattice? One simple suggestion is that this stress corresponds to the point where atoms in the lattice cannot reconfigure fast enough to maintain a minimum potential energy state, such that new mechanisms are invoked to absorb or dissipate the energy of deformation. Can this suggested condition be translated into a nucleation criterion?

Work on the failure initiation criteria of viscoelastic polymers and brittle inorganic materials has suggested that the stress for the reconfiguration condition is quantitatively equal to the loss modulus of a material,[40] since this is the stress at which atoms can naturally dissipate energy during the reconfiguration under strain. Since loss modulus is much smaller than the storage modulus, this suggests a relation in the loss tangent, tan δ

$$\sigma_i \cong B \tan \delta \qquad [9.3]$$

This suggestion immediately moves dislocation theory out of the realm of simple elastic potential functions and instability criteria. To a first approximation, loss tangent is the conversion of mechanical energy to heat in a deformation cycle. A simple relation can be derived to estimate tanδ as a thermomechanical process, where input of mechanical energy changes the elastic modulus, which then changes the Debye temperature, θ_D, and changes the heat capacity to convert mechanical to thermal energy in an irreversible cycle. To a first approximation, this process is embodied in the simple relation

$$\tan \delta \approx 0.12\ \alpha\theta_D \qquad [9.4]$$

where α is the volumetric thermal expansion coefficient. The two parameters α and θ_D can be predicted from atomistic calculations (Section 9.1) or measured. For the copper example, $\alpha \approx 5 \times 10^{-5}$ K^{-1} and $\theta_D \approx 343$ K, such that tan $\delta \approx$ 0.002 and the estimated stress for dislocation nucleation is $\sigma_i \approx 260$ MPa. The predicted loss tangent is not unreasonable for a metal, and the predicted nucleation stress agrees with that from experimental data used in the composite model for Hall–Petch.[38] Note that the expression for tan δ is not appropriate for low temperatures, where extra loss terms may be induced by the gradient

of heat capacity with temperature in the Debye functions, which are observed as Bordoni relaxation peaks.[41]

The combination of scaling of energy density for nucleation and energy dissipation for initiation is considered for polymers and ceramics later in this chapter as a more general concept. Clearly it does not embody all the effects required for a comprehensive treatment of defect/dislocation contributions to failure; for example, does the kind of inter-atomic bonding (covalent, metallic, van der Waal's) determine the tendency to void formation or shear dislocation? However, it is introduced here to broaden the arguments about where to go in the search for better models of structural materials beyond the over-simplistic use of purely elastic potential functions.

9.4 Polymers

9.4.1 Background

A polymer usually consists of a chain of chemically bonded atoms (C, O, N, Si), which interacts with its neighbours by relatively weak van der Waal's forces; a characteristic group of atoms in the chain is called a mer unit. The elastic modulus in the chain axis is of the order of 100 GPa (a simple *trans*-poly(methylene), $–CH_2–$, chain modulus is about 350 GPa,[42] for example), which reduces with the greater cross-sectional area of bulky side groups. Perpendicular to the chain axis, the modulus is of the order of a few GPa, and this modulus dominates the mechanical properties of most bulk polymers due to the additivity rules of compliance. Above the glass transition temperature, T_g, the polymer modulus drops to an order MPa, characteristic of rubbers. Polymers have two important structural reference time scales: below T_g, that of vibrations of individual atomic groups along the chain, $\tau_0 \approx 10^{-13}$ s; above T_g, the cooperative mode vibrations of the whole chain, which scales with τ_o approximately in chain length to the power three.

Texts such as that of Ward give a good introduction to the mechanical properties of polymers.[43] Reviews edited by Bicerano[13] and Kotelyanskii and Theodorou[12] present the different modelling methods for polymers up to the continuum scale, and a compilation of references relevant to FEA modelling of polymers is a useful guide.[44]

Perhaps the most important practical reference for polymers is the Group Contribution method of van Krevelen,[45] which is an elegant and comprehensive demonstration of the potential simplicity of modelling using largely empirical relations developed from a strong foundation in experiment and theory of polymer physics and chemistry. The remarkable simplicity of the concept of additivity of the contributions of small groups of atoms to the overall properties of a complex polymer suggests that the physics of the thermomechanical properties of polymers has an underlying simplicity. This simplicity is again

evident in the empirical method of Bicerano,[46] which uses little more than the connectivity of chemical bonding between atoms in a mer unit to predict many important physical properties.

The objective here is to show that most of the important structural properties of polymers can be predicted quantitatively across all scales of space and time/rate to an acceptable degree of accuracy for engineering simulations; even the highly non-linear dependence on strain, strain rate, and temperature. Thus, polymers are a good reference point for modelling structural properties of materials, and can offer fresh perspectives on modelling strategy that has been traditionally based on metals.

The following section presents a brief outline of how each of the tools discussed elsewhere in this book can be applied to different aspects of structural properties of polymers as a hierarchy of scales. A more personal approach to applied polymer modelling is then presented as an overall scheme of how energy storage and dissipation control mechanical properties.

9.4.2 Spatial scales

At the atomic scale, the main use of *ab initio* or semi-empirical quantum mechanics for polymers is to predict the structure and energetics of strong chemical bonding in the polymer chain. Stretching of atom–atom bonds is important to quantify the thermal stability of the polymer, since chain cleavage is the main mechanism for deterioration of mechanical properties over time at elevated temperatures. Bending of bond angles along the polymer chain controls the chain stiffness and the vibrational frequency for groups of atoms in the chain, $1/\tau_o$. This is important for understanding structural properties, since this frequency determines the main Debye temperature of skeletal modes, θ_1, which controls the zero point energy of polymers and heat capacity, which critically has a one-dimensional Debye functional form.[47]

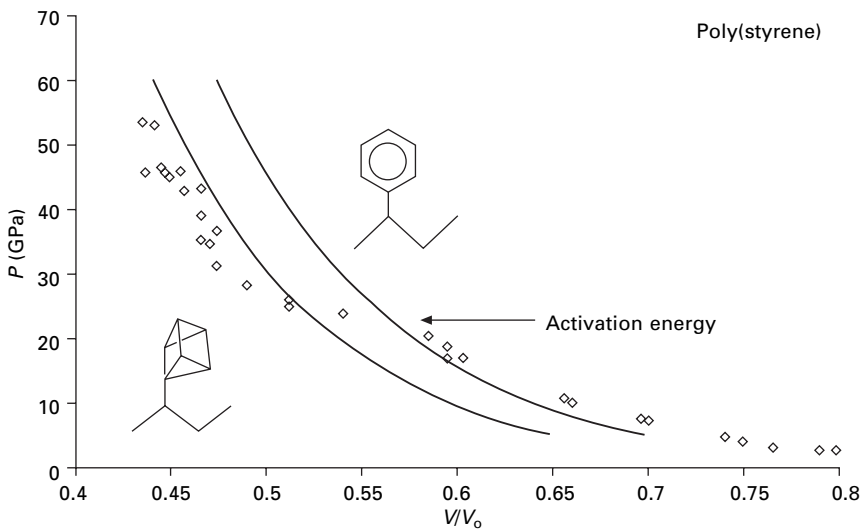
Quantum mechanics is useful to predict the conformations that are possible in a polymer chain, in terms of the stable isomeric states that can be adopted with different combinations of torsional angles in the chain backbone. The long-term stability of these higher energy states allows polymers to have stable amorphous structures. Similarly, the activation energy for the onset of vibrational modes in the polymer chain plays a key role in the low temperature relaxation spectra of tough engineering polymers.[48]

The main disadvantage of quantum methods is the small number of atoms that can be used in a model, which is particularly important for amorphous polymers. The ONETEP method looks promising for polymer and hybrid systems, in that the scaling between the number of atoms and computational size has been reduced to a level where thousands of atoms might be used in a simulation, and takes quantum methods into the dimensional scale of nanometers.[49] It is not yet clear whether this approach will be able to make

accurate predictions of mechanical properties that depend predominantly upon weak intermolecular forces; conventional quantum tools are often unable to combine these effects adequately with simulations of strong chemical bonds along the chain backbone.

Molecular mechanics and dynamics have been the main practical modelling tools for polymers. That atomic-scale simulations can predict many bulk properties of polymers is due to the relative lack of dislocations and defects at the micron scale in amorphous polymer structures. The use of empirical forcefields for weak intermolecular interactions allows properties such as elastic modulus to be predicted quite well, but simulation times of the order pico- to nanoseconds do limit the ability of molecular dynamics to reach equilibrium conditions or to model the important transition from shock to pseudo-equilibrium conditions.

Initial attempts here to model the equation of state properties of polymers under shock conditions using molecular dynamics were promising and allowed effective pressure–volume–temperature relations to be generated, leading to an understanding of some unusual effects in important polymer systems.[16] Figure 9.3 shows pressure–volume simulations for poly(styrene), with experimental data generated from plate impact experiments shown for reference. The 'kink' at about 20 GPa pressure is suggested to be due to collapse of the aromatic ring, and was reproduced by quantum mechanics calculations of activation energy between the two structures and molecular dynamics simulations of planar and collapsed ring groups; the same 'kink' is observed in many polymers that contain aromatic rings.[50]



9.3 Molecular mechanics predictions of impact pressure–volume–temperature.[16]

Molecular mechanics methods have been extended to replace individual atoms by groups of atoms along a polymer chain. This increases the size and simulation times that are possible in polymers and is useful in important problems such as comparing the penetration of small molecules into polymers over times of the order microseconds.[51] However, since failure is often (even in polymers) associated with events at the micron scale, simulations of mechanical failure even by these coarser grained approaches are not feasible, and this is an important limitation on the methods.

Personally, the author uses molecular dynamics simulations of structural properties of polymers with some reservations. Commercial potential functions such as COMPASS seem to give good predictions of volumetric properties (such as density), cohesive energy, and bulk elastic modulus, but the thermo-mechanical coupling and energy dissipation processes at a molecular level involved in engineering moduli and time-dependent viscoelastic properties should be viewed with a healthy scepticism.[52]

Nanoscale modelling of structural properties of polymers has been the subject of great interest in recent years, due mainly to the promise of great gains in strength and toughness of hybrid nanocomposite materials composed of nanoparticles such as clay or carbon nanotubes, dispersed in a polymeric matrix. The compilation edited by Baltá-Calleja and Michler gives a good background to nanostructural effects on the mechanical properties of polymers.[53] In practice, structural properties of synthetic nanocomposites have not yet realised this great promise, and successes have been due mainly to improvements in more 'functional' attributes such as fire retardation, increased barrier properties, or electrical conductivity. More conventionally, reductions in the domain size of the phases in semicrystalline polymers to the nanometre scale can improve failure characteristics considerably, and nature uses the nanoscale to create materials such as bone and silk with excellent mechanical properties; silk is used as an example in Section 9.4.3 below. Modelling may be able to help nanocomposites achieve their potential for structural properties by identifying where weaknesses occur and how they might be remedied.

Modelling at the nanoscale is now possible using large-scale molecular dynamics simulations, which have been able to identify problems in the dispersion of nanoparticles in polymers. For example, Odegard was able to predict the elastic properties of silica nanoparticle/polyimide nanocomposites and show that properties are controlled effectively by an interface region in the polymer with a thickness of the order 1 nm.[54] If the polymer macromolecules are unable to pack efficiently around the nanoscale inclusions, this interface region can effectively negate any potential advantages of dispersing more rigid nanoparticles. From a multiscale perspective the Odegard paper demonstrates an important effect on moving through modelling scales, since the perturbation of the molecular structure of the polymer around the

nanoparticles invalidates the use of continuum approaches to micromechanics, such as that of Mori–Tanaka, which assume two phases that are perfectly bonded together.[55]

A more pragmatic approach to modelling at the nanoscale is to consider the change in the intrinsic properties of materials due to the thermodynamic influence and constraints of one material upon the other around an interface, and how far these changes reach into the material. The author has applied this approach to bone,[56] where hydroxyapatite mineral and hydrated tropocollagen polymer are assembled in layers with thicknesses of about 2 and 5 nm respectively. By quantifying the enthalpy distribution through the interface as a pairwise sharing process between adjacent molecular layers, it was suggested that the energy sharing effect decreased as 2^n in the nth molecular layer. This simple approach allows continuum level models (such as composite stiffness) to be applied to nanocomposites by including the effect of interfaces on the individual properties of the components; for bone, the model shows that the morphology of compact bone has evolved with almost perfect layer thicknesses for optimal energy sharing, while retaining the distinct identity of the mineral and polymer phases.

For structural materials, modelling at the micron scale is all about fracture and failure. The case of composite materials is the subject of a sister publication.[3] The standard textbook in this area for polymers is by Kinloch and Young, which is also an excellent introduction to the structural properties of polymers.[57] Once a cavitation site has been nucleated, the standard fracture mechanics approach of elastic energy to generate a propagating crack can (in principle) be used. However, polymers are highly nonlinear viscoelastic materials, and caution is urged on anyone wishing to use classical fracture mechanics on polymers.

The energy density relation for separation of nucleated defects of equation [9.2] applies quite well to polymers; noting that the lower defect energy is that of void formation or cavitation in polymers. For a generic polymer glass (similar to poly(styrene), for example), material parameter values can be taken approximately to be $E \approx 3$ GPa and $\Gamma \approx 0.15$ J/m^2. The important case of defect generation for typically brittle polymers is for stresses below the macroscopic yield stress, which can be scaled by the empirical rule of thumb, $\sigma_y \approx 0.02E$, to suggest values of inter-cavity distance, $d \approx 1.5$ μm.

The second nucleation condition of equation [9.3] in loss tangent effectively defines a brittle–ductile transition condition for polymers of tan $\delta > 0.028$ for ductility.[40] Here, the bulk yield stress due to the ideal elastic instability condition of van der Waal's forces between chains is less than the nucleation stress for voids. Polymers with strong relaxation peaks in loss tangent at low temperature (such as polycarbonate) have high enough values of tan δ to impart considerable toughness. Thus, atomic scale relaxation processes link directly with microscale defect nucleation to determine failure properties.

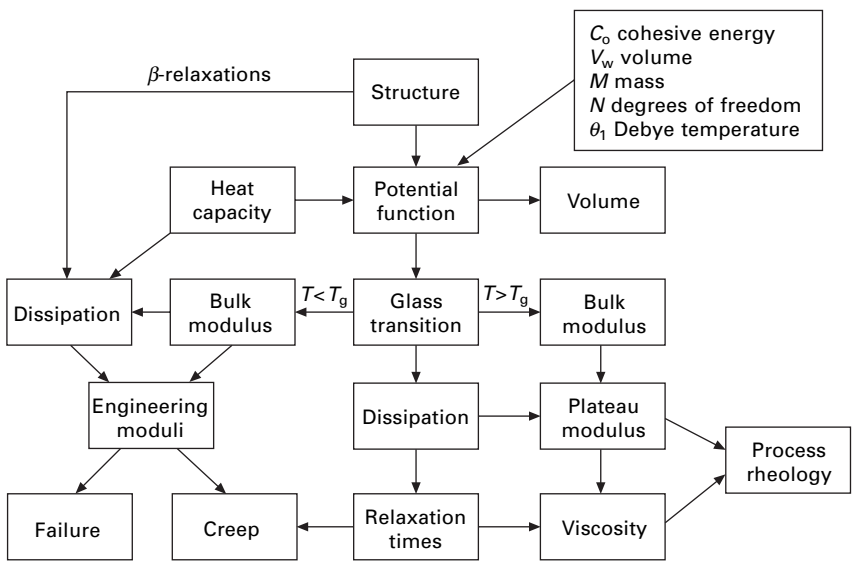
The author has successfully used the energy balance approach in equation [9.2] in a wide range of applied problems, such as the design of rubber particles to toughen intrinsically brittle polymers[40] and failure by gradual accumulation of damage in particulate composites.

9.4.3 A general modelling framework for polymers

Starting from the implication of the van Krevelen[45] and Bicerano[46] methods that the physical properties of polymers must have an underlying simplicity, the author developed the method of group interaction modelling (GIM),[40] which is shown schematically in Fig. 9.4.[58] The method calculates a small number of parameters for a characteristic group of atoms in a polymer that are sufficient to solve an ensemble-average potential function for interatomic groups normal to the chain axis

$$-E_{coh} + H_c + H_T = E_{coh}\left(\left(\frac{V_o}{V}\right)^6 - 2\left(\frac{V_o}{V}\right)^3\right) \qquad [9.5]$$

Negative values of cohesive energy, E_{coh}, are countered by positive thermal energy, H_T, and the zero point energy of configuration, H_c, to give the volume, V, relative to the volume at the minimum of the potential well, V_o. Physical properties are calculated using analytical structure–property relations derived directly from the potential function.



9.4 Model framework for Group Interaction Modelling[40,58].

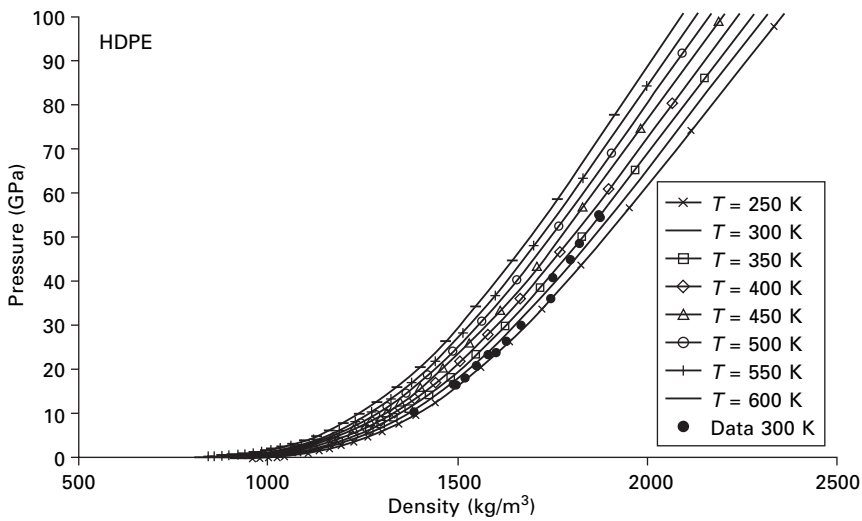
This relation suggests that temperature effects can be included directly into the potential energy well for thermomechanical properties via the thermal energy of skeletal mode vibrations in the polymer chain as a one-dimensional Debye function.[47] The second key point is that the zero point energy of the skeletal vibrations takes specific fractional values of 0.04 and 0.106 of the cohesive energy for crystal and amorphous structures respectively, which are fractionally additive for semicrystalline polymers; these values are the lowest 'quantised' energy levels of $k\theta_1/2$ and $3k\theta_1/2$, where k is Boltzmann's constant and θ_1 is the Debye temperature of skeletal modes.

The Born criterion for elastic instability[25] is used directly to predict the glass transition, T_g, and crystal melt, T_m, temperatures for group interactions

$$\left.\begin{matrix} T_g \\ T_m \end{matrix}\right\} = 0.224\ \theta_1 + \left\{\begin{matrix} 0.0513 \\ 0.084 \end{matrix}\right\} \frac{E_{coh}}{N} \qquad [9.6]$$

where N is the number of degrees of freedom per group. Not only is this a predictive relation for transition temperatures, which has been validated for T_g using a set of 250 polymers, but it also makes direct quantitative relations between different states of matter, which is a critical issue for structural materials.

The potential energy relation predicts pressure and bulk modulus directly from energy differentials of the potential function, and Fig. 9.5 shows in-house predictions for the PVT equation of state for semicrystalline high-density polyethylene, with experimental data from plate impact experiments at 300 K shown for reference.



9.5 GIM predictions of equation of state for high-density polyethylene.

The viscoelastic properties of polymers are apparently very complex, and GIM calculates these properties simply in terms of energy stored and energy dissipated during deformation; elastic energy is stored as volumetric changes against the bulk modulus and elastic modulus reduces as energy is dissipated by loss mechanisms at a molecular level.

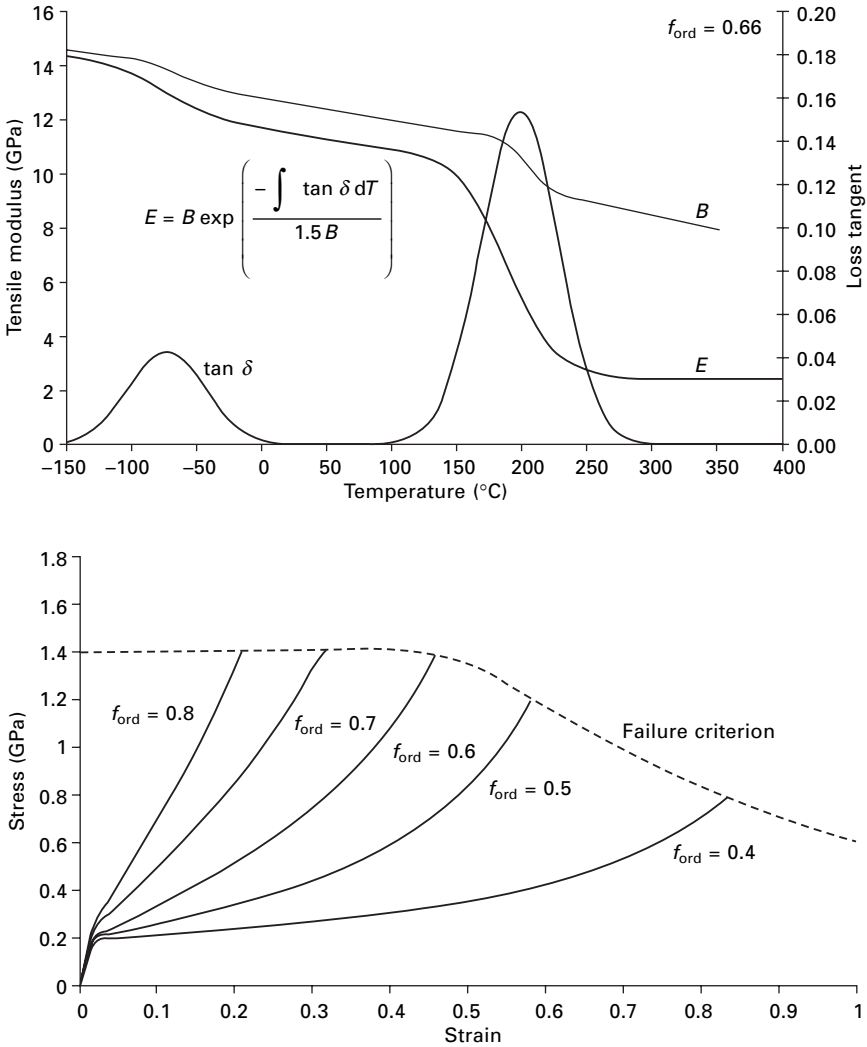
Polymers such as natural silk are also complex in their chemical and morphological structure, and provide an excellent example of how these multiple degrees of complexity can be rationalised into a straightforward modelling process.[59] First, the structural complexity is reduced by defining the fraction of ordered (crystal) and disordered (amorphous) states of a characteristic peptide segment within the nanoscale morphology. Bulk modulus, B, is calculated as a function of temperature using the potential function and the loss tangent through the glass transition peaks in the disordered fraction is used to predict the tensile modulus, E, using a relation derived from the loss due to thermomechanical coupling

$$E = B \exp\left(\frac{-\int \tan\delta \, dT}{A \cdot B}\right) \qquad [9.7]$$

where $A \approx 1.5$ GPa^{-1} is a term that quantifies the effect of a number of dimensional parameters of the peptide groups on dissipation. For an amorphous polymer, the tensile modulus in equation [9.7] becomes the rubber-like plateau modulus, such that elastic modulus is continuous through the glass transition zone and the dramatic change in modulus is quantitatively predicted from the energy dissipated by the change in heat capacity (entropy).

Figure 9.6a shows the predicted properties B, E, and tan δ as a function of temperature for a spider dragline silk with an ordered fraction f_{ord} of 66 %. The predictions of dynamic mechanical properties are then transformed into stress–strain curves to failure by using temperature as a dummy variable to predict strain and stress from the predicted thermal expansion coefficient. Other key points are that the yield condition at about 2 % strain is defined by the same Born instability condition as the glass transition temperature and that post-yield work hardening is the gradual transformation of rubberlike disordered states back to crystal or glassy states by mechanical work. Figure 9.6b shows the predicted stress–strain map for a number of model silks with different ordered fractions.

Detailed discussion of other aspects of GIM are beyond the scope of this article, but the brief outline presented here is intended to suggest that alternative approaches to modelling structural materials can raise important issues and questions about the underlying mechanisms that control structural properties that are beyond the simple framework of elastic simulations.



9.6 Model for spider silk.[59] (a) Predicted dynamic mechanical properties. (b) Predicted stress–strain relations.

9.5 Ceramics

This disproportionately short section on ceramics does not reflect their importance as structural materials. Unfortunately, ceramics do not lend themselves very well generally to multiscale modelling approaches, mainly because their predominantly brittle bulk fracture properties are not dependent as much on the microscale as metals or polymers. The main exceptions to this gross oversimplification are nanograined ceramics[60] and 'superhard'

materials made from nanometer thick layers,[61] but the general lack of microscale effects still pertains.

As a general reference, *Introduction to Ceramics* by Kingery, Bowen, and Uhlmann is invaluable, with Chapter 15 giving a good introduction to structural properties.[62] More specifically for inorganic glasses, Feltz gives an excellent overview of general physico-chemical properties.[63]

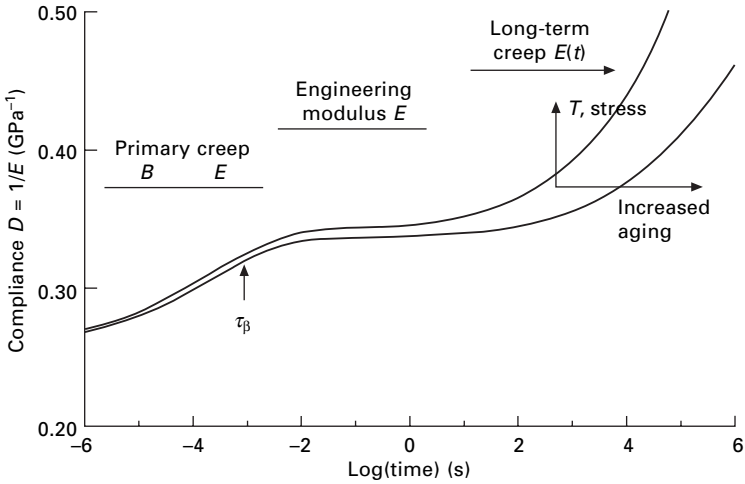
All the atomistic methods discussed for the intrinsic properties of metals and polymers are generally applicable to ceramics, provided that care is taken to ensure that quantum or molecular mechanics codes are able to cope with the specific bonding in any ceramic of interest. Some quantum codes might have trouble with ionic bonding, and the author has found particular problems with quantum codes applied to silica-based materials. Oxygen bonding to the silcon atom seems to be so mobile that large numbers of different structures are possible without significant changes in the bond lengths; this can have a significant effect at high stress, where structural transitions are likely.[64]

Ceramics are mainly brittle; the covalent and ionic bonding seems to have a lower energy to create free surface than induce a dislocation. The loss tangent model for failure stress of equations [9.3] and [9.4] can be scaled with generic parameters of the order $\alpha \approx 25 \times 10^{-6}$ K^{-1} and $\theta_D \approx 300$ K, giving tan δ for the ratio of failure stress to modulus of about 0.001. Basic 'strength' is not much different to metals perhaps, but the toughness, as area under the stress–strain curve to failure, is small.[4]

The intrinsic brittleness of ceramics due to low energy dissipation suggests that the elastic instability criterion is an interesting condition at which to consider the spacing of defects nucleated at this stress from equation [9.2]. Taking the extreme case of transition metal nitrides with a high modulus of about 700 GPa and using an ideal elastic instability stress criterion of about $M/30$ and a generic $\Gamma \approx 1$ J/m^2 suggests that the optimum hardness should be at a spacing of about 15 nm. This is the thickness of layers of different nitrides reported to give maximum hardness in 'superhard' nanolayered materials.[61] Thus, potentially weak interlayer boundaries can promote dislocation nucleation and allow a ceramic to use the elastic instability criterion to have properties such as hardness approaching that of 'ideal' materials.

9.6 Time scales

Unfortunately, modelling time effects in structural materials is largely empirical, with little attention drawn to the problem in literature. For modelling purposes, time can be split very broadly into three main scales that reflect different energy dissipation or relaxation mechanisms: impact, quasi-static, and creep. Figure 9.7 shows the specific case of the tensile compliance (inverse of modulus) of polymers, since polymers have a relatively more straightforward



9.7 Time domains in polymers.[40]

generic form than metals. The eventual aim of multiscale modelling is to be able to predict the complete time response by combining the shortest timescale of atomic level vibrations with effects under load (such as atomic redistribution, dislocation nucleation rate and velocity, and phase transitions) to predict the time- and rate-dependence of mechanical properties. Modelling of time scales must be linked intimately with scales of spatial dimension and the deformation mechanism at each scale.

Impact or shock events occur at times less than about 10^{-3} s, which is often called primary creep, since a material relaxes from an instantaneous response to load by means of a redistribution of atoms to attain a lower total energy state. Although large-scale atomic simulations have made some links with spatial dimensions in this regime,[65] the author finds little practical insight into important structural effects with modelling time scales of less than 1 microsecond. In practice, this scale is very important in many engineering applications and experiments such as plate impact tests are generally modelled as instantaneous volumetric deformation in the deformation axis (mode 1 bulk, for example). In this mode, mechanical properties are taken to be dominated by the bulk modulus, and dynamic finite-element calculations can be made using PVT relations for an equation of state, without a full constitutive model. Examples in sections 9.3.1 and 9.4.3 for metals and polymers show how impact equations of state can be generated directly from atomistic simulations.

Quasi-static loading from 10^{-3} to 10^{3} s is the relatively stable 'plateau' region in time after molecules have redistributed following initial loading and before the onset of creep. This is the time scale of most laboratory experiments. Perhaps surprisingly, simulations of intrinsic properties using

methods such as quantum mechanics can predict the full stiffness matrix and properties such as Poisson's Ratio in these time scales very well; the surprise being that stable properties often do not develop until after about 10^{-3} s in practice, and simulations jump straight to the relaxed state in a purely elastic mode.

Creep is the slow and generally irreversible displacement of molecules under load. Many empirical or phenomenological models have been developed to 'predict' creep, mainly as a means to extrapolate deformation to times of, say, one order of magnitude longer than available experimental data. Specialist literature explores this field in depth for metals[66] and polymers,[67] but most models do not fit comfortably within the remit of multiscale modelling. As an extension to the creep topic, fatigue is probably the single most important mechanical property that cannot be modelled within the multiscale arena, and the author does not see much hope of significant advances in the immediate future. The author has suggested for polymers that creep and physical ageing can be related directly to the Debye frequency by the combination of mechanical and thermal energy effects in the activation function for the time- and rate-dependence of the glass transition, which may be a guide for application to other material types if this elastic instability condition can be applied to dislocations.[40]

9.7 Future trends

The need for improved predictive modelling of structural materials increases as the combination of properties required for application performance becomes more demanding. The commercial benefits of reduced costs and risks and improved effectiveness and efficiency on the technical side need to be combined with the use of modelling to explore new opportunities for material compositions and multiscale structures.[68]

This chapter has already pointed to some major weaknesses in modelling structural materials: dislocation nucleation and dynamics and bridging between time scales, for example. Also, ways to tackle these problems have been suggested; by greater emphasis on energy dissipation mechanisms than purely elastic processes. However, many more general aspects need to be addressed at the different scales. On a practical level, experiments need to be formulated to validate simulations directly with as few uncertain factors as possible.

At the atomic level there is perhaps too much complacency on the methods after a period of considerable improvements, particularly in the quantum methods based on density-functional theory. Integration of dissipation mechanisms and thermal effects is a key requirement for structural properties. The author firmly believes that analytical methods based upon mean-field approximations, using energy terms derived from relatively small-scale quantum simulations, would be a useful development and make practical

simulations possible for even the smaller and application-oriented development groups, rather than just major government laboratories: see Section 9.4.3 for an example of this approach for polymers. One particular effect that stems from the polymer approach in Section 9.4.3 is the need for a way to quantify the state of matter (crystal, glass, rubber, liquid, etc.) as an atomic scale parameter in calculations at higher scales; perhaps simply as distinct zero point energies with well-defined values within the potential energy well.

At the meso scale around the micron level, there needs to be a move away from brute-force methods, although they will still be pursued at the national level as a yardstick from which more refined techniques can evolve. The physics of coupling between scales depends critically on this scale, and not just concurrent approaches to modelling strategy. In particular, the link between time and distance scales is dominated by the evolution of defects and dislocations in structural materials. It is at this scale that cooperation between the different disciplines of chemistry, physics, and engineering needs to search for the simplest possible links between chemical composition and evolution of structural morphology during deformation. The author has suggested some simple approaches to this problem in Section 9.3.3, and it is here where the smaller application-oriented laboratories may make the greatest contribution.

The macro scale is superficially the strongest area for structural materials, due to extensive development motivated by commercial use in engineering design and problem solving. The coupling physics at the lower scales into macro simulations such as finite-element methods may be the greatest practical limiting factor for structural modelling.

Overall, the main suggestion of this article is to stand back from the massive supercomputer simulation approaches and think about properties of structural materials at a much more basic level. Perhaps the key question should be 'how can I calculate the property the engineer needs using the minimum number of parameters and smallest possible simulations?'

This work was carried out as part of the Weapons and Platform Effectors Domain of the MoD Research Programme.

9.8 References

1. Griffith A.A., *Philos. Trans. R. Soc. London A* **221**, 163–168 (1920).
2. Kassner M.E. *et al.*, 'New directions in mechanics', *Mech. Mater.* **37**, 231–259 (2005).
3. Soutis C. and Beaumont P.W.R., eds, *Multi-scale Modelling of Composite Material Systems: The Art of Predictive Damage Modelling*, Cambridge: Woodhead, (2005).
4. Ashby M.F., *Materials Selection in Mechanical Design*, Oxford: Pergamon, (1992).
5. *Materials Research to Meet 21st Century Needs*, DoD report available on http://www.nap.edu/execsumm/0309087007.html
6. Westmoreland P. R., *International Comparative Study on Applying Molecular and Materials Modelling*, http://www.wtec.org/loyola/molmodel/

7. Shercliff H. R., *Predictive Modelling of Materials in the UK: A review of activity and needs in the UK in response to the Technology Foresight exercise* http://www-materials.eng.cam.ac.uk/hrs/publications/modelling-report.html
8. Diaz de la Rubia T. and Bulatov V.V. (Guest Editors), *MRS Bull.* **26**, March (2001).
9. Ghoniem N.M. and Cho K., *Comput. Model. Eng. Sci.* **3**, 147 (2002).
10. Ghoniem N.M., Busso E.P., Kioussis N. and Huang H., *Philos. Mag.* **83**, 3475 (2003).
11. Lu G. and Kaxiras E., in *Handbook of Theoretical and Computational Nanotechnology*, eds. M. Rieth and W. Schommers, Stevenson Ranch, C.A.: American Scientific Publisher, Vol. X, I, (2005).
12. Kotelyanskii M.J. and Theodorou D., eds., *Simulation Methods for Polymers*, New York: Marcel Dekker, (2004).
13. Bicerano J., ed., *Computational Modelling of Polymers*, New York: Marcel Dekker, (1992).
14. Information on CASTEP can be found on the supplier website http://www.accelrys.com
15. Rose J.H., Smith J.R., Guinea F. and Ferrante J., *Phys. Rev. B* **29**, 2963–2969 (1984).
16. Porter D. and Gould P. J., *J. Phys. IV France* **110**, 809–814 (2003).
17. Anderson O.L., *J. Phys. Chem. Solids* **24**, 909–917 (1963).
18. Kittel C., *Introduction to Solid State Physics*, 5th Edn, Chapter 5, New York: John Wiley and Sons, (1975).
19. Mayer B., Anton H., Bott E., Meth Fessel M., Sticht J., Harris J. and Schmidt P.C., *Intermetallics* **11**, 23–32 (2003).
20. Brenner S.S., *J. Appl. Phys.* **27**, 1484–1956 (1956).
21. Roundy D., Krenn C.R., Cohen M.L. and Morris J.W., *Phys. Rev. Lett.* **82**, 2713–2716 (1999).
22. Porter D., *Int. Mater. Rev.* **47**, 225–232 (2002).
23. Born M., *J. Chem. Phys.* **7**, 591–601 (1939). Also discussed in Jin Z-H, Gumbsch P., Lu K., Ma E., *Phys. Rev. Lett.* **87**, 055703–055703-4 (2001).
24. Guinea F., Rose J.H., Smith J.R. and Ferrante J., *Appl. Phys. Lett.* **44**, 53–60 (1984).
25. Wang J., Li J., Yip S., Wolf D. and Phillpot S., *Physica A* **240**, 396–403 (1997).
26. http://www.thermocalc.com/
27. http://www.calphad.org/
28. http://www.thermotech.co.uk/jmatpro.html
29. Cottrell A.H., *Dislocations and Plastic Flow in Crystals*, London: Oxford University Press, (1953).
30. Read W.T., *Dislocations in Crystals*, New York and Maidenhead: McGraw-Hill, (1953).
31. Honeycombe R.W.K., *The Plastic Deformation of Metals*, London: Edward Arnold, (1968).
32. Hellmans A., *Science* **281**, 943–944 (1998).
33. Information on ParaDis: http://www-cms.llnl.gov/s-t/dislocation_dynamics.html
34. van Vliet K.J., Li J., Yip S. and Suresh S., *Phys. Rev. B* **67**, 104105-1–104105-15 (2003).
35. Miller R.E. and Acharya A., *J. Mech. Phys. Solids* **52** 1507–1525 (2004).
36. Cottrell A.H., *Trans. AIME*, **212**, 192–203 (1958).
37. Wang Y., Chen M., Zhou F. and Ma E., *Nature* **419**, 912–915 (2002).
38. Jiang B. and Weng G.J., *Metal. Mater. Trans. A* **34A**, 765–772 (2003).
39. Schiøtz J. and Jacobsen K.W., *Science* **301**, 1357–1359 (2003).
40. Porter D., *Group Interaction Modelling of Polymer Properties*, New York: Marcel Dekker, (1995).

41. Fantozzi G., Esnouf C., Benoit W., and Ritchie I.G., *Prog. Mater. Sci.* **27**, 311–451 (1982).
42. Meier J.J., *Macromolecules* **26**, 4376–4378 (1993).
43. Ward I.M., *Mechanical Properties of Solid Polymers*, Chichester: Wiley-Interscience, (1971).
44. Mackerle J., *Model. Simul. Mater. Sci. Eng*. **11**, 195–231 (2003).
45. van Krevelen D. W., *Properties of Polymers*, 3rd Edn, Amsterdam: Elsevier, (1990).
46. Bicerano J., *Prediction of Polymer Properties*, New York: Marcel Dekker, (1993).
47. Wunderlich B., Cheng S.Z.D. and Loufakis K., Thermodynamic properties, in *Encyclopedia of Polymer Science and Engineering*, vol. 16, New York: Wiley Interscience, 767–807, (1989).
48. Meier D.J., ed., *Molecular Basis of Transitions and Relaxations*, New York: Gordon and Breach Science Publishers, (1978).
49. Skylaris C.-K., Haynes P.D., Mostofi A.A., Payne M.C., *J. Chem. Phys*., **122**, 084119-1–084119-10 (2005).
50. Carter W.J. and Marsh S.P., *Hugoniot Equation of State for Polymers*, Los Alamos report LA-13006-A, (1997).
51. Panagiotopoulos Z., Gibbs ensemble and histogram reweighting gran canonical Monte Carlo methods, in *Simulation Methods for Polymers*, eds Kotelyanskii M.J. and Theodorou D., New York: Marcel Dekker, 2004, Chapter 9.
52. Information on COMPASS can be found on the supplier website http://www.accelrys.com
53. Baltá-Calleja F.J. and Michler G.H., eds., *Mechanical Properties of Polymers Based on Nanostructure and Morphology*, Boca Raton: FL, CRC Press, (2005).
54. Odegard G.M., Clancy T.C. and Gates T.S., *Polymer* **46**, 553–562 (2005).
55. Mori T. and Tanaka K., *Acta Metall.* **21**, 571–574 (1973).
56. Porter D., *Mater. Sci. Eng.* **A365**, 38–45 (2004).
57. Kinloch A.J. and Young R.J., *Fracture Behaviour of Polymers*, London: Elsevier Applied Science Publishers, (1973).
58. Marsh G., *Materials Today*, October, 32–36 (2002).
59. Porter D., Vollrath F. and Shau Z., *Eur. Phys. J. E* **16**, 199–206 (2005) and Vollrath F. and Porter D., *Appl. Phys. A* **83**, 205–212 (2006).
60. Kuntz J.D., Zhan G.-D. and Mukherjee A.K., *MRS Bull.* **29**, 22–28 (2004).
61. Barnett S. and Madan A., *Physics World*, January, 45–48 (1998).
62. Kingery W.D., Bowen H.K. and Uhlmann D.R., *Introduction to Ceramics*, New York: John Wiley, (1976).
63. Feltz A., *Amorphous Inorganic Materials and Glasses*, Weinhein and New York: VCH, (1993).
64. Mao H., Sundman B. and Wang Z.W., *J. Alloys Comp.* **327**, 253–262 (2001).
65. Horstemeyer M.F., Baskes M.I. and Plimpton S.J., *Acta Mater.*, **49**, 4363–4374 (2001).
66. Kassner M. and Perez-Prado M.-T., *Fundamentals of Creep in Metals and Alloys*, Amsterdam: Elsevier, (2004).
67. Findley W.N., Lai J.S. and Onaran K., *Creep and Relaxation of Nonlinear Viscoelastic Materials*, New York: Dover, (1989).
68. Stoneham M., *Materials Modelling – Growth and Innovation Prospects 2005*, DTI Innovation and Growth Team report, September (2005).

Index

ab initio electronic structure calculations 1–20
  basic equations 2–3
  density functional theory (DFT) 3–5, 189, 240–1
  interatomic potentials generation 12–15
  and the magnetism of iron 7–10
  for perfect solids 5–7
  positron annihilation studies 15–19
  and theoretical tensile strength in metals 10–12
ADEPT simulation 156, 179
Allen–Cahn equation 74
aluminium films 48
artificial fields 64
as-drawn circuit structures 149
asymptotic homogenization and finite elements 137–9
atomic configuration 12–15
atomic superposition method 16–17
atomistic simulations 86, 94–7, 178–80
automated serial sectioning 124
automatic adaption 198–201
axial loading simulation 227–39
  axial compression 234–7
  axial tension 229–34

Bain's path 12
ballistic transport and reaction model (BTRM) 166, 168, 177–8
Bauschinger effect 50
Beer–Lambert law 125
Bloch theorem 5–6
Boltzmann transport equation (BTE) 178
bond-order potentials (BOPs) 14–15
boundary conditions in computational micromechanics 133–5
Boussinesq–Cerruti Green's function formalism 39
Bravais lattice 196
Brillouin zone (BZ) 5
Brown's formula 27–8
BTE (Boltzmann transport equation) 178
BTRM (ballistic transport and reaction model) 166, 168, 177–8
bulk plasticity 44–7
Burgers vector 205, 207

CADD *see* coupled atomistic and discrete dislocation (CADD)
Cahn–Hilliard equations 74, 159
CALPHAD approach 68–9, 78, 268
carbon nanostructure modelling 220–58
  axial loading simulation 227–39
    axial compression 234–7
    axial tension 229–34
  concurrent TB/MD overlap coupling 225–7
  density of point 226–7
  dynamics of carbon nanotubes 220–1
  hydrogen interactions 239–41, 249–58
  minimization schemes 238–9
  molecular dynamic (MD) method 220–7
  nearest distance between carbon atoms 226–7
  ONIOM method 240, 241–9
  overlap TB/MD multiscale model 221–7

switch function effects 227–9
tight-binding (TB) method 220–7
CASTEP method 265–8
ceramics 262, 281–2
chemical vapor deposition (CVD) 156, 165–6
chemisorption mechanisms 239–40, 249–58
composition fields 64
computational micromechanics 1–2, 131–6
boundary conditions 133–5
constitutive equations and damage 135–6
domain decomposition method 133
numerical strategies 131–3
see also finite element and homogenization modelling
concurrent TB/MD overlap coupling 225–7
condensed matter theory 1
conformal voxel extraction (CVE) method 177
constitutive equations and damage 135–6
continuum micromechanics 121, 143–4
continuum region 195–8
continuum representations 157–9
copper films 48
copper-based MLM structures 152–3
core models 25–6
cosine re-emission 166
Coulomb interaction 5, 16
coupled atomistic and discrete dislocation (CADD) 192, 202–13
CADD formulation 202–5
detecting dislocations 205–8
passing dislocations 208–10
coupled atomistic/continuum modelling 189–215
automatic adaption 198–201
continuum region 195–8
embedded atom method (EAM) 190
exact atomistic problem 194–5
pad atoms 204
quasicontinuum (QC) method 192, 193–202, 210–12
representative atoms 195
critical RVE 122–4
cross-slip 54–5
crystalline solids, plastic deformation 25
cubic RVE 134–5
curvature-driven grain boundary migration 172
CVD (chemical vapor deposition) 156, 165–6
CVE (conformal voxel extraction) method 177

DDD simulations *see* dislocation behaviour modelling
deformation gradient 196, 200–1
density functional theory (DFT) 3–5, 189, 240–1
density of point 226–7
deposition model 165–7
detecting dislocations 205–8
diffuse re-emission 166
diffusion-accommodated grain rotation 97–8
dimensional analysis of grain growth 101–3
discrete dislocation dynamics (DDD) 25–7
discrete representations 155–6
discretization and motion 28–30
dislocation behaviour modelling 25–57
approximations and rules 34–7
Brown's formula 27–8
and bulk plasticity 44–7
core models 25–6
coupled atomistic and discrete dislocation (CADD) 192, 202–13
discretization and motion 28–30
elastic force over unit length 26
extensions of 54–7
and film plasticity 48–51
finite-element method 28
history of 27–8
local line velocity 29
and microelectronic devices 51–4
non-local and image fields 37–40
parametric method 28
Peach–Koehler formalism 27–8
periodic boundary conditions 40–4
regularization of the self-interaction 31–4
standard core model 33

dislocation nucleation 56
dislocations and failure in metal structures 269–71
domain decomposition method 133
dynamics of carbon nanotubes 220–1

EAM (embedded atom method) 190
elastic force over unit length 26
elastic instability 267–8
electroless deposition (ELD) model 163
electromigration 172
electronic structure calculations *see ab initio* electronic structure calculations
embedded atom method (EAM) 190
embedded cell methods 141–3
encapsulation procedures 160
energy definition (ONIOM method) 245–6
Eshelby method 127–8
Euler differencing 30
Euler phase-field crystal model 78
EVOLVE 157, 167, 170
exact atomistic problem 194–5

FACET 159
field variables 63–5
film plasticity 48–51
finite element and homogenization modelling 28, 121–44
  asymptotic homogenization and finite elements 137–9
  computational micromechanics 1–2, 131–6
  future directions 143–4
  and grain growth 116
  homogenization models 125–31
  microstructure generation and representation 124–5
  multiscale coupling 136–43
  representative volume element (RVE) 121, 122–5
  Voronoi Cell finite element method 132
free energy models 66–71
FVIEW 169, 170

generalized gradient approximation (GGA) 5
geometry optimization 246–7
ghost forces 191
Ginzberg–Landau equations 159
glue models 190
gradient energy 71–2
grain boundary migration and grain rotation 108–12
grain evolution 172–4
grain formation 167–70
grain growth 84–117
  atomic-level simulations 86, 94–7
  diffusion-accommodated grain rotation 97–8
  dimensional analysis 101–3
  finite-element (FEM) method 116
  grain boundary migration and grain rotation 108–12
  Herring relation 85, 86, 110
  length analysis 101–3
  mechanisms of grain growth 88–93
  mesoscale representations 98–9
  mesoscale simulation 85–6, 97–103
    results 106–14
    validation 103–4
  molecular dynamics simulation 87–97
  rotation-coalescence mechanism 91, 106–7
  and stress distribution 116
  T1 events 93–4
  T2 events 93–4
  timescales 101–3
  topological discontinuities 93–4
  variational function approach 99–101
  von Neumann–Mullins relation (VNM) 112–14
grain-continuum modelling 148–80
  atomistic simulators 178–80
  continuum representations 157–9
  copper-based MLM structures 152–3
  deposition model 165–7
  discrete and continuum 160
  discrete representations 155–6
  grain evolution 172–4
  grain formation 167–70
  kinetic lattice Monte Carlo (KLMC) 155–6
  molecular dynamic (MD) methods 155–6

multiple level set (MLS) representation 161–2, 175
and multiscale modelling 148
PLENTE software 162–5, 167, 168–9, 178–80
porous dialectrics 170–2
quasi-continuum approach 160
virtual wafer fab 148–51
voxel extraction (VE) method 177
grain-continuum representation 159–60
grain-focused modelling 151–2
Greed function 6

Hall–Petch effect 50
Hamiltonian equation 2, 3–4, 6
Herring relation 85, 86, 110
hierarchical approaches 190
homogenization models 125–31
asymptotic homogenization and finite elements 137–9
basic equations 126
linear behaviour 127–9
non-linear behaviour 130–1
hydrogen interactions with carbon nanostructures 239–41, 249–58
ONIOM calculations 254–8
weakening of C-C bond 249–54

image fields 37–70
integrated circuit (IC) performance 148–51
interatomic potentials generation 12–15
interfacial energy 71–2
inverse rule of mixtures 127
iron magnetism 7–10
isotrain approach 127
isotress approach 127

kinetic lattice Monte Carlo (KLMC) 155–6
Kohn–Sham equation 2, 3, 6
Korringa–Kohn–Rostoker (KKR) method 6

Landau–Lifshits–Gilbert equation 74
lattice statics 194
Lawrence Livermore National Laboratories 46
length of grain analysis 101–3
level-set methods 159–60
link atoms 242–5
local bulk chemical free energy density 66–71
local line velocity 29
local-density approximation (LDA) 4–5

macroscopic behaviour of materials 121
magnetism of iron and iron overlayers 7–10
MD (molecular dynamic) methods 87–97, 155–6, 220–7
mechanisms of grain growth 88–93
MEMS (micro-electromechanical structures) 38, 53
mesoscale representations 98–9
mesoscale simulation 85–6, 97–103
results 106–14
validation 103–4
mesh superposition methods 139–41
mesh-spectral model 78–9
Message Passing Interface (MPI) 164
metal structural materials 262, 264–73
bridging of scales 271–3
CASTEP method 265–8
dislocations and failure 269–71
elastic instability 267–8
intrinsic properties 266–8
microscale defects 269–71
potential wells 265–6
pressure–volume–temperature (PVT) equation 265–7
thermal instability 268
micro-electromechanical structures (MEMS) 38, 53
microelectronic devices 51–4
microstructures
evolution equations 62, 74–5
generation and representation 63–5, 124–5
multiscale models 77
thermodynamics 65–74
minimization schemes 238–9
molecular dynamic (MD) methods 87–97, 155–6, 220–7
molybdenum 14–15
Moore's Law 149
Mori–Tanaka method 128
multilevel metallization 149

multiple level set (MLS) representation 161–2, 175
multiscale coupling 136–43
  asymptotic homogenization and finite elements 137–9
  embedded cell methods 141–3
  mesh superposition methods 139–41

nanocrystalline materials 15–19
nanoscale plasticity 189
Navier equations 140
Navier–Stokes equation 103
non-local and image fields 37–40
non-locality criterion 200
numerical strategies in computational micromechanics 131–3

ONETEP model 274–5
ONIOM method 240, 241–9
  and chemisorption of hydrogen atoms 254–9
  energy definition 245–6
  geometry optimization 246–7
  link atoms 242–5
  second derivatives 247–9
  theoretical development 241–2
order-disorder transformations 64
overlap TB/MD multiscale model 221–7

pad atoms 204
parametric method 28
passing dislocations 208–10
Peach–Koehler formalism 27–8
perfect solids 5–7
periodic boundary conditions 40–4
periodic RVE 134–5
phase-field modelling 62–79, 159
  advantages and disadvantages 76–7
  artificial fields 64
  CALPHAD approach 68–9, 78
  composition fields 64
  developments and opportunities 77–9
  Euler phase-field crystal model 78
  field variables 63–5
  and gradient energy 71–2
  input parameters 75
  and interfacial energy 71–2
  local bulk chemical free energy density 66–71
  mesh-spectral model 78–9
  microstructure evolution 62
    equations 74–5
    multiscale models 77
  microstructure representation 63–5
  microstructure thermodynamics 65–74
  non-local interactions 72–4
  numerical solutions 75–6
  physical fields 64
  and solid-liquid interfaces 77
  three-dimensional simulations 78
physical fields 64
physical vapor deposition (PVD) 156, 165–6
physisorption mechanisms 239–40
plasticity
  bulk plasticity 44–7
  film plasticity 48–51
  nanoscale plasticity 189
  plastic deformation 25
  strain gradient plasticity theories 135–6
PLENTE software 162–5, 167, 168–9, 178–80
polymer structural materials 262, 273–81
  modelling framework 278–81
  ONETEP model 274–5
  spatial scales 274–8
  viscoelastic properties of polymers 280
porous dialectrics 170–2
positron annihilation studies 15–19
  atomic superposition method 16–17
positron lifetime spectroscopy (PLS) 15–16
potential wells 265–6
Potts models 155
pressure–volume–temperature (PVT) equation 265–7
prismatic RVE 134
PVD (physical vapor deposition) 156, 165–6

quantum mechanics/molecular mechanics (QM/MM) model 225
quasicontinuum (QC) method 160, 192, 193–202, 210–12

random re-emission 166

reaction-rate theory 97
reactive empirical bond-order (REBO) 220
Read–Schockley equation 94
regularization of the self-interaction 31–4
representative atoms 195
representative volume element (RVE) 121, 122–5
  boundary conditions 133–5
  critical RVE 122–4
  cubic RVE 134–5
  definition 122
  microstructure generation and representation 124–5
  periodic RVE 134–5
  prismatic RVE 134
resistancing schemes 162
Reuss model 127, 129
Reynolds number 103
roadmapping semiconductors 149
rotation–coalescence mechanism 91, 106–7
rule of mixtures 127

Schrödinger equation 2, 3, 16–17, 223
secant formulations 130
self-interaction corrected (SIC) methods 5
semiconductor roadmap 149
Shockley partials 198
solid–liquid interfaces 77
spatial scales 274–8
standard core model 33
strain concentration tensors 128
strain gradient plasticity theories 135–6
streamline upwinding Petrov-Galerkin (SUPG) methods 162
stress distribution and grain growth 116
structural materials modelling 261–85
  ceramics 262, 281–2
  concurrent approach 264
  future trends 284–5
  metals 262, 264–73
  polymers 262, 273–81
  properties and scales 262–3
  sequential approach 264
  time effects 282–4
switch function effects 227–9

T1 events 93–4
T2 events 93–4
Technology Roadmap for Semiconductors 149
tensile strength in metals 10–12
Tersoff–Brenner potential 220
thermal instability 268
ThermoCalc 268
three-dimensional simulations 78
tight-binding (TB) method 220–7
time effects on structural materials 282–4
timescales and grain growth 101–3
topological discontinuities 93–4
tungsten 11

uniform deformation 200
Universal Force Field (UFF) model 241

variational function approach 99–101
virtual atom cluster (VAC) 225
virtual wafer fab 148–51
viscoelastic properties of polymers 280
viscous force law 97, 103
Voigt model 127, 129
von Neumann–Mullins relation (VNM) 112–14
Voronoi Cell finite element method 132
Voronoi construction 88, 99
voxel extraction (VE) method 177

weighted-density approximation (WDA) 5

X-ray computed microtomography 124–5