This book presents our current knowledge and understanding of continuum-based concepts behind computational methods used for microstructure and process simulation of engineering materials above the atomic scale.

Divided into three main parts, the volume provides an excellent overview on the different methods, comparing the different methods in terms of their respective particular weaknesses and advantages. This trains readers to identify appropriate approaches to the new challenges that emerge every day in this exciting domain. The first part is a basic overview covering fundamental key methods in the field of continuum scale materials simulation. The second one then goes on to look at applications of these methods to the prediction of microstructures, dealing with explicit simulation examples, while the third part discusses example applications in the field of process simulation.

By presenting a spectrum of different computational approaches to materials, the book aims to initiate the development of corresponding virtual laboratories in the industry in which these methods are exploited. As such, it addresses graduates and undergraduates, lecturers, materials scientists and engineers, physicists, biologists, chemists, mathematicians, and mechanical engineers.
Dr. Frédéric Barlat received a PhD in Mechanics from the "Institut National Polytechnique de Grenoble," France, in 1984. The same year, he joined Alcoa Technical Center; Pittsburgh, Pennsylvania, USA, the research facility of Alcoa Inc. (formerly the Aluminum Company of America). Dr. Barlat is currently a Technology Specialist in their Materials Science Division. He is responsible for conceptualizing, importing and implementing mathematical models that predict the mechanical behavior of materials for long-term development applications in the areas of metal plasticity, fracture and material performance. His work is used for the design of alloys and processes in support of Alcoa’s major business units, including packaging, automotive and aerospace. Dr. Barlat is also an invited professor at the University of Aveiro’s Center for Mechanical Technology and Automation, Portugal, where he directs activities on the fundamentals of plasticity and forming. He has actively participated in the scientific committees of various international conferences, has been a regular reviewer in a number of scientific journals and serves as a member of the Advisory Board of the International Journal of Plasticity. Dr. Barlat is published as an author or co-author in approximately 80 papers of peer-reviewed scientific journals and has delivered more than 60 technical presentations at conferences worldwide. In 1995, he was the honored recipient of the ASM Henry Marion Howe Medal of the Material Society for the best technical paper published in Metallurgical Transactions A. He holds three US patents with co-inventors from Alcoa Inc. and Kobe Steel, Ltd., Japan.

Professor Dierk Raabe received his Ph.D. (1992) and habilitation (1997) at RWTH Aachen, Germany, in the fields of Physical Metallurgy and Metal Physics. He is currently Director and Executive at the Max-Planck Institut für Eisenforschung, Düsseldorf, Germany, after working some time as researcher at Carnegie Mellon University, USA, the High Magnetic Field Laboratory in Tallahassee, USA, and serving as Senior researcher and lecturer at the Institut für Metallkunde und Metalphysik, RWTH Aachen, Germany. His research fields are Computer simulation of Materials, Composites, Textures, Micromechanics, in which he authored more than 100 papers in peer-reviewed magazines and three books. He teaches various courses on computational materials science, materials mechanics, history of metals, and textures at RWTH Aachen (Germany) and at Carnegie Mellon University Pittsburgh (USA). His work was already awarded with several prizes, among them the Adolf-Martens Award, Masing Award, Heisenberg Award, and the Leibniz Award.

Dr. Franz Roters studied Physics in Braunschweig, where he got his diploma degree in 1993. From 1994 to 1998 he was scientist at the Institute for Metal Physics and Physical Metallurgy at the RWTH Aachen. He got his PhD. degree in 1999 in the field of constitutive modelling of aluminium. From 1999 till 2000 he was researcher at the R&D centre of VAW (today Hydro Aluminium Deutschland GmbH) in Bonn. Since 2000 he is senior scientist at the Max-Planck-Institut für Eisenforschung in Düsseldorf, where he is the leader of the research group "Theory and Simulation" in the department for Microstructure Physics and Metal Forming. Dr. Roters published more than 50 papers in the field of constitutive modelling and simulation of forming. He is head of the Technical Committee "Computersimulation" of the Deutsche Gesellschaft für Materialkunde e.V. (DGM).

Long-Qing Chen is Professor of Materials Science and Engineering at Penn State. He received his B.S. in Ceramics from Zhejiang University in China in 1982, a M.S. in Materials Science and Engineering from State University of New York at Stony Brook in 1985, and a Ph.D. degree in Materials Science and Engineering from MIT in 1990. He worked with Armen G. Khachaturyan as a postdoc at Rutgers University from 1990 to 1992. Professor Chen joined the Department of Materials Science and Engineering at Penn State as an assistant professor in 1992 and was promoted to associate professor in 1998. His main research interests include materials theory and computational materials science. Professor Chen received the Young Investigator Award from the Office of Naval Research (ONR) in 1993, the research creativity award from the National Science Foundation (NSF) in 1999, the Wilson Award for Excellence in Research in the College of Earth and Mineral Sciences in 2000, and the University Faculty Scholar Medal at Penn State in 2003.

Dr. Frédéric Barlat received a PhD in Mechanics from the "Institut National Polytechnique de Grenoble," France, in 1984. The same year, he joined Alcoa Technical Center; Pittsburgh, Pennsylvania, USA, the research facility of Alcoa Inc. (formerly the Aluminum Company of America). Dr. Barlat is currently a Technology Specialist in their Materials Science Division. He is responsible for conceptualizing, importing and implementing mathematical models that predict the mechanical behavior of materials for long-term development applications in the areas of metal plasticity, fracture and material performance. His work is used for the design of alloys and processes in support of Alcoa’s major business units, including packaging, automotive and aerospace. Dr. Barlat is also an invited professor at the University of Aveiro’s Center for Mechanical Technology and Automation, Portugal, where he directs activities on the fundamentals of plasticity and forming. He has actively participated in the scientific committees of various international conferences, has been a regular reviewer in a number of scientific journals and serves as a member of the Advisory Board of the International Journal of Plasticity. Dr. Barlat is published as an author or co-author in approximately 80 papers of peer-reviewed scientific journals and has delivered more than 60 technical presentations at conferences worldwide. In 1995, he was the honored recipient of the ASM Henry Marion Howe Medal of the Material Society for the best technical paper published in Metallurgical Transactions A. He holds three US patents with co-inventors from Alcoa Inc. and Kobe Steel, Ltd., Japan.

The Editors
Contents

Preface ................................................................. XXI
List of Contributors .............................................. XXIII

I Fundamentals and Basic Methods ................................ 1

1 Computer Simulation of Diffusion Controlled Phase Transformations
   (A. Schneider and G. Inden) .......................................... 3
   1.1 Introduction ..................................................... 3
   1.2 Numerical Treatment of Diffusion Controlled Transformations . 5
       1.2.1 Diffusion .................................................. 5
       1.2.2 Boundary Conditions ..................................... 8
       1.2.3 Cell Size .................................................. 14
   1.3 Typical Applications ........................................... 15
       1.3.1 LE, LENP and PE in Fe-Mn-C ............................ 15
       1.3.2 LE, LENP and PE in Fe-Si-C ............................ 17
       1.3.3 PE in Fe-Ni-C .............................................. 20
       1.3.4 Effect of Traces on the Growth of Grain Boundary Cementite . 21
       1.3.5 Continuous Cooling ....................................... 22
       1.3.6 Competitive Growth of Phases: Multi-Cell Calculations . 23
       1.3.7 Gas-Metal-Reactions: Carburization .................... 26
   1.4 Outlook ......................................................... 33

2 Introduction to Phase-field Method of Microstructure Evolution
   (L.-Q. Chen) ......................................................... 37
   2.1 Introduction ..................................................... 37
   2.2 Origin of the Model ............................................ 38
   2.3 Theoretical Fundamentals of the Method ....................... 38
       2.3.1 Representation of a Microstructure ....................... 38
       2.3.2 Thermodynamics of Microstructures ...................... 40
       2.3.3 The Evolution Equations ................................ 46
   2.4 Advantages and Disadvantages of the Method .................. 47
   2.5 Typical Fields of Applications and Examples .................. 47

References ............................................................. 34
## Contents

2.6 Summary and Opportunities ............................................. 48
References ........................................................................... 50

3 Cellular, Lattice Gas, and Boltzmann Automata
(D. Raabe) ........................................................................ 56
3.1 Cellular Automata ............................................................... 56
  3.1.1 Introduction ................................................................. 56
  3.1.2 Formal Description and Classes of Cellular Automata ............ 57
  3.1.3 Cellular Automata in Materials Science ............................ 59
  3.1.4 Recrystallization Simulations with Cellular Automata .......... 60
3.2 Cellular Automata for Fluid Dynamics .................................... 66
  3.2.1 Introduction ................................................................. 66
  3.2.2 The HPP and FHP Lattice Gas Cellular Automata ............... 66
  3.2.3 The Lattice Boltzmann Automaton .................................... 69
3.3 Conclusions and Outlook .................................................... 72
References ........................................................................... 73

4 The Monte Carlo Method
(A. D. Rollett and P. Manohar) ................................................ 76
4.1 Introduction ....................................................................... 76
4.2 History of the Monte Carlo Method ...................................... 76
  4.2.1 Ising and Potts Models .................................................... 77
  4.2.2 Metropolis Algorithm .................................................... 79
  4.2.3 n-fold Way Algorithm .................................................... 80
4.3 Description of the Monte Carlo Method for Grain Growth & Recrystallization ............................................. 84
  4.3.1 Discretization of Microstructure ........................................ 84
  4.3.2 Evolution of the Microstructure ....................................... 85
  4.3.3 Inert Particles ............................................................... 86
  4.3.4 Lattices ........................................................................ 86
  4.3.5 Boundary Conditions ..................................................... 88
  4.3.6 Parallelization of the Monte Carlo Algorithm ....................... 88
4.4 Nucleation in Recrystallization ............................................. 91
4.5 Initialization of MC Simulations ........................................... 92
4.6 Verification of the Monte Carlo Model .................................... 93
4.7 Scaling of Simulated Grain Size to Physical Grain Size .............. 96
4.8 Recrystallization Kinetics in the Monte Carlo model ................. 97
4.9 Results of Simulation of Recrystallization by Monte Carlo Method ............................................. 98
  4.9.1 Abnormal Grain Growth ............................................... 98
  4.9.2 Static Recrystallization .................................................... 98
  4.9.3 Grain Growth in the Presence of Particles ......................... 100
  4.9.4 Recrystallization in the Presence of Particles ....................... 100
  4.9.5 Texture Development ..................................................... 102
  4.9.6 Texture ...................................................................... 104
  4.9.7 Dynamic Recrystallization .............................................. 108
4.10 Summary ........................................................................ 109
References ........................................................................... 110
5 Crystal Plasticity
(P.R. Dawson) 114
5.1 Introduction ................................................. 114
5.2 Theoretical Background ..................................... 114
  5.2.1 Mechanical Response of Single Crystals ................. 114
  5.2.2 Lattice Orientation Distributions for Polycrystals .... 119
  5.2.3 Mechanical Response of Polycrystals .................... 121
5.3 Macroscopic Criteria for Anisotropic Strength ............ 123
  5.3.1 Generalities ........................................... 123
  5.3.2 Yield Surfaces Defined by Expansions ................. 125
  5.3.3 Yield Surfaces Defined by Hyperplanes ............... 126
  5.3.4 Isoparametric Flow Surface .......................... 128
  5.3.5 Direct Polycrystal Plasticity Implementation .......... 130
5.4 Numerical Implementations ................................ 131
  5.4.1 Balance Laws ......................................... 131
  5.4.2 Finite Element Formulations .......................... 131
5.5 Applications .............................................. 133
  5.5.1 Application to Explosive Forming .................... 133
  5.5.2 Application to the Limiting Dome Height Test ....... 134
  5.5.3 Bending of a Curved Component ....................... 138
5.6 Summary .................................................. 138
References ..................................................... 140

6 Yield Surface Plasticity and Anisotropy
(F. Barlat, O. Cazacu, M. Życzkowski, D. Banabic, and J.W. Yoon) 143
6.1 Introduction ................................................. 143
6.2 Classical Plasticity Theory ................................. 144
  6.2.1 Isotropic Yield Conditions for Perfect Plasticity .... 144
  6.2.2 Flow Rules ............................................. 147
  6.2.3 Subsequent Yield Surfaces during Plastic Hardening ...... 148
  6.2.4 Anisotropic Plasticity ................................. 150
  6.2.5 Direct Generalizations of Isotropic Yield Conditions ... 151
6.3 Material Structure and Plastic Anisotropy .................. 152
  6.3.1 Texture ................................................. 152
  6.3.2 Dislocations ........................................... 154
  6.3.3 Porosity and Second Phases ............................ 158
6.4 Yield Functions for Metals and Alloys ...................... 159
  6.4.1 Quadratic Yield Functions ............................. 159
  6.4.2 Non-Quadratic Yield Functions ........................ 160
  6.4.3 Yield Functions in Polar Coordinates .................. 163
  6.4.4 Other Anisotropic Yield Functions ..................... 163
  6.4.5 BBC2000 Yield Criterion ............................. 163
  6.4.6 Yld2000-2d Yield Criterion ........................... 165
  6.4.7 CB2001 Yield Criterion ............................... 165
  6.4.8 Strain Rate Potentials ............................... 167
Contents

6.5 Application to Sheet Forming and Formability ................................................. 167
  6.5.1 Mechanical testing ................................................................. 167
  6.5.2 Analysis and Treatment of the Test Results ......................................... 170
  6.5.3 Application to 3103-O Aluminum Alloy Sheet Sample ................................. 172
  6.5.4 Plastic Flow Localization ......................................................... 172
  6.5.5 Cup Drawing Simulation ......................................................... 173

6.6 Conclusions ......................................................................................... 175

References .................................................................................................... 176

7 Artificial Neural Networks
  (E. Broese and H.-U. Löffler) ............................................................................. 182
  7.1 Introduction ............................................................................................. 182
  7.2 Basic Terms ............................................................................................. 183
  7.3 Fields of Application ............................................................................... 183
    7.3.1 Pattern Recognition/Classification ..................................................... 183
    7.3.2 Empirical Modeling ............................................................................ 183
  7.4 Implementation ......................................................................................... 184
    7.4.1 Software ............................................................................................ 184
    7.4.2 Hardware ........................................................................................... 185
  7.5 Types of Artificial Neural Networks .......................................................... 185
    7.5.1 Multilayer Perceptron ......................................................................... 185
    7.5.2 Radial Basis Function Networks ......................................................... 188
    7.5.3 More Network Types ......................................................................... 190
  7.6 Kinds of Learning ...................................................................................... 191
    7.6.1 Unsupervised Learning ..................................................................... 191
    7.6.2 Supervised Learning ......................................................................... 191
    7.6.3 Reinforcement Learning ................................................................... 191
    7.6.4 Bayesian Learning ............................................................................ 192
  7.7 Application Details .................................................................................... 192
    7.7.1 Network Type Selection and Configuration ....................................... 192
    7.7.2 Input Selection .................................................................................... 193
    7.7.3 Data Preprocessing and Input Scaling ................................................ 193
    7.7.4 Prevention of Overfitting .................................................................. 193
    7.7.5 Optimization of Training Parameters ............................................... 194
    7.7.6 Diagnostics of the Internal State ......................................................... 194
  7.8 Future Prospects ...................................................................................... 195

References .................................................................................................... 195

8 Multiscale Discrete Dislocation Dynamics Plasticity
  (H. M. Zbib, M. Hiratani, and M. Shehadeh) .................................................... 197
  8.1 Introduction ............................................................................................. 197
  8.2 Theoretical Fundamentals of the Method .................................................... 199
    8.2.1 Kinematics and Geometric Aspects .................................................... 199
    8.2.2 Kinetics and Interaction Forces ......................................................... 199
    8.2.3 Dislocation Equation of Motion ......................................................... 200
11.4 Elastic Energy of a Microstructure .................................. 271
11.5 Bulk Microstructures with Periodic Boundary Conditions ......... 271
11.6 A Single Crystal Film with Surface and Substrate Constraint ...... 273
11.7 Elastic Coupling of Structural Defects and Phase Transformations . 274
11.8 Phase-Field Model Applied to Solid State Phase Transformations ... 275
11.9 Isostructural Phase Separation ......................................... 275
11.10 Precipitation of Cubic Intermetallic Precipitates in a Cubic Matrix ... 277
11.11 Structural Transformations Resulting in a Point Group Symmetry Reduction 279
11.12 Ferroelectric Phase Transformations ................................ 281
11.13 Phase Transformation in a Reduced Dimensions: Thin Films and Surfaces 283
11.14 Summary .................................................................. 285
References .................................................................. 287

12 Irregular Cellular Automata Modeling of Grain Growth
(K. Janssens) ................................................................ 292
12.1 Introduction .................................................................. 292
12.2 Irregular Cellular Automata ............................................. 292
  12.2.1 The Concept ......................................................... 292
  12.2.2 Shapeless or Point Cellular Automata ......................... 293
12.3 Irregular Shapeless Cellular Automata for Grain Growth .......... 293
  12.3.1 Curvature Driven Grain Growth ................................. 294
  12.3.2 In the Presence of Additional Driving Forces ................. 297
12.4 A Qualitative Example: Static Annealing of a Cold Rolled Steel .... 299
  12.4.1 The Deformation Model ........................................... 299
  12.4.2 The Annealing Model ............................................. 300
12.5 Conclusion .................................................................. 302
References .................................................................. 302
### 13 Topological Relationships in 2D Trivalent Mosaics and Their Application to Normal Grain Growth

(R. Brandt, K. Lücke, G. Abbruzzese, and J. Svoboda)  
304

13.1 Introduction .................................................. 304
13.2 Individual Grains and their Distributions (One-Grain Model) .......... 307
  13.2.1 Definition of Parameters .................................. 307
  13.2.2 The Grain Sizes and Shapes and their Distributions ................. 308
  13.2.3 The Coordination and its Distributions .......................... 309
13.3 Topological Relationships of Trivalent Mosaics ............................ 309
  13.3.1 Grain Boundaries (GBs) and Triple Points (TPs) .................... 309
  13.3.2 The Geometry of the GB (Function $p_{ij}$) ...................... 310
  13.3.3 Size Correlations of Nearest Neighbor Grains (Function $k_{ij}$) ...... 310
  13.3.4 Space Filling (Function $q_{ij}$) ................................ 311
13.4 Cases of Randomness ............................................ 312
  13.4.1 “Abbruzzese–Lücke Equations” (ALE, Full Randomness) ............. 312
  13.4.2 “Weaire–Aboav Equation” (WAE, Partial Randomness) ............... 312
13.5 Curvature Driven GG ............................................ 314
  13.5.1 Direct Simulations ......................................... 314
  13.5.2 Simulations by the Statistical Theory ............................ 315
13.6 Summarizing Remarks ............................................ 318

References .......................................................... 320

### 14 Motion of Multiple Interfaces: Grain Growth and Coarsening

(B. Nestler)  
321

14.1 Introduction .................................................. 321
14.2 The Diffuse Interface Model .................................... 323
14.3 Free Energies .................................................. 325
14.4 Numerical Simulations ......................................... 327
  14.4.1 Grain Growth and Coarsening ................................ 328
  14.4.2 Multicomponent Multiphase Solidification ........................ 329
14.5 Outlook ......................................................... 334

References .......................................................... 335

### 15 Deformation and Recrystallization of Particle-containing Aluminum Alloys

(B. Radhakrishnan and G. Sarma)  
337

15.1 Background .................................................... 337
  15.1.1 Formation of Deformation Zones .............................. 338
  15.1.2 Formation and Growth of Particle Stimulated Nuclei ............... 339
15.2 Computational Approach ........................................ 342
15.3 Simulations .................................................... 344
15.4 Results and Discussion ........................................ 344
  15.4.1 Microstructure and Kinetics .................................. 344
  15.4.2 Texture .................................................... 349
15.5 Summary ....................................................... 351

References .......................................................... 353
16 Mesoscale Simulation of Grain Growth
(D. Kinderlehrer, J. Lee, I. Livshits, and S. Ta’asan) 355
16.1 Introduction ........................................ 355
16.2 Discretization ........................................ 359
16.3 Numerical Implementation ............................ 360
16.4 Numerical Results ................................... 363
16.5 Conclusion .......................................... 364
References .............................................. 365

17 Dislocation Dynamics Simulations of Particle Strengthening
(V. Mohles) 368
17.1 Introduction .......................................... 368
17.2 Simulation Method .................................... 370
  17.2.1 Basis of the Method ............................. 370
  17.2.2 Dislocation Segmentation ....................... 371
  17.2.3 Dislocation Self-Interaction .................... 372
  17.2.4 Simulation Procedure and Accuracy .......... 374
17.3 Particle Arrangement ................................ 375
17.4 Strengthening Mechanisms ........................... 377
  17.4.1 Dispersion Strengthening ...................... 377
  17.4.2 Order Strengthening ............................ 379
  17.4.3 Lattice Mismatch Strengthening .............. 381
17.5 Summary and Outlook ................................ 386
References .............................................. 386

18 Discrete Dislocation Dynamics Simulation of Thin Film Plasticity
(B. von Blanckenhagen and P. Gumbsch) 389
18.1 Thin Film Plasticity .................................. 389
18.2 Simulation of Dislocations in Thin Films ......... 391
  18.2.1 Boundary Conditions ............................ 392
18.3 Thin Film Deformation, Models and Simulation ... 394
  18.3.1 Mobility Controlled Deformation .............. 394
  18.3.2 Source Controlled Deformation ............... 395
References .............................................. 401

19 Discrete Dislocation Dynamics Simulation of Crack-Tip Plasticity
(A. Hartmaier and P. Gumbsch) 404
19.1 Introduction .......................................... 404
19.2 Model .................................................. 405
19.3 Crack-Tip Plasticity .................................. 410
19.4 Scaling Relations ..................................... 412
19.5 Discussion ............................................ 415
19.6 Conclusions ......................................... 416
References .............................................. 416
20 Coarse Graining of Dislocation Structure and Dynamics
(R. LeSar and J. M. Rickman) 419
20.1 Introduction ............................................. 419
20.2 Dynamics of Discrete Dislocations .......................... 420
  20.2.1 Dislocation Dynamics Methods .................. 420
  20.2.2 Phase-Field Methods ............................ 420
20.3 Static Coarse-Grained Properties .......................... 421
  20.3.1 Continuous Dislocation Theory .................. 422
  20.3.2 Extensions to the Continuous Theory ............... 425
20.4 Dynamic Coarse-Grained Properties .......................... 429
20.5 Conclusions .............................................. 431
References .................................................. 432

21 Statistical Dislocation Modeling
(R. Sedláˇcek) 435
21.1 Introduction ............................................. 435
21.2 One-parameter Models ..................................... 438
  21.2.1 Phenomenological Model .................. 438
  21.2.2 Materials Science Approach .................. 440
21.3 Multi-parameter Models ................................... 441
  21.3.1 Various Approaches .......................... 441
  21.3.2 Composite Framework ........................ 443
21.4 Conclusions .............................................. 445
References .................................................. 446

22 Taylor-Type Homogenization Methods for Texture and Anisotropy
(P. Van Houtte, S. Li, and O. Engler) 448
22.1 Introduction ............................................. 448
22.2 Local Constitutive Laws (Mesoscopic Scale) .............. 449
22.3 The Taylor Ambiguity .................................... 451
22.4 Full Constraints (FC) Taylor Theory ................... 452
22.5 Classical Relaxed Constraints (RC) Models .............. 453
22.6 Multi-grain RC Models ................................... 454
  22.6.1 Introduction .................................. 454
  22.6.2 The Lamel Model .................................. 455
  22.6.3 The Advanced Lamel Model .................. 456
  22.6.4 The Grain Interaction (GIA) Model ............... 458
22.7 Validation of the Models ................................ 458
22.8 Conclusions .............................................. 458
References .................................................. 460

23 Self Consistent Homogenization Methods for Texture and Anisotropy
(C.N. Tomé and R.A. Lebensohn) 462
23.1 Introduction ............................................. 462
23.2 Viscoplastic Selfconsistent Formalism ..................... 464
23.2.1 Local Constitutive Behavior and Homogenization ................. 464
23.2.2 Green Function Method and Fourier Transform Solution .......... 466
23.2.3 Viscoplastic Inclusion and Eshelby Tensors ....................... 467
23.2.4 Interaction and Localization Equations ............................ 469
23.2.5 Selfconsistent Equations ...................................... 470
23.2.6 Secant, Tangent and Intermediate Approximations ............... 471
23.2.7 Algorithm .............................................. 471
23.3 Implementation of a Texture Development Calculation .................. 472
23.3.1 Kinematics ............................................ 472
23.3.2 Hardening ............................................. 474
23.3.3 Twinning Reorientation .................................... 475
23.4 Applications ............................................. 476
23.4.1 Tension and Compression of FCC ............................... 476
23.4.2 Torsion (Shear) of FCC .................................... 477
23.4.3 Twinning and Anisotropy of HCP Zr ............................ 481
23.4.4 Compression of Olivine (MgSiO$_3$) ............................ 482
23.5 Further Selfconsistent Models and Applications .......................... 484
References .................................................. 486

24 Phase-field Extension of Crystal Plasticity with Application to Hardening
Modeling
(B. Svendsen) 489
24.1 Introduction ............................................... 489
24.2 Basic Considerations and Results ................................ 490
24.3 The Case of Small Deformation ................................ 494
24.4 Simple Shear of a Crystalline Strip ................................ 495
References .................................................. 498

25 Generalized Continuum Modelling of Single and Polycrystal Plasticity
(S. Forest) 500
25.1 Introduction ............................................. 500
25.1.1 Scope of this Chapter .................................... 500
25.1.2 Motivations for Generalized Continuum Crystal Plasticity ...... 501
25.2 Generalized Continuum Crystal Plasticity Models ..................... 502
25.2.1 Cosserat Single Crystal Plasticity ............................ 502
25.2.2 Second Gradient Single Crystal Plasticity ........................ 504
25.2.3 Gradient of Internal Variable Approach ........................ 505
25.3 From Single to Polycrystals: Homogenization of Generalized Continua 506
25.3.1 Introduction to Multiscale Asymptotic Method .................. 506
25.3.2 Extension of Classical Homogenization Schemes ................. 509
25.4 Simulations of Size Effects in Crystal Plasticity ..................... 510
25.4.1 Constrained Plasticity in Two-Phase Single Crystals .......... 510
25.4.2 Plasticity at the Crack Tip in Single Crystals ................. 510
25.4.3 Grain Size Effects in Polycrystalline Aggregates .............. 511
25.5 Conclusion ............................................... 513
References .................................................. 513
26 Micro-Mechanical Finite Element Models for Crystal Plasticity
(S. R. Kalidindi)  515
26.1 Introduction ........................................ 515
26.2 Theoretical Background ............................. 515
  26.2.1 Crystal Plasticity Framework ................. 516
  26.2.2 Total Lagrangian vs. Updated Lagrangian Schemes . 516
  26.2.3 Fully Implicit Time Integration Procedures .... 518
  26.2.4 Polycrystal Homogenization Theories .......... 520
26.3 Micro-Mechanical Finite Element Models .......... 520
26.4 Examples ...................................... 521
  26.4.1 Predictions of Deformation Textures .......... 521
  26.4.2 Predictions of Micro-Texture .................. 521
References ........................................ 528

27 A Crystal Plasticity Framework for Deformation Twinning
(S. R. Kalidindi)  529
27.1 Introduction ....................................... 529
  27.1.1 Slip vs. Deformation Twinning .............. 529
  27.1.2 Major Consequences of Deformation Twinning . 530
27.2 Historical Perspective ........................... 532
  27.2.1 Twin Reorientation Schemes .................. 533
  27.2.2 Volume Fraction Transfer Scheme ............ 533
27.3 Incorporation of Deformation Twinning ........... 534
  27.3.1 Relaxed Configuration ....................... 534
  27.3.2 Elastic Response ................................ 536
  27.3.3 Plastic Flow Rule ............................. 536
  27.3.4 Evolution of Twin Rotations ................. 538
  27.3.5 Slip-Twin Hardening Functions .............. 540
27.4 Examples ...................................... 542
References ........................................ 545

28 The Texture Component Crystal Plasticity Finite Element Method
(F. Roters)  547
28.1 Introduction ....................................... 547
28.2 The Texture Component Method ..................... 547
  28.2.1 Approximation of X-Ray Textures using Texture Components . 548
  28.2.2 Representation of Texture Components in a Crystal Plasticity FEM . 548
28.3 The Crystal Plasticity Model ....................... 551
28.4 Application of the TCCP-FEM to Forming Simulation . 552
  28.4.1 R-value Prediction ............................. 552
  28.4.2 Prediction of Earing Behavior ............... 552
28.5 Outlook ....................................... 557
References ........................................ 558
29 Microstructural Modeling of Multifunctional Material Properties:  
The OOF Project  
(R. Edwin García, Andrew C.E. Reid, Stephen A. Langer, and W. Craig Carter)  
29.1 Introduction ........................................ 559  
29.2 Program Overview .................................. 561  
29.3 Modeling of Piezoelectric Microstructures ............... 564  
29.4 Modeling of Electrochemical Solids: Rechargeable Lithium Ion Batteries  
29.5 The OOF2O Project: A Preview ....................... 571  
References .............................................. 573  

30 Micromechanical Simulation of Composites  
(S. Schmauder)  
30.1 Introduction ........................................ 574  
30.2 Matricity .......................................... 575  
30.2.1 Matricity Model .................................. 575  
30.2.2 Adjusting Matricity in the Model ................... 577  
30.2.3 Realisation of the Adjustability of Matricity by Weighting Factors .... 577  
30.2.4 Calculation of Stress-strain Curves ................ 578  
30.2.5 Mechanical Constants ............................. 579  
30.2.6 Yield Stress ...................................... 580  
30.3 Results and Discussion .............................. 580  
30.3.1 Comparison to Cluster Parameter r ................ 580  
30.4 Conclusion .......................................... 589  
References .............................................. 590  

31 Creep Simulation  
(W. Blum)  
31.1 Introduction ........................................ 592  
31.2 Empirical Relations .................................. 593  
31.3 Basic Dislocation Processes ........................... 594  
31.3.1 Homogeneous Glide Activity ....................... 594  
31.3.2 Heterogeneous Glide Activity ..................... 595  
31.4 Models ............................................. 596  
31.4.1 Two-parameter Model for Homogeneous Glide .......... 597  
31.4.2 Composite Model for Heterogeneous Glide ........... 599  
31.5 Concluding Remarks ................................ 601  
References .............................................. 603  

32 Computational Fracture Mechanics  
(W. Brocks)  
32.1 Introductory Remarks on Inelastic Material Behaviour .... 606  
32.2 FE Meshes for Structures with Crack-Like Defects .......... 608  
32.2.1 General Aspects and Examples ...................... 608  
32.2.2 Singular Elements for Stationary Cracks ............. 609  
32.2.3 Regular Element Arrangements for Extending Cracks .... 610
32.3 The J-Integral as Characteristic Parameter in Elasto-Plastic Fracture Mechanics ........................................... 611
32.3.1 Foundation ........................................ 611
32.3.2 The Domain Integral or VCE Method ..................... 612
32.3.3 Path Dependence of the J-Integral in Incremental Plasticity ...... 613
32.4 The Cohesive Model ........................................ 614
32.4.1 Fundamentals ........................................ 614
32.4.2 Example: Simulation of Ductile Tearing in a Laser Weld ............ 617
32.5 Summary .................................................. 618
References ................................................. 619

33 Rheology of Concentrated Suspensions: A Lattice Model (Y. Brechet, M. Perez, Z. Neda, J. C. Barbe, and L. Salvo) 623
33.1 Introduction .............................................. 623
33.2 Behaviour of Suspensions: The Generation of Clusters .................... 624
33.3 Conclusions .............................................. 627
References ................................................. 628

III Application to Engineering Materials Processes 631

34 Solidification Processes: From Dendrites to Design (J. A. Dantzig) 633
34.1 Introduction .............................................. 633
34.2 Dendritic Microstructures ................................... 634
34.3 Inverse Problems and Optimal Design ................................ 636
34.4 Conclusion .............................................. 638
References ................................................. 639

35 Simulation in Powder Technology (H. Riedel and T. Kraft) 641
35.1 Introduction .............................................. 641
35.2 Powder Production ........................................ 642
35.3 Die Filling ............................................... 642
35.4 Powder Compaction ....................................... 642
35.4.1 The Drucker-Prager-Cap Model and Finite Element Implementations 643
35.4.2 Experiments to Determine the Drucker-Prager-Cap Parameters ........ 645
35.4.3 Example ............................................... 647
35.5 Sintering ................................................. 648
35.5.1 Models for Solid-State Sintering .......................... 649
35.5.2 Liquid-Phase Sintering ................................... 651
35.5.3 Parameters of the Liquid-Phase Sintering Model for an Alumina Cer-amic ....................................................... 652
35.5.4 Finite-Element Implementations and Applications ................... 653
35.6 Sizing and Post-Sintering Mechanical Densification ...................... 654
35.7 Fatigue .................................................. 655
35.8 Conclusions ............................................. 655
References .................................................. 655

36 Integration of Physically Based Materials Concepts
(M. Crumbach, M. Goerdeler, M. Schneider, G. Gottstein, L. Neumann, H. Aretz,
R. Kopp, B. Pustal, and A. Ludwig) 659
36.1 Through-process Modeling of Aluminum Alloy AA2024 from Solidification
through Homogenization and Hot Rolling .......................... 661
36.2 Through-process Texture Modeling of Aluminum Alloy AA5182 during Indus-
trial Multistep hot Rolling, Cold Rolling, and Annealing .......... 665
36.3 Through-thickness Texture Evolution during Hot Rolling of an IF-Steel . . . 667
36.4 Conclusions ............................................. 668
References .................................................. 668

37 Integrated Through-Process Modelling, by the Example of Al-Rolling
(K. F. Karhausen) 670
37.1 Introduction ............................................... 670
37.2 Features of the Al Production Chain for Rolled Products .............. 671
37.3 TP Modelling of the Al Process Chain for Rolled Products .......... 673
37.4 Application of Through Process Modelling ........................ 674
37.4.1 Tracing of Dislocation Density ........................... 675
37.4.2 Tracing of Texture ....................................... 682
37.4.3 Tracing of Microchemistry ............................... 684
37.5 Conclusions ............................................. 686
References .................................................. 686

38 Property Control in Production of Aluminum Sheet by Use of Simulation
(J. Hirsch, K.F. Karhausen, and O. Engler) 688
38.1 Introduction ............................................... 688
38.2 Optimization Strategies in Sheet Processing and Material Quality .... 689
38.3 Processing and Microstructure Features of Aluminum Sheet ....... 690
38.4 Thermomechanical Simulation of Rolling Processes ................ 691
38.5 Microstructure Evolution During hot Rolling ....................... 694
38.6 Material Properties of Industrially Processed Aluminum Sheet .... 700
38.7 Simulation of Anisotropic Sheet Properties ........................ 702
38.7.1 Strength Anisotropy ...................................... 703
38.7.2 Tensile Test and r-Value Simulation ........................ 703
38.7.3 Earing During Cup Deep Drawing ........................... 704
38.8 Formability of Aluminum Sheets ................................ 706
38.9 Summary and Outlook ...................................... 707
References .................................................. 708

39 Forging
(Y. Chastel and R. Logé) 709
39.1 Introduction ............................................... 709
39.2 Case I: Microstructure Evolution During Complex Hot Forging Sequences 710
  39.2.1 Equations for Microstructure Evolutions 710
  39.2.2 Integration into a Finite Element Code 712
  39.2.3 2D Simulation Results 713
  39.2.4 Extension to 3D Forging and Dynamic Recrystallization 715
39.3 Case II: Warm Forming of Two-Phase Steels 715
39.4 Case III: Texture Evolution in an Hexagonal Alloy 718
  39.4.1 Calibrating the Polycrystalline Model with Simple Mechanical Tests 720
  39.4.2 Using the Texture-Induced Anisotropic Plastic Flow to Validate the FEM Results 721
  39.4.3 Application to Hot Forming 721
39.5 Conclusions 722
References 723

40 Numerical Simulation of Solidification Structures During Fusion Welding (V. Pavlyk and U. Dilthey) 727
  40.1 Introduction 727
  40.2 Model of Dendrite Growth under Constrained Solidification Conditions 729
    40.2.1 Solidification Problem with the Sharp Interface 729
    40.2.2 Numerical Solution 730
  40.3 Verification of the CA-FDM Solidification Model 736
  40.4 Model Application under Welding Conditions 736
    40.4.1 Macroscopic Modelling of Solidification Conditions 736
    40.4.2 Microscopic Simulation of Solidification Structures 738
  40.5 Conclusions 739
References 742

41 Forming Analysis and Design for Hydroforming (K. Chung) 744
  41.1 Introduction 744
  41.2 Ideal Forming Design Theory for Tube Hydroforming 747
  41.3 Strain-Rate Potential: Srp98 750
  41.4 Preform Design for Hydroforming Parts 751
  41.5 Summary 753
References 753

42 Sheet Springback (R. H. Wagoner) 757
  42.1 Introduction 757
  42.2 Review of Simulation Literature 758
  42.3 Review of the Experimental Literature 760
  42.4 Draw-Bend Springback 762
  42.5 Conclusions 768
References 768
## Contents

### 43 The ESI-Wilkins-Kamoulakos (EWK) Rupture Model

(A. Kamoulakos) 775

- 43.1 Background .................................................. 775
- 43.2 The EWK Fracture Model ..................................... 777
- 43.3 Academic Validation ........................................... 778
- 43.4 Semi-Industrial Validation .................................... 779
- 43.5 Conclusions .................................................... 782

References .......................................................... 783

### 44 Damage Percolation Modeling in Aluminum Alloy Sheet

(M.J. Worswick, Z.T. Chen, A.K. Pilkey, and D. Lloyd) 785

- 44.1 Introduction .................................................... 785
- 44.2 Experiment ..................................................... 787
- 44.3 Material – Characterization of Second Phase Particle Fields .... 787
- 44.4 GTN-based FE Model ............................................ 788
- 44.5 Coupled damage percolation model ............................ 791
- 44.6 Results ......................................................... 792
- 44.7 Discussion ....................................................... 794

References .......................................................... 796

### 45 Structure Damage Simulation

(D. Steglich) 797

- 45.1 Introduction .................................................... 797
- 45.2 Plastic Potentials and Porosity ................................... 798
- 45.3 Model Parameter Identification ................................ 801
- 45.4 Strain Softening, Damage and Lengthscale ..................... 803
- 45.5 Hints for Application .......................................... 805

References .......................................................... 806

### 46 Microstructure Modeling using Artificial Neural Networks

(H.-U. Löffler) 809

- 46.1 Introduction .................................................... 809
- 46.2 Artificial Neural Networks in Process Simulation .......... 812
- 46.3 Joint Microstructure Model and Neural Network System .... 816
  - 46.3.1 Physical Model .............................................. 816
  - 46.3.2 Physical Model plus Neural Network ....................... 818
  - 46.3.3 Off-line System, on-line System and in-line System ..... 819
  - 46.3.4 Results from Hot Strip Mills .............................. 821
- 46.4 Conclusions .................................................... 822

References .......................................................... 823

Index 825
Preface

This book presents our current knowledge and understanding of continuum-based concepts behind computational methods used for microstructure and process simulation of engineering materials above the atomic scale. While the area of ground-state and molecular dynamics simulation techniques has recently been reviewed in various books, no such collection exists for continuum scale materials simulation concepts. This book tries to fill that gap.

By presenting for the first time a wide spectrum of different continuum-based computational approaches to materials microstructure simulations within a single volume, we also hope to initiate the development of corresponding scientific centers in academic research institutions as well as virtual laboratories in the industry in which these methods are exploited. Moreover, we feel that other fields such as computational bio–materials science, where modeling approaches developed in the materials community have been used increasingly, might substantially benefit from the methods presented in this book.

We think that students and scientists who increasingly work in the field of continuum-based materials simulations should have a chance to compare the different methods in terms of their respective particular weaknesses and advantages. Such a critical evaluation is important since continuum models, as a rule, do not emerge directly from ab-initio calculations. In other words, continuum simulations of materials rely on approximate constitutive models which are usually not derived through quantum mechanics. This means that one should carefully check the underlying model assumptions of such approaches with respect to their applicability to a given problem. We hope that this volume provides a good overview on the different methods and allows a reader to identify appropriate approaches to the new challenges emerging every day in this exciting domain.

Continuum-based simulation approaches cover a wide class of activities in the materials research community ranging from basic thermodynamics and kinetics to large scale structural materials mechanics and microstructure-oriented process simulations. This spectrum of tasks is matched by a variety of simulation methods. The volume, therefore, consists of three main parts. The first one presents basic overview chapters which cover fundamental key methods in the field of continuum scale materials simulation. Prominent examples are the phase field model, cellular automata, crystal elasticity-plasticity finite element methods, Potts models, lattice gas approaches, discrete dislocation dynamics, yield surface plasticity, as well as artificial neural networks.

The second one presents applications of these methods to the prediction of microstructures. This part deals with explicit simulation examples such as phase field simulations of solidification, modeling of dendritic structures by means of cellular automata, statistical theory of grain growth, curvature-driven grain growth, deformation and recrystallization of particle-
containing aluminum alloys, cellular automaton simulations of grain growth, thermal activation in discrete dislocation dynamics, coarse graining of dislocation dynamics, and texture component crystal plasticity finite element methods to name but a few.

The third part presents applications in the field of process simulation. Examples are the integration of physically based materials concepts into process simulations, modeling of casting and solidification, integrated simulations of rolling, sheet forming and hydroforming simulations, springback simulation, and automotive crash simulation.

This book is intended for students at the undergraduate and graduate levels, lecturers, materials and mechanical scientists and engineers, physicists, biologists, chemists, and mathematicians. The editors would greatly appreciate any suggestions, criticisms, advice, or examples that might improve the content of this volume.

Dierk Raabe, Franz Roters, Frédéric Barlat, Long-Qing Chen

February 2004
List of Contributors

- **Guiseppe Abbruzzese, Ch. 13**  
  Centro Sviluppo Materiali  
  Terni Research Unit  
  Structural and Special Materials Dept. - Terni  
  Viale B. Brin, 218  
  05100 Terni, Italy  
  e-mail: g.abbruzzese@c-s-m.it

- **Holger Aretz, Ch. 36**  
  LASSO Ingenieurgesellschaft mbH  
  Markomannenstr. 11  
  70771 Leinfelden-Echterdingen  
  Germany  
  e-mail: aretz@lasso.de

- **Dorel Banabic, Ch. 6**  
  Technical University of Cluj Napoca  
  15 C. Daicoviciu Str.  
  3400 Cluj Napoca  
  Romania  
  e-mail: banabic@tcm.utcluj.ro

- **Jean Charles Barbe, Ch. 33**  
  LETI / CENG  
  17 Avenue des Martyrs  
  38054 Grenoble Cedex  
  France  
  e-mail: jean-charles.barbe@cea.fr

- **Frédéric Barlat, Ch. 6**  
  Materials Science Division  
  Alcoa Technical Center  
  100 Technical Drive  
  Alcoa Center, PA 15069-0001  
  USA  
  e-mail: frederic.barlat@alcoa.com

- **Wolfgang Blum, Ch. 31**  
  Universität Erlangen-Nürnberg  
  Technische Fakultät  
  Institut für Werkstoffwissenschaften LS 1  
  Martensstr. 5  
  91058 Erlangen  
  Germany  
  e-mail: wolfgang.blum@ww.uni-erlangen.de

- **Robert Brandt, Ch. 13**  
  Muhr und Bender KG  
  Schlachtwiese 4  
  57439 Attendorn  
  Germany

- **Yves Jean Marc Brechet, Ch. 9, 33**  
  L.T.P.C.M., Groupe “Physique du Métal”  
  Institut National Polytechnique de Grenoble  
  Domaine Universitaire, BP75  
  38402 Saint Martin d’Heres Cedex  
  France  
  e-mail: yves.brechet@ltpcm.inpg.fr

- **Wolfgang Brocks, Ch. 32**  
  Institute for Materials Research  
  GKSS Research Centre  
  Max Planck Str  
  21502 Geesthacht  
  Germany  
  e-mail: brocks@gkss.de

- **Einar Broese, Ch. 7**  
  Untere Bergstr. 33  
  50962 Heroldsb erg  
  Germany  
  e-mail: einar.broese@web.de
List of Contributors

• W. Craig Carter, Ch. 29
  Massachusetts Institute of Technology
  Dept. of Materials Science and Engineering
  13-5018 77 Massachusetts Ave
  Cambridge, MA 02139-4307
  USA
  e-mail: ccarter@mit.edu

• Oana Cazacu, Ch. 6
  Department of Mechanical and Aerospace Engineering
  Graduate Engineering & Research Center
  University of Florida
  1350 N. Poquito Road
  Shalimar, FL 32579
  USA
  e-mail: cazacu@gerc.eng.ufl.edu

• Yvan Chastel, Ch. 39
  Centre de Mise en Forme des Matériaux (CEMEF)
  UMR CNRS 7635, Ecole des Mines de Paris
  BP 207
  06904 Sophia Antipolis Cedex
  France
  e-mail: Yvan.Chastel@ensmp.fr

• Long-Qing Chen, Ch. 2, 11
  102 Steidle Building
  Penn State University
  University Park, PA 16802.
  USA
  e-mail: lqc3@psu.edu

• Zengtao Chen, Ch. 44
  Department of Mechanical Engineering
  University of Waterloo
  Waterloo, Ontario, N2L 3G1
  Canada
  e-mail: ztchen@lagavulin.uwaterloo.ca

• Kwansoo Chung, Ch. 41
  School of Materials Science and Engineering
  Seoul National University
  56-1, Shinlim-dong, Kwanak-gu
  151-744, Seoul
  Korea
  e-mail: kchung@snu.ac.kr

• Mischa Crumbach, Ch. 36
  Institut für Metallkunde und Metallphysik
  RWTH-Aachen
  52056 Aachen
  Germany
  e-mail: crumbach@imm.rwth-aachen.de

• Jonathan A. Dantzig, Ch. 34
  Department of Mechanical & Industrial Eng’g
  University of Illinois
  1206 West Green Street
  Urbana, IL 61801
  USA
  e-mail: dantzig@uiuc.edu

• Paul R. Dawson, Ch. 5
  196 Rhodes Hall
  Cornell University
  Ithaca, New York 14853
  USA
  e-mail: prd5@cornell.edu

• Ulrich Dilthey, Ch. 40
  RWTH Aachen
  ISF - Welding Institute
  Pontstrasse 49
  52062 Aachen
  Germany
  e-mail: di@isf.rwth-aachen.de

• Olaf Engler, Ch. 38, 22
  Hydro Aluminium Deutschland GmbH
  R+D Center Bonn
  Georg-von-Boeselager-Strasse 21
  53117 Bonn
  Germany
  e-mail: olaf.engler@hydro.com
List of Contributors

- Samuel Forest, Ch. 25
  Centre des Materiaux / UMR 7633
  Ecole des Mines de Paris / CNRS
  BP 87
  91003 Evry Cedex
  France
  e-mail: Samuel.Forest@ensmp.fr

- Charles-André Gandin, Ch. 10
  Ecole Nationale Supérieure des Mines de Nancy
  LSGMM
  Parc de Saurupt
  54042 Nancy
  France
  e-mail: gandin@lsg2m.org

- R. Edwin Garcia, Ch. 29
  Center for Theoretical and Computational Materials Science
  National Institute of Standards and Technology
  100 Bureau Drive, Mail Stop 8500
  Gaithersburg MD 20899-8500
  USA
  e-mail: Edwin.Garcia@nist.gov

- Matthias Goerdeler, Ch. 36
  Institut für Metallkunde und Metallphysik
  RWTH-Aachen
  52056 Aachen
  Germany
  e-mail: goerdeler@imm.rwth-aachen.de

- Günter Gottstein, Ch. 36
  Institut für Metallkunde und Metallphysik
  RWTH-Aachen
  52056 Aachen
  Germany
  e-mail: gg@imm.rwth-aachen.de

- Peter Gumbsch, Ch. 18, 19
  Fraunhofer Institute for Mechanics of Materials
  Woecherstrasse 11
  79108 Freiburg
  Germany
  e-mail: peter.gumbsch@iwm.fraunhofer.de

- Alexander Hartmaier, Ch. 19
  Max Planck Institute for Metals Research
  Heisenbergstrasse 3
  70569 Stuttgart
  Germany
  e-mail: Hartmaier@mf.mpg.de

- Masato Hiratani, Ch. 8
  Lawrence Livermore National Laboratory
  7000 East Av, Livermore, P.O.Box 808
  L-353, CA 94551
  USA
  e-mail: hiratani1@llnl.gov

- Jürgen Hirsch, Ch. 38
  Hydro Aluminium Deutschland GmbH
  R+D Center Bonn
  P.O. Box 2468
  53014 Bonn
  Germany
  e-mail: Juergen.Hirsch@hydro.com

- Shenyang Hu, Ch. 11
  305 Steidle Building
  Department of Materials Science and Engineering,
  Pennsylvania State University
  University Park, PA 16802.
  USA
  e-mail: sxh61@psu.edu

- Gerhard Inden, Ch. 1
  Max-Planck-Institut für Eisenforschung GmbH
  Max-Planck-Str. 1
  40237 Düsseldorf
  Germany
  e-mail: inden@mpie.de

- Alain Jacot, Ch. 10
  Laboratoire de Simulation des Matériaux
  Institut des Matériaux, Bat. G
  Ecole Polytechnique Fédérale de Lausanne
  1015 Lausanne
  Switzerland
  e-mail: Alain.Jacot@epfl.ch
List of Contributors

• Koenraad G.F. Janssens, Ch. 12
  Sandia National Laboratory
  P.O. Box 5800 MS 1411
  Albuquerque NM 87185-1411
  USA
  e-mail: kgjanss@sandia.gov

• Surya R. Kalidindi, Ch. 26, 27
  Department Head and Professor
  Department of Materials Science and Engineering
  Drexel University
  Philadelphia, PA 19104
  USA
  e-mail: skalidin@coe.drexel.edu

• Argiris Kamoulakos, Ch. 43
  E.S.I. Group
  SILIC 112
  99 rue des Solets
  94513 RUNGIS CEDEX
  France
  e-mail: ak@esi-group.com

• Kai F. Karhausen, Ch. 37, 38
  Hydro Aluminium Deutschland GmbH
  R+D Center Bonn
  P.O. Box 2468
  53014 Bonn
  Germany
  e-mail: Kai-Friedrich.Karhausen@hydro.com

• David Kinderlehrer, Ch. 16
  Department of Mathematical Sciences
  Carnegie Mellon University
  Pittsburgh, PA 15213
  USA
  e-mail: davidk@cmu.edu

• Reiner Kopp, Ch. 36
  Institut für Bildsame Formgebung
  RWTH-Aachen
  52056 Aachen
  Germany
  e-mail: kopp@ibf.rwth-aachen.de

• Torsten Kraft, Ch. 35
  Fraunhofer-Institut für Werkstoffmechanik
  IWM
  Wöhlerstr. 11
  79108 Freiburg
  Germany
  e-mail: tkr@iwm.fraunhofer.de

• Stephen A. Langer, Ch. 29
  Mathematical and Computational Sciences Division
  National Institute of Standards and Technology
  100 Bureau Drive, Mail Stop 8910
  Gaithersburg, MD 20899-8910
  USA
  e-mail: Stephen.Langer@nist.gov

• Ricardo Anibal Lebensohn, Ch. 23
  MST-8 MS G755
  Materials Science and Technology Division
  Los Alamos National Laboratory
  Los Alamos, NM 87545
  USA
  e-mail: lebenso@lanl.gov

• Jeehyun Lee, Ch. 16
  Mathematics Department
  Yonsei University
  134 Sinchon-Dong, Seodaemum-Gu
  Seoul, 463-928
  South Korea
  e-mail: jeehyun@empal.com

• Richard A. LeSar, Ch. 20
  Group T-12 Mail Stop B268
  Los Alamos National Laboratory
  TA3 Bldg 471 Rm 118
  Los Alamos, NM 87545
  USA
  e-mail: ral@lanl.gov
• Saiyi Li,  Ch. 22
Los Alamos National Laboratory
LANSCE, MS H805
Materials Science and Technology Division
(MST-8)
Los Alamos, NM 87545
USA
e-mail: saiyi@lanl.gov

• Irene Livshits,  Ch. 16
Department of Mathematics, UCA Box 4912
201 Donaghey Avenue
University of Central Arkansas
Conway AR 72032-0001
USA
e-mail: irenel@mail.uca.edu

• David Lloyd,  Ch. 44
Alcan International Limited
Kingston Research and Development Centre
Kingston, Ontario, K7L 5L9
Canada
e-mail: David.Lloyd@alcan.com

• Hans-Ulrich Löffler,  Ch. 7, 46
Siemens AG, I&S IP PEP AD1
P.O. Box 3240
91050 Erlangen
Germany
e-mail: Hans-Ulrich.Loeffler@erl9.siemens.de

• Andreas Ludwig,  Ch. 36
Metallurgical Processes Department of Ferrous Metallurgy
University of Leoben
Franz-Josef-Str.18
8700 Leoben
Austria
e-mail: andreas.ludwig@notes.unileoben.ac.at

• Kurt Lüke†,  Ch. 13
Institut für Metallkunde und Metallphysik
RWTH-Aachen
52056 Aachen
Germany

• Priya Manohar,  Ch. 4
Materials Research Division
Modern Industries, Inc.
850 Poplar Street
Pittsburgh, PA - 15220
USA
e-mail: priyam@mi-erie.com

• Volker Mohles,  Ch. 17
Institut für Materialphysik
Universität Münster
Wilhelm-Klemm-Strasse 10
48149 Muenster
Germany
e-mail: mohles@uni-muenster.de

• Zoltan Neda,  Ch. 33
Babes Bolyai University
Department of Theoretical Physics
3400, Cluj-Napoca
Romania
e-mail: zneda@nd.edu

• Britta Nestler,  Ch. 14
Karlsruhe University of Applied Sciences
Moltkestrasse 30
76133 Karlsruhe
Germany
e-mail: britta.nestler@fh-karlsruhe.de

• Roland Logé,  Ch. 39
Centre de Mise en Forme des Matériaux (CEMEF)
UMR CNRS 7635
Ecole des Mines de Paris
BP 207
06904 Sophia Antipolis Cedex
France
e-mail: Roland.Loge@ensmp.fr
XXVIII

List of Contributors

- **Luc Neumann, Ch. 36**
  Institut für Bildsame Formgebung
  RWTH-Aachen
  52056 Aachen
  Germany
  e-mail: neumann@ibf.rwth-aachen.de

- **Vitali Pavlyk, Ch. 40**
  Institut für Schweisstechnische Fertigungsverfahren
  RWTH Aachen
  Pontstrasse 49
  52062 Aachen
  Germany
  e-mail: pak@isf.rwth-aachen.de

- **Michel Perez, Ch. 33**
  GEMPPM
  INSA Lyon
  Avenue A.Einstein
  69621 Villeurbanne
  France
  e-mail: michel.perez@gemppm.insa-lyon.fr

- **Keith Pilkey, Ch. 44**
  Department of Mechanical Engineering
  Queen’s University
  Kingston, Ontario, K7L 3N6
  Canada
  e-mail: pilkey@me.queensu.ca

- **Björn Pustal, Ch. 36**
  Gießerei-Institut
  RWTH-Aachen
  52056 Aachen
  Germany
  e-mail: b.pustal@gi.rwth-aachen.de

- **Dierk Raabe, Ch. 3**
  Max-Planck-Institut für Eisenforschung GmbH
  Mikrostrukturphysik und Umformtechnik
  Max-Planck-Str. 1
  40237 Düsseldorf
  Germany
  e-mail: raabe@mpie.de

- **Bala Radhakrishnan, Ch. 15**
  Computer Science and Mathematics Division
  Oak Ridge National Laboratory
  P.O. Box 2008
  Oak Ridge, TN 37831-6140
  USA
  e-mail: radhakrishnb@ornl.gov

- **Michel Rappaz, Ch. 10**
  Laboratoire de simulation des matériaux
  Ecole Polytechnique Fédérale de Lausanne
  MXG, Ecublens
  1015 Lausanne
  Switzerland
  e-mail: michel.rappaz@epfl.ch

- **Andrew C. E. Reid, Ch. 29**
  Center for Theoretical and Computational Materials Science
  National Institute of Standards and Technology
  100 Bureau Drive, Mail Stop 8500
  Gaithersburg MD 20899
  USA
  e-mail: Andrew.Reid@nist.gov

- **Jeffrey Rickman, Ch. 20**
  Department of Materials Science and Engineering
  244 Whitaker Laboratory
  Lehigh University
  Bethlehem, PA 18015
  USA
  e-mail: jmr6@lehigh.edu

- **Hermann Riedel, Ch. 35**
  Fraunhofer-Institut für Werkstoffmechanik
  IWM
  Wöhlerstr. 11
  79108 Freiburg
  Germany
  e-mail: riedel@iwm.fraunhofer.de
• Anthony D. Rollett, Ch. 4
  Materials Science and Engineering Department
  Carnegie Mellon University
  Wean Hall 4315
  5000 Forbes Avenue
  Pittsburgh, PA 15213
  USA
  e-mail: Rollett@andrew.cmu.edu

• Franz Roters, Ch. 28
  Max-Planck-Institut für Eisenforschung GmbH
  Mikrostrukturphysik und Umformtechnik
  Max-Planck-Str. 1
  40237 Düsseldorf
  Germany
  e-mail: roters@mpie.de

• Luc Salvo, Ch. 33
  GPM2
  Institut National Polytechnique de Grenoble
  Domaine Universitaire, BP46
  38402 Saint Martin d’Heres Cedex
  France
  e-mail: luc.salvo@gpm2.inpg.fr

• Gorti B. Sarma, Ch. 15
  Computational Materials Science Group
  Computer Science and Mathematics Division
  Oak Ridge National Laboratory
  P.O. Box 2008
  Oak Ridge, TN 37831-6359
  USA
  e-mail: sarmag@ornl.gov

• Siegfried Schmauder, Ch. 30
  Institut für Materialprüfung
  Werkstoffkunde und Festigkeitslehre
  (IMWF) Universität Stuttgart
  Pfaffenwaldring 32
  70569 Stuttgart
  Germany
  e-mail: siegfried.schmauder@mpa.uni-stuttgart.de

• André Schneider, Ch. 1
  Max-Planck-Institut für Eisenforschung GmbH
  Max-Planck-Str. 1
  40237 Düsseldorf
  Germany
  e-mail: schneider@mpie.de

• Manfred Schneider, Ch. 36
  Institut für Metallkunde und Metallphysik
  RWTH-Aachen
  52056 Aachen
  Germany
  e-mail: mschneid@imm.rwth-aachen.de

• Randan Sedláček, Ch. 21
  TU München
  Werkstoffkunde und Werkstoffmechanik
  Arcisstr. 21
  80333 München
  Germany

• Mutasem Shehadeh, Ch. 8
  School of Mechanical and Materials Engineering
  Washington State University
  Pullman, WA 99164-2920
  USA

• Dirk Steglich, Ch. 45
  GKSS Research Centre Geesthacht
  Max-Planck-Str.
  21502 Geesthacht
  Germany
  e-mail: dirk.steglich@gkss.de

• Bob Svendsen, Ch. 24
  Chair of Mechanics
  Department of Mechanical Engineering
  University of Dortmund
  44227 Dortmund
  Germany
  e-mail: bob.svendsen@udo.edu
• Jiri Svoboda, Ch. 13
Institute of Physics of Materials
Academy of Sciences of the Czech Republic
Zizkova 22
616 62 Brno
Czech Republic

• Shlomo Ta’asan, Ch. 16
Department of Mathematical Sciences
Carnegie Mellon University
Pittsburgh, PA 15213
USA
e-mail: shlomo@andrew.cmu.edu

• Carlos Norberto Tomé, Ch. 23
MST-8 MS G755
Materials Science and Technology Division
Los Alamos National Laboratory
Los Alamos, NM 87545
USA
e-mail: tome@lanl.gov

• Paul Van Houtte, Ch. 22
Département MTM
Katholieke Universiteit Leuven
Kasteelpark Arenberg 44
3001 Leuven
Belgium
e-mail: Paul.Vanhouette@mtm.kuleuven.ac.be

• Burghard von Blanckenhagen, Ch. 18
Universität Karlsruhe
Institut für Zuverlässigkeit von Bauteilen und Systemen
Kaiserstraße 12
76131 Karlsruhe
Germany
e-mail: Burghard.vonBlanckenhagen@izbs.uni-karlsruhe.de

• Robert H. Waggoner, Ch. 42
Department of Materials Science and Engineering
The Ohio State University
2041 College Road
Columbus, OH 43210-1179
USA
e-mail: wagoner.2@osu.edu

• Michael Worwick, Ch. 44
Department of Mechanical Engineering
University of Waterloo
Waterloo, Ontario, N2L 3G1
Canada
e-mail: worswick@lagavulin.uwaterloo.ca

• Jeong Whan Yoon, Ch. 6
Materials Science Division
Alcoa Technical Center
Alcoa Center
PA 15069-0001
USA
e-mail: JeongWhan.Yoon@alcoa.com

• Hussein M. Zbib, Ch. 8
School of Mechanical and Materials Engineering
Washington State University
Pullman, WA 99164-2920
USA
e-mail: zbib@wsu.edu

• Michal Zyczkowski, Ch. 6
Politechnika Krakowska
Cracow University of Technology
Institute of Applied Mechanics
Al. Jana Pawla II 37
31-864 Kraków
Poland
e-mail: m-1@institute.pk.edu.pl
Part I

Fundamentals and Basic Methods for Microstructure Simulation above the Atomic Scale
1 Computer Simulation of Diffusion Controlled Phase Transformations

André Schneider and Gerhard Inden

The numerical treatment of diffusion controlled phase transformations becomes mandatory in multi-component systems and for complex conditions like heat treatments. The treatment presented here is based on (a) the concept of sharp interfaces, (b) local equilibrium at the interface, (c) thermodynamic driving forces derived from the CALPHAD approach for computational thermodynamics and (d) mobilities. It embraces isothermal and non-isothermal conditions, and various geometries. In order to treat microstructure formation and the competition between growing phases, volume may be subdivided into cells and regions which are coupled by appropriate boundary conditions.

The class of materials studied will essentially be steels. For these materials thermodynamic databases as well as data for the kinetic parameters are available. Steels offer a large variety of different transformation regimes owing to large differences in the kinetic parameters between different phases and between substitutional and interstitial elements. The regimes are reflected by typical boundary conditions: local equilibrium, local equilibrium with negligible partitioning and para-equilibrium. Several examples of diffusion controlled phase transformations in binary, ternary and higher order systems will be discussed.

1.1 Introduction

The principles of diffusion controlled phase transformations have been known for a long time. Making simplifying assumptions and considering binary systems it was possible to derive analytical solutions which provided very valuable guidelines in the field of materials development. In practice, however, materials have to be very complex multi-component systems in order to fulfill all the requirements under service conditions. In those cases the simplifying assumptions can no longer be made. This holds particularly for the thermodynamic properties which control driving forces, boundary conditions and kinetic parameters like diffusivities.

The boundary conditions cannot be taken constant during the transformation, the diffusivities become complex functions of temperature and composition (and thus also of space), and diffusion can no longer be treated on the basis of concentration gradients. In multi-component systems up-hill diffusion is a frequent case. It is thus mandatory to move to computer assisted solutions, provided the required input data are available.

A first important input is the thermodynamics of multi-component systems. About 35 years ago computational thermodynamics started with the pioneering work of Larry Kaufman (Kaufman 1970), evolving to what is known today as the CALPHAD approach (Saunders
Computer Simulation of Diffusion Controlled Phase Transformations

and Miodownik 1998, Hack 1996). The method is based on the concept of deriving, by numerical optimization techniques, the thermodynamic functions of a system from all available experimental data. The thermodynamic functions are written as polynomials of composition and temperature. The numerical values of the polynomial coefficients are the result of the optimization. Although this method aims primarily at calculating complex phase diagrams, i.e. focusing on equilibrium states, the thermodynamic functions also provide information for metastable states. Such information is needed for the treatment of phase transformations. In multi-component systems the condition of local equilibrium at a moving interface differs very much from the global equilibrium represented in a phase diagram. It is thus very important to get information about the thermodynamics of phases in their ranges of metastability. This is made available by the CALPHAD approach. Meanwhile, a variety of thermodynamic databases and software has been developed covering large parts of the wide field of materials and phases: THERMO-CALC (Sundman et al. 1985), CHEMSAGE and FACTSAGE (Eriksson and Hack 1990, Bale et al. 2002), PANDAT and MTDATA (Davies et al. 1990, 1994). In the present instance the thermodynamic calculations were performed with the software THERMO-CALC coupled with the database SSOL.

A second major input is the kinetic parameters. The thermodynamics making available true driving forces, mobilities may be taken as kinetic parameters instead of diffusivities. This reduces the number of parameters considerably. The numerical values of the mobilities can be made available from diffusion constants and composition profiles in diffusion couples using optimization routines similar to those used in computational thermodynamics. The numerical results presented here were obtained using the software DICTRA for solving the moving boundary problem and the database MOB2 for the mobilities. The concepts of numerical treatments of diffusion in multi-component systems were outlined by (Andersson and Ågren 1992) forming the basis of the DICTRA program (DICTRA). The treatment of the moving boundary problem in DICTRA is outlined in (Crusius et al. 1992 a,b). Applications are presented in several review articles (Inden and Neumann 1996, Inden 1997, Borgenstam et al. 2000, Inden 2002).

The aim of computer simulations of phase transformations is the prediction of (a) microstructural changes in multi-component systems, (b) the effect of complex heat treatments, (c) the effect of alloying elements on the kinetics of phase formation, and (d) the behavior under long term service conditions. It is only by performing computer simulations that these reactions can be described quantitatively. The numerical results often lead to unexpected results. It is only by making use of these techniques that the full spectrum of processing possibilities can be explored.

In this report simulations will be made for steels. Small amounts of alloying elements may have a strong influence on the kinetics of phase transformations. The mobilities of alloying elements may vary by orders of magnitude in different phases, e.g. in ferrite (α) compared to austenite (γ), and there is a significant difference between the kinetic parameters of substitutional and interstitial elements. Therefore, steels offer a large variety of different transformation regimes depending on time and temperature of a heat treatment. The regimes are reflected by typical boundary conditions: local equilibrium (LE), local equilibrium with negligible partitioning (LENP) and para-equilibrium (PE). Because of the fast kinetics, the formation and dissolution of metastable phases is possible. These processes occurring at intermediate stages may consume a large part of the available driving force, thus increasing the time to reach equi-
1.2 Numerical Treatment of Diffusion Controlled Transformations

1.2.1 Diffusion

Diffusion is generally treated on the basis of Fick’s equation, i.e. the movement of atoms in a concentration gradient, concentration $c_i$ of an element $i$ being defined as the number of moles of $i$ per molar volume $V_m$. Concentration is directly accessible by local analytical methods. In this case the kinetic parameters coupling concentration gradients with atomic fluxes are the diffusivities. From a thermodynamic point of view concentration gradients do not represent true forces. This becomes evident in e.g. the Darken experiment which clearly shows the diffusion of atoms against their concentration gradient. In multi-component systems this effect is very common. The situation is even more complicated in the case of phases with several sublattices, the sublattices being specified by the preferential occupation of given species. This case will not be treated here. A review of the treatment of diffusion in phases with several sublattices is given by Ågren (Ågren 1982a).

If a kinetic database is to be set up, this should not be done on the basis of diffusivities. One has to start with the fundamental equations relating fluxes to true thermodynamic forces. In this approach the kinetic parameters are the atomic mobilities. This reduces drastically the number of parameters to be stored and their composition and temperature dependence is much less complex.

For small deviations from equilibrium the flux $J_X$ of an extensive quantity $X$ may be taken proportional to the driving force given by the gradient of the intensive variable $\Phi$ conjugate to $X$ (Onsager 1931): $J_X = -\sum J_{XJ} \nabla \Phi_J$. In diffusion the most common variables are $X = V, S, n_k$ and their conjugate potentials $\Phi = P, T, \mu_k$. For a choice of conjugate variables and driving forces the $L$-matrix is symmetric: $L_{XJ} = L_{JX}$ (Onsager 1931). The off-diagonal terms in the coefficient matrix $L_{XJ}$ describe the coupling of fluxes, e.g. a particle flux due to a temperature gradient (Soret effect) or the transport of heat due to concentration gradients (Dufour effect). For the present purpose we will limit the discussion to isothermal and isobaric conditions. In this case the extensive quantities are the numbers $n_k$ of species $k$ and the conjugate variables are the chemical potentials $\mu_k$.

1.2.1.1 Lattice Fixed Reference Frame

Flux being defined as the number of species crossing a unit area per time, a reference frame has to be introduced defining the position of this unit area. We start with the lattice fixed (or Kirkendall) reference frame. Since in the solid state substitutional elements diffuse via a vacancy exchange mechanism this reference frame is defined by $\sum_{k=1}^s J^V_k + J^P_k = 0$ where $s =$ number of substitutional elements, $V =$ vacancy and $K$ stands for Kirkendall. For the
fluxes we should then write
\[ J^K_k = - \sum_{j=1}^{s+1} L^K_{kj} \frac{\partial \mu_j}{\partial \xi} \quad \text{with} \quad j \in 1, 2, \ldots, S, V \quad \text{and} \quad J^K_I = -L^K_{II} \nabla \mu_I \]

where \( I \) defines an interstitial element and \( \xi \) the distance (linear geometry assumed). If the variation of lattice parameter with composition can be neglected \( \xi \) may be replaced simply by the space variable \( z \).

The matrix of coefficients \( L^K_{ij} \) being symmetric, the relation defining the reference frame allows to eliminate one flux equation. Usually the flux of vacancies is eliminated and one gets for the substitutional elements
\[ J^K_k = -\left\{ \sum_{j=1}^s L^K_{kj} \nabla (\mu_j - \mu_V) \right\} \quad \text{and} \quad J^K_I = -L^K_{II} \nabla \mu_I \quad (1.1) \]

With a sufficient density of vacancy sources and sinks we may assume equilibrium for the vacancies and set \( \mu_V = 0 \). In the lattice fixed frame of reference there is no or little coupling between diffusional steps of substitutional elements. Therefore, the corresponding cross terms can also be neglected.

We then end up with a set of flux equations
\[ J_k = -L^K_k \nabla \mu_k \quad \text{for} \quad k \in 1, 2, \ldots, S \quad \text{and} \quad J^K_I = -L^K_{II} \nabla \mu_I \quad (1.2) \]

We may thus conclude: in the lattice fixed frame of reference elements diffuse only according to their own chemical potential gradient, and in an \( n \)-component system diffusion can be described with \( n \) kinetic coefficients \( L^K_{kk} \).

If instead of chemical potential gradients concentration gradients were used, one gets instead of (1.2) the following expressions:
\[ J^K_k = -L^K_k \left( \sum_{j=1}^{n-1} \frac{\partial \mu_k}{\partial c_j} \frac{\partial c_j}{\partial z} \right) = -\sum_{j=1}^{n-1} D^K_{kj} \frac{\partial c_j}{\partial z} \quad (1.3) \]

where component \( n \) was chosen as the dependent composition variable (since \( \sum_{j=1}^n c_j = 1/V_m \)). The diffusivity matrix \( D^K_{kj} = L^K_{kk} \frac{\partial \mu_k}{\partial c_j} \) is by no means diagonal due to the thermodynamic coupling coming from the dependence of \( \mu_k(c_1, c_2, \ldots, c_{n-1}) \) on all independent concentrations. Furthermore, the diffusivity matrix is not symmetric. In this case, the treatment of diffusion in an \( n \)-component system requires \((n-1)(n-1)\) diffusivities.

### 1.2.1.2 Mobilities

It thus turns out that in order to build up a most efficient database for kinetic parameters of \( n \)-component systems one should use the \( n \) parameters \( L^K_{kk} \) describing the diffusion of \( k \) under a chemical potential gradient of \( k \). The parameters \( L^K_{kk} \) can be expressed as atomic mobility \( M_k \) using the definition
\[ J^K_k = -M_k c_k \frac{\partial \mu_k}{\partial z} \]
1.2 Numerical Treatment of Diffusion Controlled Transformations

The mobility contains terms related to the vacancy concentration and to the exchange rate of atoms with a vacancy. Both terms refer to thermally activated processes. Therefore, the mobility will be given the form

\[ M_k = \frac{M^0_k}{RT} \exp \left( \frac{-\Delta G^*_k}{RT} \right) \]

where \( M^0_k \) is related to a frequency factor and \( \Delta G^*_k \) is a Gibbs energy barrier. Both terms depend on composition. In the mobility database these terms are represented by a Redlich-Kister polynomial (see e.g. Hillert 1998) very much the same way as it is done for the thermodynamic functions in the corresponding databases. For magnetic systems a magnetic factor has also to be introduced.

A kinetic database has been established covering a large number of elements and phases (Andersson and Ågren 1992, Jönsson 1994, Engström and Ågren 1996, Franke and Inden 1997). It is now available as MOB2 (MOB2).

1.2.1.3 Change of Driving Forces

The coefficients \( L_{kj} \) are not uniquely defined. Due to the Gibbs-Duhem equation \( \sum_{i=1}^{n} x_i d\mu_i = 0 \), the flux \( J_k \) does not change if \( L_{kj} \) is replaced by \( L'_{kj} = L_{kj} + A_k x_j \). This fact can be used to eliminate the potential gradient \( \nabla \mu_n \) of the dependent variable \( n \) such that the fluxes become \( J_k = \sum_{j=1}^{n-1} L'_{kj} (\nabla \mu_j - \nabla \mu_n) \). For this choice of driving forces the coefficient \( A_k \) has to take the form \( A_k = -\sum_{i=1}^{n} L_{ki} \) in order to have conjugate fluxes and driving forces. It is thus always possible to discuss diffusion in terms of \((n-1)\) independent driving forces \( \nabla (\mu_i - \mu_n) \).

1.2.1.4 Number Fixed Reference Frame

So far we have treated diffusion in the Kirkendall reference frame. Experimentally, however, diffusion is analyzed in other reference frames, e.g. a frame fixed with one end of a sample. The DICTRA software treats diffusion in the number fixed frame. If the partial molar volumes of the elements do not depend on composition this reference frame is equivalent to a volume fixed frame. In systems with interstitial elements no partial volume is usually attributed to these elements. Volume is thus build up by the substitutional atoms only: \( V \approx V_s = V_m / \sum_{k \in s} x_k \). The concentration is then given by

\[ c_k = \frac{x_k}{V_m} = \frac{x_k / \sum_{k \in s} x_k}{V_s} = \frac{u_k}{V_s}. \]

Therefore the reference frame is fixed by the number of substitutional atoms. The fluxes \( J'_k \) in this new reference frame must fulfill the condition \( \sum_{k=1}^{s} J'_k = 0 \).

The \( L'_{ij} \) can be expressed in terms of the parameters \( L_{ij} \). The new frame moves with some velocity \( v \) relative to the Kirkendall frame. Consequently the new fluxes \( J'_k \) must contain a contribution coming from this movement: \( J'_k = J^K_k - v \nabla \mu^*_k = J^K_k - v \frac{\Delta \mu}{RT} \). The velocity is
obtained from the condition $\sum_{k=1}^{s} J'_{k} = 0$, i.e.: $v = V_s \sum_{i=1}^{s} J^K_i$. For the new fluxes one gets

$$J'_k = J^K_k - u_k \sum_{i=1}^{s} J^K_i \quad (1.4)$$

Introducing (1.2) into (1.4) gives

$$L'_{kj} = \delta_{kj} - u_k L_{jj}$$

$$L'_{Ij} = -u_I L_{jj}$$

(1.5)

For a quaternary system with $S = A, B, C$ and an interstitial element $I$ one gets

$$J'_A = -[(1 - u_A) L_{AA} \nabla \mu_A - u_A L_{BB} \nabla \mu_B - u_A L_{CC} \nabla \mu_C]$$

$$J'_B = -[-u_B L_{AA} \nabla \mu_A + (1 - u_B) L_{BB} \nabla \mu_B - u_B L_{CC} \nabla \mu_C]$$

$$J'_C = -[-u_C L_{AA} \nabla \mu_A - u_C L_{BB} \nabla \mu_B + (1 - u_C) L_{CC} \nabla \mu_C]$$

$$J'_I = -[-u_I L_{AA} \nabla \mu_A - u_I L_{BB} \nabla \mu_B - u_I L_{CC} \nabla \mu_C + L_{II} \nabla \mu_I]$$

Performing the same conversion with Equation (1.3) for the diffusivities and defining $D'_{jk}$ in the new reference frame by

$$J'_k = -\sum_{n=1}^{s} D'_{kj} \frac{\partial c_j}{\partial z}$$

one obtains

$$J'_k = -\sum_{j=1}^{n-1} D^K_{kj} \frac{\partial c_j}{\partial z} + u_k \sum_{i=1}^{s} \sum_{j=1}^{n-1} D^K_{ij} \frac{\partial c_j}{\partial z}$$

$$= - (1 - u_k) \sum_{j=1}^{n-1} D^K_{kj} \frac{\partial c_j}{\partial z} + u_k \sum_{i=1}^{s} \sum_{j=1}^{n-1} D^K_{ij} \frac{\partial c_j}{\partial z}$$

and thus $D'_{kj} = (1 - u_k) D^K_{kj} - u_k \sum_{i=1}^{s} D^K_{ij} \delta_{kj}$ for $k \in (1, 2, \ldots, S)$ and $D'_{Ij} = D^K_{Ij} - u_k \sum_{i=1}^{s} D^K_{ij}$ for the interstitial $I$.

In summary, starting from $n$ mobilities it is possible to calculate any kind of kinetic parameter like diffusivities. The advantage of working with the mobilities is three-fold: 1. the number of independent parameters is smaller compared to any other choice, 2. the utilization of true driving forces makes sure that effects like up-hill diffusion result automatically, 3. the number of independent parameters to be varied while analyzing experimental data like diffusivities and composition profiles is smallest.

1.2.2 Boundary Conditions

In the numerical treatments of phase transformations volume is usually subdivided into two half spaces, one for the growing phase, the other for the shrinking phase, see Figure 1.1. The diffusion problem is then treated in each of these regions separately. In order to treat more complex situations space is subdivided in the DICTRA program into cells with sizes to be defined. Various geometries are possible like linear, cylindrical, ellipsoidal or spherical geometry. The cells are then coupled by appropriate boundary conditions. Within each cell
regions may be defined which are also connected by boundary conditions. The boundaries between regions are mobile. Within each region a phase has to be specified. The region will grow or shrink during the diffusion process. The phase may be set dormant such that it appears only after a critical driving force is attained. The spatial coordinate is discretized by grid points, either equidistant or with increasing density as obtained from a geometric series (Crusius 1992a). Figure 1.1 shows how position, cell size and particle size are defined in this report. In the case of two regions per cell, the symmetry in linear geometry reduces diffusion calculations to the half cell.

Figure 1.1: Definition of cell parameters in linear and spherical geometry. Here the case of two regions within one cell is shown: the growing particle $\alpha$ and the surrounding matrix $\gamma$.

The mass balance at the moving interface controls its velocity. Let us consider a moving $\alpha/\gamma$ interface between two phases $\alpha$ and $\gamma$. The flux balance is given by

$$
\frac{v^\alpha}{V^\alpha} u_k^\alpha - \frac{v^\gamma}{V^\gamma} u_k^\gamma = J_k^\alpha - J_k^\gamma \quad k = 1, 2, \ldots, n
$$

By using the two relations $\sum_{k \in S} u_k = 1$ and $\sum_{k \in S} J_k = 0$ one velocity and one flux equation can be eliminated. We thus end up with $n$ variables, the interface velocity $v$ and $(n - 1)$ independent flux equations:

$$
v \cdot (u_k^\alpha - u_k^\gamma) = J_k^\alpha - J_k^\gamma \quad k = 1, 2, \ldots, (n - 1)
$$

(1.6)

with the velocity $v = \frac{v^\alpha}{V^\alpha} = \frac{v^\gamma}{V^\gamma}$.

In the case of interdiffusion, i.e. diffusion of substitutional elements on a common sublattice, the diffusion potential may be written $\Phi_k = \mu_k - \mu_n$, where $n$ is the selected dependent element. In the case of an interstitial element the corresponding potential is $\Phi_k = \mu_k$. Diffusional equilibrium at an $\alpha/\gamma$ interface means $\Phi_k^\alpha = \Phi_k^\gamma, k = 1, 2, \ldots, (n - 1)$. We may thus fix one of the $\Phi_k$ and determine $v$ and the $(n - 2)$ remaining $\Phi_k$ by means of the $(n - 1)$ Equations (1.6). It is to be emphasized that the solution will generally not lead to $\mu_k^\alpha = \mu_k^\gamma$, the condition of (full) local equilibrium. This is schematically illustrated in Figure 1.2 for a binary substitutional system and in Figure 1.3 for a binary system like Fe-C with an interstitial component.

In the case of substitutional elements the local diffusional equilibrium condition is met with a parallel tangent construction. For a selected value of $\Phi_B^\gamma = \mu_B^\gamma - \mu_A^\gamma$ corresponding to a composition $x^{\gamma/\alpha} > x_0$ at the interface, the corresponding composition $x^{\alpha/\gamma}$ in $\alpha$ at the interface is obtained by the parallel tangent. The driving force for diffusion in $\gamma$ is related to the difference $\Phi_k^{\gamma/\alpha} - \Phi_k^{\gamma/\alpha} = \mu_k^{\gamma/\alpha} - \mu_k^{\gamma/\alpha}$. It becomes evident from Figure 1.2 that this driving force increases the more $x^{\gamma/\alpha}$ approaches $x_0^\gamma$ and increases further beyond $x_0^\gamma$. 
The same situation is illustrated in Figure 1.3 for the case of an interstitial binary alloy. In this case diffusional equilibrium at the interface means $\mu_\alpha^\gamma = \mu_\gamma^\alpha$.

The condition of local equilibrium applies to interfaces that are highly (or infinitely) mobile. All potentials being the same across the interface, there is no force acting on the interface and its movement is solely controlled by diffusion in the matrix phase $\gamma$.

In cases different from local equilibrium shown in Figure 1.3 there exists a chemical potential difference of Fe across the interface. This difference exerts a force on the interface.


Figure 1.3 this force is directed in the growth direction towards $\gamma$ and can be used to overcome a friction of the interface or to provide the force required to overcome a barrier.

In the case of a finite interface mobility the value of $\mu_C^{\gamma/\alpha}$ has to be selected such that the C-flux at the interface leads to a velocity that is the same as that following from the potential difference across the interface. Such effects shall not be considered here. They require more sophisticated approaches (Hillert 1999) including concepts of diffuse interfaces.

![Figure 1.4: Boundary condition at the interfaces of a particle with one mobile and one pinned interface (linear geometry).](image)

There is, however, one case where a velocity constraint is imposed. This situation is encountered when a particle, precipitated at a grain boundary, grows asymmetrically, i.e. only into one of the adjacent grains. Usually the interface with a high coherency exhibits a low mobility or is even pinned, i.e. $v = 0$. At a pinned interface the fluxes on either side of the interface must be equal such that the mass balance does not require an interface movement. The resulting conditions are schematically drawn in Figure 1.4. The deviation of the boundary condition from local equilibrium leads to a difference in diffusion potential through the growing particle $\alpha$. There is, of course, a force exerted on the pinned interface given by the difference in chemical potential of the substitutional component Fe, but due to the pinning this force does not lead to a movement.

This kind of situation is particularly interesting in the case of Fe-C since C diffuses about 100 times faster in $\alpha$ than in $\gamma$. The chemical potential (and thus composition) gradient through $\alpha$ leads to a short-circuit diffusion such that the mobile interface moves twice as fast as it did if both interfaces were moving under LE conditions. Consequently, the growth rate of a grain
boundary allotriomorph with only one mobile interface is practically the same as that of a symmetrically growing particle.

![Fe-Mn-C phase diagram](image)

**Figure 1.5:** Isothermal section of the Fe-Mn-C phase diagram. Alloy compositions located on an C-iso-activity line are shown. The shaded areas define ranges where different reaction regimes, with or without partitioning of Mn, occur.

### 1.2.2.1 Local Equilibrium with Partitioning (LE)

LE at an $\alpha/\gamma$ interface means $\mu_{\alpha/\gamma}^{\alpha} = \mu_{\alpha/\gamma}^{\gamma}$, $k = 1, 2, \ldots, n$, with $\mu_{\alpha/\gamma}^{\alpha}$ = chemical potential at the interface of $\alpha$ in contact with $\gamma$. The degrees of freedom in this 2-phase equilibrium is $n - 2$. Thus, the $n - 2$ free chemical potentials and $v$ are determined by the $(n - 1)$ Equations (1.6). In a binary system the degree of freedom is zero, i.e. LE has to be identical with the global equilibrium and the compositions on either side of the boundary are given by the corresponding tie-line in the phase diagram. The situation in higher order systems is more complicated. This shall be illustrated for a ternary system like Fe-Mn-C shown in Figure 1.5. In a two-phase equilibrium the degree of freedom is one. Therefore the operating tie-line is not yet fixed by the LE hypothesis.

The Fe-Mn-C system has been selected as example in Figure 1.5 to discuss the situation with a fast diffusing interstitial element, like C. Given an alloy composition in the two-phase field $\alpha + \gamma$ say 1 to 3 in Figure 1.5, the operating tie-line is defined by the criterion that the interface velocities obtained from the flux balance of every diffusing species are all equal. To meet this condition in the presence of a fast diffusing element like C, the driving force for C-diffusion must be very much smaller than that for the substitutional element, here Mn. The starting operating tie-line must thus correspond to a C-activity very close to that of the bulk material (heavy line in Figure 1.5). The reaction proceeds with this boundary condition until the driving force for C-diffusion has decreased to an extent that Mn diffusion becomes competitive. Then the operating tie-line shifts towards the global equilibrium tie-line.
1.2 Numerical Treatment of Diffusion Controlled Transformations

The alloys 1 to 4 in Figure 1.5 all correspond to the same C-activity. The starting operating tie-line is thus the same for all these alloys. This tie-line requires a partitioning of Mn between $\alpha$ and $\gamma$ since the Mn composition on either side of the interface is very much different. Therefore the velocity of transformation is essentially controlled by the substitutional diffusion leading to a slow reaction. In Figure 1.5 the range of compositions where the reaction requires partitioning is shown as shaded area labelled LE.

1.2.2.2 Local Equilibrium with Negligible Partitioning (LENP)

In the case of alloy 4 (Figure 1.5) the operating LE tie-line defines the same Mn content in both phases. For this alloy and all compositions below the broken line a reaction without partitioning of Mn is possible. This leads to a fast reaction and the boundary condition is called local equilibrium with negligible or no partitioning (LENP).

The dotted line in Figure 1.5 defines a sharp limit between the two regimes. This limit is given by the intersection of the line of constant Mn content (white broken line) with the line of constant activity in $\gamma$.

These basic concepts of LE and LENP were introduced long ago (Hultgren 1920, Hillert 1953, Kirkaldy 1962, Popov 1959, Darken 1961). For an overview see (Hillert 1998). With the new computational instruments the kinetics of these phase transformations can be calculated quantitatively for any given time-temperature profile and composition.

1.2.2.3 Para-Equilibrium (PE)

So far fast and slow reactions were the result of the thermodynamic conditions related to the composition of the alloy at a given temperature. There were no kinetic constraints involved.

Figure 1.6: Para-equilibrium phase diagram superimposed onto the full equilibrium diagram. The tie-lines of PE are horizontal lines since equilibrium was determined within a vertical section of constant Mn fraction for substitutional elements.
The situation becomes different if transformations at lower temperatures are considered where substitutional atoms definitely become more and more immobile compared to the velocity of the interface. In that case equilibrium can no longer be established with respect to these elements, but it may still be attained with respect to the fast diffusing interstitial elements. This situation was discussed by Hultgren in 1947 (Hultgren 1947). The resulting equilibrium he called para-equilibrium. Multicomponent systems like Fe-X-C are regarded as pseudo-binary systems M-C where all substitutional elements (Fe, ...) are replaced by one average component M. The thermodynamic properties of M have to be constructed such that the Gibbs energy function of the pseudo-binary system M-C is identical to the Gibbs energy function within the corresponding vertical section of the system Fe-X-C.

The conditions for para-equilibrium are defined by the equations:

\[ \mu_\alpha^C = \mu_\gamma^C \]
\[ u_\alpha^{Fe} \mu_\alpha^{Fe} + u_\alpha^{X} \mu_\alpha^{X} = u_\gamma^{Fe} \mu_\gamma^{Fe} + u_\gamma^{X} \mu_\gamma^{X} \]

The first equation imposes the local equilibrium for C, the second equation imposes the constraint that this equilibrium has to be within a section of constant fraction of substitutional element X. The resulting phase diagram is shown in Figure 1.6. The tie-lines are horizontal if u-fraction is used as coordinate for component X.

A reaction according to para-equilibrium cannot lead to the final global equilibrium. PE leads to phase fractions different from those of global equilibrium. It is important to notice that PE gives equal chemical potentials of C on either side of the boundary, but not for the potentials of substitutional components. This chemical potential difference exerts a force on the interface, but the net force is zero.

The PE reaction in diffusion controlled phase transformations has been treated for various examples (e.g. Ghosh and Olson 2001, Hutchinson et al. 2003).

### 1.2.3 Cell Size

Cell size plays an important role in the kinetics of phase transformations. It is thus mandatory to have a reliable value when it comes to comparison with experiments. In Figure 1.7 the results for the binary Fe-C system are shown. Very good agreement is obtained between experiment and calculation thanks a good knowledge of the density of nuclei obtained from metallographic analysis of serial sections every 3 \( \mu \text{m} \) (Zurek et al. 1993). Using this information no fitting had to be applied to obtain this good agreement. The dotted lines in Figure 1.7a represent the difference resulting from the indicated error range of the density of nuclei.

The importance of cell size becomes evident by means of Figure 1.8. A series of calculations is shown for the same conditions as in Figure 1.7, however within cells of different sizes. Looking at the particle sizes in Figure 1.8a one can see the common growth law and the levelling off due to soft impingement of the C-composition profiles with the cell boundary. If instead one looks at the volume fraction of ferrite formed the kinetic curves seem to suggest a faster kinetics. It is thus very important to make sure that when comparing experiments with calculations the cell size corresponds to the particle distances in the experiments. Information on the kinetics requires both quantities, average particle distance and volume fraction.
1.3 Typical Applications

1.3.1 LE, LENP and PE in Fe-Mn-C

The reactions according to the regimes LE, LENP and PE shall be illustrated taking the system Fe-Mn-C as example. At first the effect of composition shall be shown. Three compositions are considered (Figure 1.9): alloy 1 which is in the domain where reactions can only occur under LE conditions, alloy 2 which can react either according to LE or PE, and alloy 3 which is in the domain PE or LENP. Figure 1.10 shows the result of the calculation assuming a spherical cell of $50 \mu m$ radius. The large difference between slow reaction according to LE and fast reaction according to LENP or PE are clearly seen. In the case of alloy 3 the difference between PE and LENP reactions is small, except for the final volume fraction at the end of the reaction. In contrast to a binary system the boundary condition in ternary and higher order systems varies strongly with time. E.g. in the case of alloy 2 the starting tie-line (marked 2 in Figure 1.9) is far away from the tie-line at the end of the reaction which has to pass through the alloy composition (point 2). As a result, very different regimes may appear during the...
Computer Simulation of Diffusion Controlled Phase Transformations

The evolution of the reaction. It is thus very important that the full thermodynamic information is incorporated, otherwise reliable results cannot be obtained.

Figure 1.9: Isothermal section of the system Fe-Mn-C showing the starting tie-lines for three alloys according to their reaction regime.

Figure 1.10: Kinetics of transformation of alloys 1–3 in Figure 1.9. Spherical geometry, cell radius 50 μm.

Figure 1.11: Kinetics of ferrite formation from austenite: fast reaction at $T = 1000$ K.

Similar differences in the transformation regimes are obtained by variation of temperature. This is illustrated in Figure 1.11 and 1.12. From the phase diagrams, Figures 1.11a and 1.12a, it can be deduced that at 1000 K the LEPN reaction is possible, while 50 K higher it is not. The composition profiles for various time steps in Figure 1.11b show that at 1000 K, during the first 1000 s, ferrite inherits the Mn-content of austenite. During this first LEPN regime the boundary condition is constant. After soft impingement of the C-profile with the cell boundary the reaction slows down and can proceed further only with Mn diffusion. During
1.3 Typical Applications

This second LE regime the boundary condition varies considerably with time and eventually equilibrium is attained. On the contrary, at 1050 K the phase diagram Figure 1.12a shows that the LENP reaction is not possible. Therefore the reaction is slow from the beginning and can proceed only with Mn diffusion. The difference between the reactions at the two temperatures are tremendous: at 1000 K, 63% of ferrite is formed within 1000s, while at 1050 K only 0.2% is formed within $10^6$ s. This shows the overwhelming possibilities of microstructure control offered by appropriate heat treatments. There may be ways of selecting just the right temperature in order to avoid the precipitation of an unwanted phase during heat treatments within industrial time scales. Fast and slow reaction regimes are also obtained in the process of dissolution of phases. This has been treated in (Inden 1997).

![Figure 1.12: Kinetics of ferrite formation from austenite: slow reaction at $T = 1050$ K.](image)

1.3.2 LE, LENP and PE in Fe-Si-C

1.3.2.1 Ferrite Precipitation in Fe-1.15Si-0.5C

First the precipitation of ferrite from austenite in a ternary Fe-1.15 %Si-0.5 %C ($u_{Si} = 0.02272$, $u_{C} = 0.02302$) alloy at $T = 749^\circ C$ shall be discussed. Figure 1.13 shows the calculated isothermal section at equilibrium together with the diagram for para-equilibrium. It is clear that the alloy is located in the domain where both LENP and PE are thermodynamically possible. The corresponding tie-lines defining the respective boundary condition are indicated by thick lines.

Experiments were performed at this temperature and analyzed metallographically by serial sectioning. It was found that nucleation took place at grain vertices and the density of nuclei was determined as $2200 \pm 400$/mm$^3$. This density fixes the average cell size to be used in the calculation (50 $\mu$m radius for spherical geometry or 80 $\mu$m half cell size in linear geometry). This size controls the time at which soft-impingement of carbon starts reducing the driving force for C diffusion, leading to a switch from LENP to LE. The results of DICTRA calculations are shown in Figure 1.14 for linear and spherical geometry. The experiments are perfectly reproduced by the calculation for spherical geometry. Both calculations clearly show
the existence of two reaction regimes. It shall be emphasized that the excellent agreement is obtained without any fitting. At the end of the fast regime the plateau defines an almost constant volume fraction for very long times, far beyond industrial time scales.
1.3 Typical Applications

It is interesting to check the second alternative: PE. The results of the PE calculation are shown in Figure 1.15. The PE calculations could only be performed in linear geometry. For comparison the LENP calculations are also shown for the same geometry. The results show that there is very little difference between the kinetics of PE and LENP. The alloy is thus inappropriate for identifying whether the fast reaction regime might have been controlled by the PE condition or not. The reason for this becomes evident from the phase diagram in Figure 1.13. The tie-lines of both PE and LENP define C-compositions in austenite, which are almost on the same iso-activity line. The driving force for C-diffusion is thus practically the same in both cases.

1.3.2.2 Ferrite Precipitation in Fe-1.73%Si-0.4%C

The thickening and lengthening kinetics of grain boundary allotriomorphic proeutectoid ferrite was measured by Bradley/Aaronson (Bradley and Aaronson 1981) at temperatures between 725°C and 825°C. Some results are shown in Figure 1.16. The particle thickening shows the expected parabolic growth behavior for diffusion controlled reactions.

![Figure 1.16: Thickening of grain boundary allotriomorphic ferrite.](image)

DICTRA calculations were performed for both spherical and linear geometry and it turns out that the thickening rate is best reproduced using linear geometry, as it may be expected for grain boundary allotriomorph thickening.

At $T = 725°C$ and $775°C$ the DICTRA calculations for linear geometry and LENP conditions agree rather well with the experimental data of ferrite thickening. At $T = 825°C$ the LENP regime is thermodynamically not possible, as can be seen from Figure 1.17. The end point of the tie-line for LENP corresponds to a C-activity which is just less than that of the matrix. Thus, C-diffusion in austenite is directed towards the growing ferrite instead of the austenite bulk. Growth is not possible under this condition. In this instance the DICTRA calculations which are based on the local equilibrium condition thus predict growth according to LE with a thickening rate very much less than obtained in the experiments.

The results at all temperatures are compared by means of the parabolic rate constant vs. temperature. This is shown in Figure 1.18. The LENP results cannot continue beyond the
limit of 822°C, above which LENP is thermodynamically not possible. The PE calculations show excellent agreement with the growth constants in the whole temperature range. Below 822°C where LENP conditions are thermodynamically possible the LENP calculations are also close to the results for PE conditions. This is again the effect of the small difference in driving force for C-diffusion in austenite for both LENP and PE, as seen in Figure 1.17. The difference between the results for LENP and PE condition differ by less than the error bars attributed to the data points in the original paper.

Figure 1.17: Calculated isothermal section of Fe-Si-C according to full equilibrium and PE (dotted lines).

Figure 1.18: Comparison of experimental and calculated parabolic growth rates for thickening of grain boundary allotriomorphs of pro-eutectoid ferrite. The calculations indicate a reaction according to PE. The limit of possible LENP reaction is at 822°C.

The experimental data of Bradley/Aaronson at 825°C have to be interpreted as being due to a reaction of PE type, since at this temperature LENP is thermodynamically not possible. However, the limit for LENP is very close and may be too close to the experimental temperature of 825°C to draw a definitive conclusion. These results, however, show the way how to select other compositions to have better conditions for drawing conclusions.

1.3.3 PE in Fe-Ni-C

The results presented so far give the impression that PE is rarely observed. In this context the recent experimental and theoretical work by Hutchinson (Hutchinson et al. 2003) has to be mentioned which indicates that in fact PE plays an important role at early reaction times. The likelihood of experimentally observing PE depends on the range of interface velocities encountered. Since the PE reaction is very fast, special experimental techniques are required to fix the reaction steps. Hutchinson used a dilatometer instrument at IRSID/France allowing short term heat treatments and quenching rates of more than 100/s. He analyzed five alloys with compositions shown in Figure 1.19 in the isothermal section of the Fe-Ni-C system at 700°C, the temperature of the annealing experiments. The alloys are partly located in the field
1.3 Typical Applications

PE/LENP, partly in the PE field. The experimental results obtained are shown in Figure 1.20. At the very beginning all alloys show a very rapid growth. This growth is consistent with the growth under PE conditions shown as solid lines in Figure 1.20. The most interesting effect is the levelling off of the growth of the three alloys with the highest Ni-contents. Obviously, in the very early stages the interface velocity is fast enough that the Ni atoms are virtually immobile and a Ni-spike at the interface cannot develop. This regime continues until the velocity has decreased to an extent that this criterion is no longer fulfilled. Once a Ni-spike is built up, the reaction switches from PE to further growth under LE or LENP condition. This transient leads to the plateau in the thickening rate.

Following this reasoning it becomes clear why the plateau level is smaller the higher the Ni content: the higher the Ni content is, the smaller is the driving force for C-diffusion leading to slower interface velocities (given by the difference in C concentration between interface and bulk austenite) and the earlier is the moment when Ni diffusion can compete with the interface velocity. In Figure 1.20 the calculated curves for the thickening under PE conditions are also shown. They do not differ much at the early stages of reaction and they agree with the experiments at short times.

Hutchinson did set up a program allowing for a continuous transition from a starting PE reaction to the subsequent continuation with LE or LENP. The results obtained describe well the experimental observations (Hutchinson et al. 2003) and in particular can explain the formation of small amounts of \( \alpha \) in composition ranges above the LE/LENP boundary.

For the two alloys with lowest Ni-contents there are no data points available for longer times. This is due to experimental difficulties since at longer time hard impingement effects occurred as a result of the small grain size used. However, measurements of the plateau \( \alpha \) fraction formed agreed well with the LENP fractions, even though the early stage kinetics is consistent with PE. This is considered as indirect evidence of a transition from PE to LENP during growth.

1.3.4 Effect of Traces on the Growth of Grain Boundary Cementite

Si is an element that has little solubility in cementite \( \text{Fe}_3\text{C} \). In the current thermodynamic databases cementite is treated as a phase with no solubility for Si. Lateral growth of grain boundary cementite was studied experimentally by Ando and Krauss (Ando and Krauss 1981) on a (nominally) ternary Fe-2.26\%Cr-1.06\%C at 738°C.

The experimental data are shown in Figure 1.21. The calculations were first performed for the nominal composition. The starting tie-line corresponds to a fast LENP reaction. In order to comply with the experimental situation in (Ando and Krauss 1981) a spherical cell with diameter 70 \( \mu \text{m} \) was chosen, cementite growing from the outer shell of the spherical grain towards the center. The calculated results for short times are shown in Figure 1.21. For 0\%Si the time to reach a given layer thickness is more than an order of magnitude shorter than the experiments. It is only after taking into account that the alloys contained a trace of 0.03\%Si that the calculations come close to the experimental data. The rejection of the little amount of Si into the matrix slows drastically down the rate of reaction.

It is interesting to note that taking three times more Si, 0.1\%Si, makes only little difference. This demonstrates the effectiveness that traces may have on the reaction rate. At the same time, this shows nicely that higher amounts of such elements do not necessarily lead to higher effects.
1.3.5 Continuous Cooling

The phenomena discussed so far under isothermal conditions are also encountered during continuous cooling or heating. In a cooling experiment an alloy enters first into the LE region of a two phase field, then into the PE region well before it enters eventually into the LENP region. By continuous cooling it is thus possible to observe these regimes within one specimen.

Such experiments were performed with two alloys, a ternary Fe-1.5Mn-0.09C and a quaternary Fe-1.37Mn-0.42Si-0.18C with a cooling rate of 1°C/s from the austenite region (Cugy
1.3 Typical Applications

The cooling experiments were interrupted at a given temperature to determine the microstructures obtained. The fraction of ferrite was determined from metallographic analysis. The experimental results are shown in Figure 1.22. Calculations were performed for the same conditions using a cell size in accordance with the austenite grain size. The calculated temperatures of LE and PE are marked by arrows. The volume fraction precipitated under LE is negligibly small. The onset of LENP reaction is characterized by a sudden increase in volume fraction.

The experimental data were determined at temperatures below the LENP starting point. They are perfectly located on the calculated LENP branches. Since the growth rates of PE and LENP reactions are very similar, the PE reaction, if occurring, should have started at the corresponding temperatures $T_{PE}$ indicated. The volume fraction curves would then be very much like those in Figure 1.22, but shifted parallel to higher temperatures. The results thus indicate that at the cooling rates employed, the reaction proceeds by LENP.

Several arguments can be put forward to explain this. The PE start temperatures being rather high the condition of immobility of Mn and Fe (relative to the velocity of the interface) is not fulfilled. Furthermore, the driving force for nucleation within the vertical section of PE is very much less than that for LE.

1.3.6 Competitive Growth of Phases: Multi-Cell Calculations

Simultaneous growth of phases can be treated by defining different cells (with the same geometry) and coupling them together by appropriate boundary conditions. Various options are...
Figure 1.23: Coupling of three spherical cells.

possible, one of those is local equilibrium at the cell surface, i.e. the chemical potentials of all components must be the same at all cell boundaries. This is schematically illustrated in Figure 1.23 for the case of three cells with different carbides growing in each cell. The size of the cells has to be defined according to the microstructure.

The multi-cell treatment allows for including not only stable phases present in the final equilibrium, but also metastable phases which may appear at intermediate stages and disappear eventually when global equilibrium is reached. This will be illustrated with a three-cell simulation of an alloy Fe-7.8Cr-0.1C at $780^\circ$C. An isothermal section of this system is shown in Figure 1.24. The alloy is located within the three phase field $\alpha + M_{23}C_6 + M_7C_3$. Cementite $M_3C$ is stable only with austenite $\gamma$, but there exists, of course, a metastable equilibrium between cementite and ferrite $\alpha$. In Figure 1.24b the metastable extension of the phase boundary $\alpha + M_{23}C_6$ and the metastable boundary $\alpha + M_3C$ are shown. The four phases $\alpha M_{23}C_6$, $M_7C_3$ and $M_3C$ may thus all take part in the reaction, in particular if the phases nucleate easily and/or grow rapidly.

In the present instance three spherical cells with radius $r = 5 \mu m$, one for $M_{23}C_6$, one for $M_7C_3$ and one for cementite, are assumed. The result of the calculation is shown in Figure 1.25 in terms of volume fraction of phase formed, volume being given by the sum of the three cells. The mutual influence of the presence of the various phases on the growth kinetics is clearly seen. The growth rate of $M_{23}C_6$ decreases when the dissolution of $M_3C$ and of $M_7C_3$ start, respectively. For comparison, the growth of $M_{23}C_6$ only, in the same cell environment but without $M_7C_3$ and $M_3C$ in cells 2 and 3, is shown as dotted line.

$M_{23}C_6$ shows the fastest growth kinetics of all phases. It thus also controls the activities in the other cells, in particular the C-activity. The final state is given by $\alpha + M_{23}C_6$ only. It may be due to numerical accuracy that $M_7C_3$ disappeared completely, even though the alloy composition is located within the three-phase field $\alpha + M_{23}C_6 + M_7C_3$, see Figure 1.24. In this context it has to be reminded that the situation would become worse if a surface energy term would be included. This term would lead to a destabilization of small particles compared to large ones and thus accelerate their disappearance.
1.3 Typical Applications

The broken line represents the C-iso-activity line in ferrite corresponding to the alloy.

Figure 1.24: Isothermal section of the system Fe-Cr-C at $T = 780^\circ C$.

Metastable phase boundaries for the equilibrium of $\alpha + M_7C_3$ and $\alpha + M_{23}C_6$.

Figure 1.25: Fraction of phases formed as function of time. The metastable phase cementite appears in an intermediate time interval. The presence of this metastable phase influences the growth kinetics of the presence of this stable phase $M_{23}C_6$. For comparison, the growth of $M_{23}C_6$ in its own cell, but without $M_7C_3$ or cementite in the other cells, is shown as dotted line.

It is to be noticed that the stable phase $M_7C_3$ also disappears eventually. This comes from the interaction between the cells and the different growth kinetics of the phases. The system does not reach equilibrium even after $6 \times 10^6$ s. In order to reach equilibrium the phase $M_7C_3$ has to nucleate again.

This interaction becomes more transparent if one looks at the various composition profiles within the cells, Figures 1.26–1.28. Cementite takes up the lowest amount of Cr, compare Figures 1.26a to 1.28a. Therefore, cementite grows faster than $M_{23}C_6$ or $M_7C_3$. Due to the formation of the carbides, soft impingement of the C-profiles with the cell boundary takes place already after little more than 1000 s. The C-level in $\alpha$ decreases, Figs Figures 1.26b to 1.28b. This is equivalent to a shift of the alloy point to the left, see Figure 1.24. Consequently, the operating tie-line shifts to higher Cr-contents, as reflected by the Cr-profiles in Figs 26a-28a for times higher than $10^3$ s. This leads to a decrease of growth velocity.
The growth of the carbides $M_{23}C_6$ and $M_7C_3$ takes up so much Cr that already after $10^4$ s Cr diffuses out of cell 3 into cells 1 and 2. This induces the shrinking of $M_3C$. A totally similar effect occurs when all the Cr liberated by the disappearing $M_3C$ has diffused out of cell 3 after $10^5$ s. Due to the higher growth rate of $M_{23}C_6$ compared to $M_7C_3$ in the interval $10^4-10^5$, Cr is stronger absorbed by $M_{23}C_6$ leading to a composition at the cell boundary that is below the level at the interface of growing $M_7C_3$. This is visualized in Figure 1.29 with the Cr-activity profile in cells 1 and 2 for three time steps. Due to the Cr-activity gradient in cell 2, Cr diffuses away from the interface leading to the process of shrinking of $M_7C_3$ even though it is a stable phase at global equilibrium.

The carbides are treated as stoichiometric with respect to C. Cr is soluble in the carbides, but with respect to diffusion they are treated as “diffusion non phases” (essentially for two reasons: lack of experimental data and of knowledge about the mechanism of diffusion on the sublattices of the carbides). This means, a Cr-profile within the carbides is not levelled out by diffusion within these phases. Therefore, in the calculations, the Cr-content varies within the carbides as given by the boundary condition at the time steps corresponding to the interface position. The Cr profile within the carbides thus reflects the variation of the boundary condition at the interfaces with time.

In the carbides diffusion is expected to be very much slower than in the ferrite matrix. It is thus possible that the composition profile imposed by the boundary condition cannot level out during the annealing. With the high resolution techniques now available this can be checked experimentally.

During shrinking of $M_7C_3$ and $M_3C$ the interface position can be recognized by a Cr-spike (Figs 27a and 28a). This spike is the result of the condition of equal chemical potential on both sides of the interface.

### 1.3.7 Gas-Metal-Reactions: Carburization

In this review the gas-metal reactions will be confined to carburization. It is immediately evident that the same procedures can equally well be applied to nitriding or nitro-carburizing as shown by Du and Ågren (Du and Ågren 1995, Du and Ågren 1996).
1.3 Typical Applications

Figure 1.27: Composition profiles in cell 2 (M₇C₃ carbide) at various time steps.

Figure 1.28: Composition profiles in cell 3 at various time steps. Cementite grows up to about $10^4$ s. This growth is both Cr- and C-diffusion controlled as seen in Figures 1.28a,b. After $10^4$ s cementite starts shrinking, until after little more than $5.6 \cdot 10^4$ s, cementite has completely disappeared, Figure 1.28a. The C-profile has already come very close to the final value, i.e. C-activity is constant in all three cells and takes a value close to equilibrium.

After cementite dissolution Cr diffuses out of cell 3, Figure 1.28c, until after $10^5$ s the Cr-profile is almost constant.
Figure 1.29: Cr-activity profile in the three cells at time steps $5 \cdot 10^4$ s, $6 \cdot 10^4$ s and $10^5$ s. The rapid growth of $\text{M}_{23}\text{C}_6$ consumes so much Cr that after $10^5$ s, when there is no more Cr supply from cell 3, the Cr-activity becomes lower at the cell boundary than at the interface $\alpha/\text{M}_7\text{C}_3$. Consequently, Cr diffuses away from the carbide $\text{M}_7\text{C}_3$ into cell 1. This leads to the shrinking and eventual disappearing of $\text{M}_7\text{C}_3$ even though it is a stable phase.

Gas carburization occurs in industrial furnaces with carbon activities $a_C < 1$. Corrosive carburization of high temperature alloys in carbon monoxide, methane or other hydrocarbon bearing gases at $a_C < 1$ results in the formation of dispersed carbide particles, e.g. Cr-rich carbides in the case of Cr-alloys, leading to embrittlement of structural material. At activities $a_C \gg 1$ the high temperature corrosion phenomenon, known as metal dusting (Grabke 2001), occurs.

Generally gas carburization of metals can be treated as follows (Sproge and Ågren 1988, Grabke 2001):

The carbon flux density from the gas to the metal depends on a mass transfer coefficient $f$ and on the carbon activities in the gas and at the surface:

$$ j_C = \frac{1}{A} \cdot \frac{dn_{\text{C}}}{dt} = f \cdot (a_{\text{gas}}^C - a_{\text{surface}}^C) . $$

The carbon diffusion in the metal matrix is driven by the gradient of the chemical potential of carbon:

$$ j_C = -M_C c_C \frac{d\mu_C}{dx} = -D_C c_C T \frac{d\mu_C}{dx} . $$

The mass transfer coefficients of the various possible chemical surface reactions, e.g. the slow carburizing gas mixture $\text{CH}_4 = \text{C} + 2\text{H}_2$ or the fast carburizing mixture $\text{CO} + \text{H}_2 = \text{C} + \text{H}_2\text{O}$, have to be introduced. They have to be determined from experiments as function of temperature and pressure. They are defined as $f = k_1' \cdot p_{\text{H}_2}^{3/2}$ or $f = k_2' \cdot \frac{p_{\text{H}_2}}{p_{\text{H}_2}^2}$, respectively. Numerical values can be found in (Grabke 2001).

1.3.7.1 Carburization of pure Fe

Examples of slow and fast carburization simulations using DICTRA are shown in Figure 1.30. Slow carburization of an iron sample with 1 mm in thickness in a $\text{CH}_4$–$\text{H}_2$ gas mixture at $700^\circ\text{C}$ at $a_C = 1$ with a relatively low mass transfer coefficient $f = 3.4 \cdot 10^{-11}$ mole$\cdot$cm$^{-2}$s$^{-1}$ leads to
1.3 Typical Applications

a surface controlled carburization reaction, as shown in Figure 1.30a. The carbon activity at the iron surface increases slowly with increasing time. The opposite situation is represented by carburization in a CO–H₂–H₂O mixture with a high mass transfer coefficient \( f = 3.4 \cdot 10^{-6} \text{ mol cm}^{-2} \text{s}^{-1} \), which leads to fast carbon transfer. In this case the carbon activity at the surface increases very rapidly and the carburization reaction can be regarded as a diffusion controlled process (Figure 1.30b).

Carburization in H₂S-containing atmospheres leads to a reduced kinetics due to a sulphur coverage \( \Theta_S \): 

\[
\dot{\Theta}_S = f \cdot (1 - \Theta_S) \cdot (a_C^{\text{gas}} - a_C^{\text{surface}}).
\]

Using this relationship the sulphur coverage can be determined indirectly by thermogravimetric experiments where the carbon flux density is measured (Grabke 1977).

![Figure 1.30](image)

**Figure 1.30:** Activity profiles obtained from DICTRA simulation for carburization of iron with 1 mm in thickness at 700°C and \( a_C = 1 \). a) slow carburization in a CH₄–H₂ mixture with \( f = 3.4 \cdot 10^{-11} \text{ mol cm}^{-2} \text{s}^{-1} \), b) fast carburization in a CO–H₂–H₂O mixture with \( f = 3.4 \cdot 10^{-6} \text{ mol cm}^{-2} \text{s}^{-1} \).

1.3.7.2 Carburization of Ni-Cr

The computer simulation of carburization of a Ni-25Cr high-temperature alloy was treated by Engström et al. (Engström 1994). Since the conditions of the corresponding experiments define a fast surface reaction, the DICTRA simulation is simplified taking the carbon activity at the alloy surface constant at \( a_C = 1 \). In this instance a transfer coefficient need not to be introduced into the calculation, consequently the reaction is just diffusion controlled. After being transferred to the material carbon diffuses within the matrix leading to precipitates of \( M_7C_3 \) and \( M_3C_2 \). In order to treat this problem with DICTRA a special module for dispersed precipitates was implemented (Engström 1994/2). The idea is to subdivide the carburized matrix into slices and to assume global equilibrium within every slice. The phase fractions are obtained from the equilibrium calculation. Within a segment more than two phases may precipitate at the same time. During the first calculation step, the so-called diffusion step, diffusion in the matrix phase is treated as a one phase problem. Each diffusion step results in
a new carbon profile. The second step is the Thermo-Calc calculation of the new equilibrium at each position of the cell. Then the diffusion step is repeated with the new composition profile. The expected sequence of carbide formation during carburization of a Ni-25Cr alloy can be read directly from the calculated isothermal section of the ternary Ni-Cr-C system in Figure 1.31a. At first M$_7$C$_3$ and with increasing carbon content M$_3$C$_2$ will be formed. The local distribution of these phases after a reaction time of 3600000 s is shown in Figure 1.31b. The profile plotted in Figure 1.31b represents the average local carbon content, the average being taken over matrix and precipitated carbides.

![Figure 1.31](image1.png)

**Figure 1.31:** Simulation of carburization of a Ni-25Cr high-temperature alloy at 850°C:

a) the expected sequence of carbide formation during carburization can be deduced from the vertical grey line in the isothermal section of the ternary Ni-Cr-C system at 850°C calculated with Thermo-Calc

b) local distribution of these phases after a reaction time of 3600000 s.

In order to understand the thermodynamics and kinetics of the complex carburization reactions in the metal dusting process of iron DICTRA simulations were performed and compared with the experiments. Here the concept of dispersed phases and a transfer coefficient were introduced in the calculations (Schneider and Inden 2003). The experiments were performed at 800°C in CO-H$_2$-H$_2$O gas mixtures with $a_C = 100$. Such conditions are known to be strongly carburizing. Starting with pure $\alpha$-Fe the first phase transformation during carburization is the $\alpha - \gamma$-transition by forming an austenite layer, which is in fact observed in the experiments (see Figure 1.32a). When the austenite becomes supersaturated with carbon cementite will form at its surface. Normally cementite layers are observed after metal dusting treatment of iron, but in the case of metal dusting with prior austenite formation cementite is observed to precipitate as particles which are selected from each other (see Figure 1.32b). Therefore it is reasonable to apply the model of diffusion in dispersed systems.

The three steps of the DICTRA simulation are shown schematically in Figure 1.33.

During the first step $\alpha$-Fe is carburized with a mass transfer coefficient $f(\alpha - \text{Fe})$ with $a_C = 100$ in the gas atmosphere. Then $\gamma$-Fe nucleates at the surface, which grows as a layer,
1.3 Typical Applications

Figure 1.32: Metallographic cross section of an iron sample carburized at 800°C in a CO-H₂-H₂O mixture for 7 h.

a) The low magnification image clearly reveals that during the carburization process a γ-Fe layer with a fairly homogeneous thickness has formed. At the end of the experiment the sample was cooled quickly in an air stream and the austenite transformed into martensite.

b) The higher magnification of the same cross section on the right side shows cementite particles at the surface covered with coke.

Figure 1.33: Schematic representation of the three steps in the DICTRA simulation of carburizing iron at 800°C with αC ≫ 1. During the first step α-Fe is carburized with a mass transfer coefficient \( f(α - Fe) \). Then γ-Fe starts to grow as a layer, and the mass transfer coefficient is changed to \( f(γ - Fe) \). Cementite, which is entered as a dispersed phase in the γ-Fe matrix, starts to precipitate after supersaturation of the austenite.

and the mass transfer coefficient is changed to \( f(γ - Fe) \). In the calculation cementite is entered as a dispersed phase in the γ-Fe matrix. After supersaturation of the austenite cementite precipitation starts at the surface near region. The calculated profiles of carbon activities and concentration at different times during these reaction steps are shown in Figure 1.33. By comparing the simulations with the experiments only the mass transfer coefficients are treated as adjustable parameters. The heavy line in Figure 1.33 represents the mass gain curve with time measured by thermogravimetry. The upper light grey curve represents a DICTRA simulation without dispersed cementite. Here the mass transfer coefficients are calculated from literature data. It is striking that the experimental and calculated curves do not fit. Entering dispersed cementite into the simulation the result is significantly improved (see dotted grey
curve in Figure 1.34). The shape of the curve is now similar to the experimental one and they both show a change in the mass gain kinetics between 0.6 and 2 hours of the reaction. It should be mentioned that the literature data of the mass transfer coefficients are very reliable for conditions with low carbon activities. At carbon activities $a_C \gg 1$ their validity is questionable. Consequently it seems to be necessary to adjust these parameters with respect to the experiments. Since the $\alpha \rightarrow \gamma$ transformation appears very quickly, after about 1 s, $f(\alpha \rightarrow \text{Fe})$ is not adjusted. Only $f(\gamma \rightarrow \text{Fe})$ was varied in that way that the slopes of the experimental and calculated curves are the same during the first hour of the reaction in the regime of linear mass gain kinetics. Now both the experimental and calculated curves (black and grey curves in Figure 1.34) of the mass gain kinetics fit satisfactorily. The different slopes of the curves after about 2 hours of the reaction can be explained by an onset of graphite deposition which can not be taken into account in the DICTRA simulation.

![Figure 1.34: Experimental mass gain curve measured by thermogravimetry in comparison with calculated ones. The light grey curve represents a simulation without dispersed cementite. The mass transfer coefficients $f(\alpha \rightarrow \text{Fe})$ and $f(\gamma \rightarrow \text{Fe})$ are calculated from literature data. The dotted curve shows the effect of dispersed cementite entered into the simulation: the shape of the curve is now more similar to the experimental one. The grey curve, which fits well with the experimental curve, is calculated after adjustment of the mass transfer coefficient $f(\gamma \rightarrow \text{Fe})$.](image)

The profiles in Figure 1.35 show that after austenite nucleation at about 1 s austenite layer growth takes place. The carbon is transferred from the gas phase to the austenite, diffuses relatively slowly to the $\gamma \rightarrow \alpha$ interface and diffuses quickly through the ferrite matrix. The increase of the carbon activity at the surface with increasing reaction time indicates a surface reaction controlled carburization. After 2000 s the carbon activity at the austenite surface reaches the value necessary for cementite precipitation. Now there is a transition to a diffusion controlled reaction. This is reflected by the change of the calculated mass gain kinetics after 2000 s ($\sim 0.6$ h) shown in Figure 1.35.
1.4 Outlook

These calculations are help understanding the complex driving forces of the gas-metal interactions and simultaneous diffusion controlled phase transformations, e.g. during high temperature corrosion.

![Figure 1.35: Calculated profiles of carbon activities (a) and concentration (b) at different times during carburization of iron at 800°C with $a_C = 100$.](image)

1.4 Outlook

The results presented show that the numerical calculation of diffusion controlled phase transformations has reached a high level of accuracy and agreement with experimental data. It has become a valuable tool for research and alloy development. The limitations are coming from deficiencies in the thermodynamic description of phases which have previously been determined only for the purpose of phase diagram calculations, discarding any aspect relevant for diffusion. Many of the phases are treated as so-called “diffusion-non-phase”, i.e. they are stoichiometric with respect to certain or even all components and do not allow diffusional transport within the phase. The mobilities can only be defined in accordance with the thermodynamic description. Consequently, they suffer from the same limitations. It can be expected that these limitations will be removed by future work.

There are other limitations which are obvious, e.g. the limit to geometries with one spatial coordinate only. The distribution of particles may be such that a real 3-dimensional treatment is required. e.g. experiments indicate that in ferritic steels the Laves phase nucleates preferentially at $M_{23}C_6$ particles. It is thus to be expected that the growth of Laves phase may is influenced by an exchange of elements between these phases. Such problems require methods like phase field. The same is true for shape instabilities and elastic contributions.

Another limitation comes from the sharp interface concept which is the basis of the DICTRA-type approach. Effects like solute drag or solute trapping cannot really be treated properly without opening the method to a diffuse interface.
The merit of the DICTRA method still is the sophisticated thermodynamic description. For many practical problems the thermodynamic part plays already a key-role and valuable results for materials development can be obtained already from this approach. The reactions in multi-component systems become extremely complex and can hardly be guessed in advance. The guidance offered by the quantitative predictions obtained with this concept have become a valuable tool.

Acknowledgements

One of the authors (G.I.) wishes to thank the CNRS/France for making available a research stay at Grenoble. This gave the opportunity for intense discussions with C. Hutchinson and Y. Bréchet on para-equilibrium reactions and for a co-operation in that subject. Thanks are due for the discussions and for making available results prior to publication.

References

P. Cugy and M. Kandel, IRSID/France, unpublished work, private communication.
References

P. Franke and G. Inden (1997a) Diffusion Controlled Transformations in Multi-Particle Systems, Z. Metallkunde 88, 917-924.
C.R. Hutchinson, private communication.
References

Science, Oxford.


Software THERMO-CALC: thermodynamic equilibrium calculations, developed at the Royal Institute of Technology, Stockholm. It is updated and implemented by the Foundation of Computational Thermodynamics, Royal Institute of Technology, Stockholm/Sweden.

DICTRA: calculation of Diffusion Controlled TRANsfomations. The first version was developed in 1988-1993 in a co-operation between the Royal Institute of Technology, Stockholm, (Group of Profs. M. Hillert and J. Ågren) and Max-Planck-Institut für Eisenforschung GmbH, Düsseldorf (Group of Prof. Inden) within a project (COSMOS) supported by Volkswagen-Stiftung and Land Nordrhein-Westfalen. Since 1993 the software is updated and implemented by the foundation THERMOCALC.

SSOL: Solid Solution Data Base provided by SGTE (Scientific Group Thermodata Europe).

PANDAT: Software for Multicomponent Phase Diagram Calculation, developed and maintained by Computherm LLC, 437 S. Yellowstone Drive, Suite 217, Madison WI 53719, USA.

MTDATA: Thermodynamic equilibrium calculation. Software developed and maintained by the Thermodynamics and Process Modeling Group, NPL Materials Centre, National Physical Laboratory, Teddington, Middlesex, TW11 0LW, UK.

CHEMSAGE, FACTSAGE: Thermodynamic equilibrium calculation, software developed and maintained by GTT-Technologies, Kaiserstrasse 100, 52134 Herzogenrath/Germany and CRTC, Ecole Polytechnique de Montreal/Canada.
2 Introduction to Phase-field Method of Microstructure Evolution

Long-Qing Chen

2.1 Introduction

Essentially all engineering materials contain certain types of microstructures, and our success of designing new materials is largely dependent on our ability to control them. Microstructure is a general term which refers to a spatial distribution of structural features which can be phases of different compositions and/or crystal structures, or grains of different orientations, or domains of different structural variants, or domains of different electrical or magnetic polarizations, as well as structural defects such as dislocations. The length scales of structural features in a microstructure are typically in the range of nanometers to a few microns, and in some cases they can be hundreds of microns. It is the size, shape, and spatial arrangement of the local structural features in a microstructure that determine the physical properties of a material such as mechanical, electrical, magnetic and optical properties.

Microstructures evolve during materials processing or in service at high temperatures. Microstructure evolution may be caused by phase transformations to reduce the bulk free energy or by particle or domain coarsening to eliminate the total interfacial energy. It may also take place to respond to external fields such as an applied stress or electrical or magnetic field. Ideally, one would like to capture the optimum microstructure during processing for desirable properties and slow down or even stop its evolution during applications. In order to realize and achieve such optimum microstructures and control their evolution, one must be able to predict the temporal microstructure evolution under given processing conditions such as temperature, cooling rates, pressure, and composition. There has been an enormous effort in developing various types of microstructure models in the recent past (Chen, Cahn et al. 1996; Bullard, Chen et al. 1998; Raabe 1998). One of the computational models that have received tremendous attention during the last decade is the phase-field method for modeling microstructure evolution processes and a number of reviews on the method have appeared (Chen and Wang 1996; Wang and Chen 1999; Karma 2001; Ode, Kim et al. 2001; Chen 2002). The phase-field method describes a microstructure as a whole using one or more field variables which are continuous functions of space. It is based on a diffuse-interface description for the thermodynamics of inhomogeneous systems (Waals 1893; Cahn and Hilliard 1958; Rowlinson 1979). The field variables, and thus the microstructure evolution, evolve with time according to a set of well established kinetic equations (Cahn 1961; Allen and Cahn 1977). One of the most attractive features for the phase-field method is the fact that one does not have to explicitly track the usually complicated interfaces during a microstructure evolution,
as one has to do in a conventional modeling in order to specify the boundary conditions at the interfaces. There are three separate chapters in this book which are devoted to the applications of the phase-field approach to three important materials processes, solidification, solid-state phase transformations, and grain growth, respectively. Therefore, this chapter will serve as an introduction to the phase-field method.

2.2 Origin of the Model

The name “Phase-field model” was first introduced in modeling solidification of a pure melt (Fix 1983; Collins and Levine 1985; Langer 1986). The idea was to avoid the explicit tracking of a solid-liquid interface during solidification by replacing the sharp solid-liquid interface using an artificial, continuous, non-conserved phase field, i.e. a diffuse interface. Instead of defining a moving boundary condition at the solid-liquid interface, the interface movement is described by the temporal evolution of the zero or a certain level-set of the phase field. Such a diffuse-interface description of an interface was first proposed more than a century ago by Van der Walls for fluid interfaces using a density field (Waals 1893; Rowlinson 1979). The magnetic domain theory by Landau and Lifshitz has essentially the same structure (Landau and Lifshitz 1935). The best known diffuse-interface theory to us today is due to Cahn and Hilliard who studied the interphase boundaries described by a composition field (Cahn and Hilliard 1958), which subsequently lead to the development of the well known Cahn’s theory of Spinodal decomposition concerning the evolution kinetics of a conserved composition field (Cahn 1961). The diffuse-interface description of an antiphase domain wall motion was first proposed by Allen and Cahn, i.e. the Allen-Cahn equation for a non-conserved long-range order parameter (Allen and Cahn 1977). It is the Allen-Cahn equation which governs the evolution of the non-conserved phase field in solidification modeling. The Cahn-Hilliard nonlinear diffusion equation (Cahn 1961) and the Allen-Cahn (sometimes also called the time-dependent Ginzburg-Landau) (Allen and Cahn 1977) equation provide the basic governing equations for phase-field models. With random thermal noises, both types of equations become stochastic, and their applications to studying critical dynamics as well as to modeling the morphological pattern formation during ordering and phase separation of quenched systems have been extensively explored long before the term “phase field” was coined (Hohenberg and Halperin 1977; Gunton, Miguel et al. 1983).

2.3 Theoretical Fundamentals of the Method

Although phase-field models may be presented in different forms for different applications, they are all based on a common set of fundamentals with regard to the representation of a microstructure, the thermodynamic description, and the kinetic evolution equations for microstructures.

2.3.1 Representation of a Microstructure

In a phase-field model, a microstructure is described using one or more physical and/or artificial field variables. These variables are uniform inside a phase or domain away from the
interfaces. The same phase or the same types of domains have the same uniform values for the field variables. 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 Figure 2.1.

![Figure 2.1](image)

Field variables can be either conserved or non-conserved, depending on if they satisfy the local conservation law, \( \frac{\partial \phi}{\partial t} = -\nabla \cdot J \) where \( \phi \) is a field variable and \( J \) 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 system.

Artificial fields are introduced for the sole purpose of avoiding tracking the interfaces. 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 (Caginalp and Xie 1993; Karma and Rappel 1996; Kim, Kim et al. 1999; Elder, Grant et al. 2001; Karma 2001). Physical fields refer to well-defined order parameter fields which can be experimentally measured. The interfacial width described by a physical field is also expected to reflect the actual interfacial width. In phenomenological theories of phase transformations, order parameter fields are used to characterize the nature and the critical temperatures of the phase transformations which produce the microstructure. 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 structure resulted from ordering. Another example is a composition field which describes the morphological evolution during phase separation either through nucleation and growth or spinodal decomposition or during precipitate coarsening.
2.3.2 Thermodynamics of Microstructures

In the phase-field approach, the thermodynamics of an inhomogeneous microstructure is described by the diffuse-interface theory of Cahn and Hilliard (Cahn and Hilliard 1958). According to the Cahn-Hilliard theory, for a system described by a composition field, the local free energy, \( f \), for an inhomogeneous system depends not only on the local composition, \( c \), but also on the neighboring compositions, i.e. composition gradients. Truncating at the second order in the free energy expansion in terms of composition gradients with respect to a uniform composition, the total chemical free energy of a system with volume \( V \) is given by (Cahn and Hilliard 1958),

\[
F = \int \left[ f(c) + \frac{\alpha}{2} (\nabla c)^2 \right] d^3r
\]

(2.1)

where \( \alpha \) is the composition gradient energy coefficient. The gradient energy term is zero inside a phase or domain where the composition is uniform, but is nonzero across interfaces where the composition field varies. Therefore, Equation (2.1) represents the total chemical free energy of a microstructure including the bulk chemical free energy as well as the excess free energy associated with all the interfaces in the microstructure.

For a more general inhomogeneous system described by a field variable described by a set of conserved field \( c_i \) and a set of non-conserved field \( \eta \), the total free energy can be written as

\[
F = \int \left[ f(c_1, c_2, \ldots, c_n, \eta_1, \eta_2, \ldots, \eta_p) + \sum_{i=1}^{n} \alpha_i (\nabla c_i)^2 + \sum_{i=1}^{3} \sum_{j=1}^{3} \sum_{k=1}^{p} \beta_{ij} \nabla i \eta_k \nabla j \eta_k \right] d^3r
+ \int \int G(r - r')d^3rd^3r'
\]

(2.2)

where \( f \) is the usual local free energy density as a function of all the field variables, \( \alpha_i \) and \( \beta_i \) are the gradient energy coefficients. The second double-integral represents the contributions from pair-wise, non-local, long-range interactions \( (G(r - r')) \) such as elastic, electrostatic, and magnetic interactions between volume elements \( d^3r \) and \( d^3r' \). For example, long-range elastic interactions arise in coherent systems where the lattice planes and directions across the interfaces are continuous, and the lattice parameters of the phases depend on the field variable. For simplicity, many-body interactions, for example, many-body elastic interactions that appear in systems with a strong elastic modulus inhomogeneity (Khachatryan, Semenovskaya et al. 1995), are ignored in the discussions below. Due to the importance of various energetic contributions in the phase-field modeling, each individual term in Equation (2.2) is briefly discussed below.

2.3.2.1 The Local Free Energy

The main difference among different phase-field models lies in the construction of \( f \) as a function of field variables. A simple and familiar system is a two-phase binary system described
2.3 Theoretical Fundamentals of the Method

A physical composition field. If one is interested in a real system with a known free energy function as a function of temperature, it can be directly used. On the other hand, if one is only interested in modeling the morphological evolution in a model system, a simple double-well potential may be used. For example,

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

(2.3)

where \( A \) and \( B \) are constants. At equilibrium, Equation (2.3) describes two phases with compositions \( \pm \sqrt{A/B} \) and \( -\sqrt{A/B} \), respectively.

For the case of isothermal solidification of a binary liquid, an artificial field, \( \phi \), is introduced, in addition to the physical field, composition \( c \). The artificial field, also called the phase field, is used to distinguish solid and liquid phases and to automatically take into account the boundary conditions at the interfaces. The local free energy as a function of composition \( c \) and \( \phi \) can be expressed as (Wheeler, Boettinger et al. 1992),

\[ f(c, \phi) = h(\phi) f^S(c) + [1 - h(\phi)] f^L(c) + w g(\phi) \]  

(2.4)

where \( f^S(c) \) and \( f^L(c) \) are the free energy densities of solid and liquid as a function of composition at a given temperature, respectively. \( g(\phi) \) is a double-well potential which is only a function of the artificial field. A possible functional form for \( g \) is

\[ g(\phi) = 16w\phi^2 (\phi - 1)^2 \]  

(2.5)

which has minima at \( \phi = 0 \) and \( \phi = 1 \). \( w \) in Equation (2.5) represents the depth of the two wells at \( \phi = 0 \) and \( \phi = 1 \) with respect to the local maximum at \( \phi = 1/2 \) as shown in Figure 2.2.

![Figure 2.2: Illustration of a double-well potential with potential depth of \( w \).](image)

The function, \( h(\phi) \) in Equation (2.2), is required to have the following properties:

\[ h(0) = 0, \quad h(1) = 1, \quad \frac{dh}{d\phi} \bigg|_{\phi=0} = \frac{dh}{d\phi} \bigg|_{\phi=1} = 0 \]  

(2.6)

There properties ensure that the equilibrium values, 0 and 1, for the phase field in the double-well potential are not affected by the chemical free energies \( f^S \) and \( f^L \). An example which satisfies the conditions given in Equation (2.6) is

\[ h(\phi) = \phi^3 (6\phi^2 - 15\phi + 10) \]  

(2.7)

Therefore, in this model, \( \phi = 0 \) represents the liquid phase since \( h(0) = 0 \) and \( f(c, 0) = f^L(c) \). Similarly, \( \phi = 1 \) describes the solid phase with \( h(1) = 1 \) and \( f(c, 0) = f^S(c) \). Across
Introduction to Phase-field Method of Microstructure Evolution

the interface, the local free energy has contributions from both the liquid and solid chemical free energies as well as the double-well potential.

Recently Kim et al. (Kim, Kim et al. 1999) proposed a variation of the above model (which they refer to as the WBM model), based on an earlier work by Steinbach et al. (Steinbach, Pezzolla et al. 1996). They considered the interfacial region to be a mixture of solid and liquid phases with compositions \( c_S \) and \( c_L \) and with the same chemical potential, i.e. \( c_S \) and \( c_L \) satisfy the following set of equations for a binary system,

\[
c = [1 - h(\phi)]c_L + h(\phi)c_S
\]

\[
\frac{\partial f^L(c_L)}{\partial c_L} = \frac{\partial f^S(c_S)}{\partial c_S}
\]

The main advantage of Kim’s approach compared to the WBM model is the fact that for an interface at equilibrium, there is no contribution of the actual chemical free energy, \( f_S \) and \( f_L \), to the total interfacial energy, i.e. \( \Delta f \) (see Figure 2.3) is eliminated at equilibrium because of conditions (2.8) and (2.9). The interfacial energy and interfacial width are entirely determined by the double-well potential (Figure 2.2) 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 (2.13)), thus increasing the length scale of a phase-field simulation even with the usual numerical methods using uniform grids. However, the implementation of this model requires the numerical solution to the above coupled Equations (2.6) and (2.7) for \( c_L \) and \( c_S \) for each set of \( c \) and \( \phi \) 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 \( f^L \) and \( f^S \)), i.e. \( \Delta f \) in Figure 2.3, for phase transformations during a phase-field simulation, and thus the interfacial width that one can use is also limited. Otherwise numerical instability may develop in a phase-field simulation, leading to incorrect path for the microstructure evolution.

For many solid-state structural phase transformations, the local free energy function is typically expressed as a polynomial of order parameters using a conventional Landau-type of expansion since the field variables correspond to well-defined physical order parameters. All the terms in the expansion are required to be invariant with respect to the symmetry operations of the high-temperature phase. For example, for precipitation of an ordered phase (L1_2) from a face-centered-cubic (FCC) matrix in a binary alloy, with expansion terms up to the fourth order, the local free energy function is given by (Lai 1990; Braun, Cahn et al. 1997; Li and
2.3 Theoretical Fundamentals of the Method

Chen 1998; Wang, Banerjee et al. 1998)

\[
f(c, \eta_1, \eta_2, \eta_3) = f_d(c, T) + \frac{1}{2} A_2(c, T) \left( \eta_1^2 + \eta_2^2 + \eta_3^2 \right) + \frac{1}{3} A_3(c, T) \eta_1 \eta_2 \eta_3
\]

\[
+ \frac{1}{4} A_{41}(c, T) \left( \eta_1^4 + \eta_2^4 + \eta_3^4 \right)
\]

\[
+ \frac{1}{4} A_{42}(c, T) \left( \eta_1^2 \eta_2^2 + \eta_2^2 \eta_3^2 + \eta_3^2 \eta_1^2 \right)
\]

(2.10)

where \( f_d(c, T) \) is the free energy of the disordered phase when all the order parameter values are zero, i.e. \( \eta = 0 \). \( A_2, A_3, A_{41}, \) and \( A_{42} \) are the expansion coefficients which are functions of temperature and composition. The free energy function (Equation (2.10)) has four degenerate minima with respect to the order parameters. If \( A_3(c, T) < 0 \), the free energy minima are located at

\[
(\eta_0, \eta_0, \eta_0), (\eta_0, -\eta_0, -\eta_0), (-\eta_0, \eta_0, -\eta_0), (-\eta_0, -\eta_0, \eta_0)
\]

(2.11)

where \( \eta_0 \) is the equilibrium value for the long-range parameter at a given composition and temperature. The four sets of long-range order parameters given in Equation (2.11) describe the four energetically equivalent antiphase domains of the \( \text{L1}_2 \) ordered phase related by a primitive lattice translation of the parent disordered FCC phase.

Another example is the free energy function employed in a phase-field model for domain structure evolution during a cubic to tetragonal ferroelectric phase transition. In this case, the local electric polarization fields, \( (P_1, P_2, P_3) \) are the natural field variables that describe a domain structure. The free energy function as a function of polarization can be written as (Devonshire 1954; Nambu and Sagala 1994; Hu and Chen 1998),

\[
f(P_1, P_2, P_3) = A_1 \left( P_1^2 + P_2^2 + P_3^2 \right) + A_{11} \left( P_1^4 + P_2^4 + P_3^4 \right)
\]

\[
+ A_{12} \left( P_1^2 P_2^2 + P_2^2 P_3^2 + P_3^2 P_1^2 \right) + A_{111} \left( P_1^6 + P_2^6 + P_3^6 \right)
\]

\[
+ A_{112} \left[ P_1^4 \left( P_2^2 + P_3^2 \right) + P_2^4 \left( P_1^2 + P_3^2 \right) + P_3^4 \left( P_1^2 + P_2^2 \right) \right]
\]

(2.12)

where \( A_1, A_{11}, A_{12}, A_{111}, A_{112}, \) and \( A_{123} \) are the expansion coefficients. The values of these coefficients determine the thermodynamic behavior of the bulk paraelectric and ferroelectric phases as well as the bulk ferroelectric properties, such as the ferroelectric transition temperature, the stability and metastability of the parent paraelectric phase, the spontaneous polarization and the susceptibility as functions of temperature, etc. In this example, \( (P_1 = 0, P_2 = 0, P_3 = 0) \) represents the paraelectric cubic state, \( (P_1 = 0, P_2 = 0, P_3 \neq 0) \) represents a tetragonal ferroelectric state with its tetragonal axis along the \( z \) direction, etc. More examples of free energy models can be found in Chapter 11.

2.3.2.2 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 gradients of the field variables. \( \alpha \)
in Equation (2.1) is the gradient energy coefficient which characterizes the energy penalty due to the field inhomogeneity at the interfaces, i.e. the interfacial energy contribution to the total free energy. Examples of interfaces include solid-liquid interphase boundaries during solidification, grain boundaries, matrix-precipitate interfaces during precipitation reactions and precipitate coarsening, and domain boundaries 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 (Cahn and Hilliard 1958)

\[
\gamma = \frac{\Delta F}{S} = \int \left[ f(c) - f_0(c) + \frac{\alpha}{2} (\nabla c)^2 \right] d^3r
\]

(2.13)

where \(S\) is the interfacial area and \(f_0(c)\) is the bulk equilibrium free energy density as a function of composition represented by the common tangent line for a binary system. For an interface at equilibrium, the total gradient energy is half of the total interfacial energy (Cahn and Hilliard 1958). Analytical expressions for the interfacial energy in terms of free energy parameters and the gradient energy coefficients are only available for very simple cases that 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, 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 \(\sigma\), interfacial width \(\lambda\), the double-well depth of a free energy model \(\Delta f\), and the gradient coefficient \(\alpha\) obey the following set of relationships,

\[
\sigma \propto \sqrt{\alpha \Delta f}, \quad \lambda \propto \sqrt{\alpha / \Delta f}, \quad \sigma \sim \lambda \Delta f
\]

(2.14)

The last relation in (2.11) is particularly useful in estimating the interfacial width if one has the knowledge on the bulk thermodynamics of a system and the interfacial energy. It is also easy to understand that why a wider interface may be employed if one can artificially eliminate the chemical free energy contribution, \(\Delta f\), and only use the depth of the double-well potential, as in the Kim’s model, to yield the same interfacial energy (Kim, Kim et al. 1999).

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 (2.13)). Another approach is to add higher order gradient energy terms (Langer 1986; Taylor and Cahn 1998; Abinandanan and Haider 2001; Dreyer and Muller 2003). 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 underline crystalline symmetry.
2.3 Theoretical Fundamentals of the Method

This is particularly true for modeling the anisotropic domain wall energies for ordered intermetallic phases (Braun, Cahn et al. 1997; Wang, Banerjee et al. 1998), ferroelectric domain structures, and ferromagnetic domain structures.

2.3.2.3 Non-local Interactions

The second integral in Equation (2.2) represents the nonlocal contributions to the total free energy from long-range interactions such as elastic interactions, electric dipole-dipole interactions, and electrostatic interactions. These long-range interactions are usually obtained by solving the corresponding mechanical and electrostatic 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,

\[
\frac{\partial \sigma_{ij}}{\partial r_j} = 0 \quad \text{with} \quad \sigma_{ij}(r) = \lambda_{ijkl}(r)\left[\varepsilon_{kl}(r) - \varepsilon_{kl}^0(r, \phi, P, \ldots)\right]
\] (2.15)

where \(\sigma_{ij}\) is the local elastic stress, \(r_j\) is the \(j\)th component of the position vector, \(r\), \(\lambda_{ijkl}(r)\) 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 \(\varepsilon_{kl}^0\) 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 resulted elastic energy is a function of phase-field variables and thus the microstructure (Khachaturyan 1983). Various levels of approximations and different approaches have been proposed to solve the elasticity Equation (2.15) with an arbitrary distribution of eigenstrains, i.e. microstructure. For the case of homogeneous approximation and periodic boundary conditions, it was shown by Khachaturyan (Khachaturyan and Shatalov 1969) 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 (Onuki 1989; Nishimori and Onuki 1990; Sagui, Orlikowski et al. 1998). 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 (Leo, Lowengrub et al. 1998; Hu and Chen 2001; Zhu, Chen et al. 2001; Wang, Jin et al. 2002). More extensive discussions on the solution to the mechanical equilibrium equation will be discussed in Chapter 11.

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

\[
\frac{\partial D_i}{\partial r_i} = 4\pi \rho_f \quad \text{or} \quad \frac{\partial E_i}{\partial r_i} = 4\pi \rho_t
\] (2.16)

where \(D_i\) and \(E_i\) are the \(i\)th components of the electric displacement field and the electric field, respectively, \(\rho_f\) is the local free charge density and \(\rho_t\) is the local total charge density which includes the effective charge associated with electric polarizations.

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 functions of both the
volume fraction and morphologies of the coexisting phases or domains (Khachaturyan 1983; Chen and Khachaturyan 1993; Bishop, Garcia et al. 2003). Therefore, they very often play dominant roles in the microstructure evolution in solid states.

### 2.3.3 The Evolution Equations

In all phase-field models, the temporal and spatial evolution of the field variables follows the same set of kinetic equations. All conserved fields, $c_i$, evolve with time according to the Cahn-Hilliard equation (Cahn 1961), or simply the diffusion equation in the case that no gradient energy is introduced for the conserved variable, whereas the non-conserved fields, $\eta_p$, are governed by the Allen-Cahn equation (Allen and Cahn 1977), i.e.

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

(2.17)

\[
\frac{\partial \eta_p(r,t)}{\partial t} = -L_{pq} \frac{\delta F}{\delta \eta_p(r,t)}
\]

(2.18)

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 (2.2).

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 (Caginalp and Xie 1993) matches the phase-field parameters at the limit of zero interfacial thickness to experimentally measured thermodynamic and kinetic properties while a thin-interface analysis (Karma and Rappel 1996; Kim, Kim et al. 1999; Karma 2001) allows the variation of the phase-field variable over a certain thickness for the interface. It is shown by Karma (Karma and Rappel 1996) that a phase-field simulation using the thin-interface asymptotics permits one to use a larger interface width and thus a larger grid size.

By numerically solving the systems of Cahn-Hilliard diffusion equations and the Allen-Cahn relaxation equations subject to appropriate initial and boundary conditions, the evolution profiles of the field variables, and thus the microstructure evolution can be obtained. 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 approaches such as the semi-implicit Fourier Spectral method (Chen and Shen 1998; Zhu, Chen et al. 1999), and adaptive-grid finite-element method (Provatas, Goldenfeld et al. 1998; Jeong, Goldenfeld et al. 2001; Jeong, Dantzig et al. 2003; Lan and Chang 2003).
2.4 Advantages and Disadvantages of the Method

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 impossible. 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. As it will be shown in the next section, it has been successfully applied to solidification, solid-state phase transformations, coarsening and growth, and many others (Table 2.1 in the next section). Third, it 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 rather 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 (Bottger, Grafe et al. 2000; Grafe, Bottger et al. 2000; Grafe, Bottger et al. 2000; Chen, Wolverton et al. 2001; Zhu, Liu et al. 2002; Kobayashi, Ode et al. 2003) or to derive the free energy models from microscopic models (Vaks 2002; Ohno and Mohri 2003). 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 of Karma (Karma and Rappel 1996) and the model by Kim (Kim, Kim et al. 1999) are 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.

2.5 Typical Fields of Applications and Examples

Existing phase-field applications have been focused on three major materials processes: solidification, solid-state phase transformation, and grain growth and coarsening. There are three separate chapters which are devoted to each of these three processes, and hence no detailed discussions and references will be provided in this introduction. Since most of the phase-field applications prior to 2002 were summarized in a review paper by the author, instead of repeating the description of examples, examples pertaining to these three areas and other applications are updated in Table 2.1 from a table in reference (Chen 2002).
2.5 Typical Fields of Applications and Examples

<table>
<thead>
<tr>
<th>Table 2.1: Applications of the phase-field method</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Solidification</strong></td>
</tr>
<tr>
<td>Multicomponent alloys</td>
</tr>
<tr>
<td>Nonisothermal solidification</td>
</tr>
<tr>
<td><strong>Solid State Phase Transformations</strong></td>
</tr>
<tr>
<td>Spinodal phase separation</td>
</tr>
<tr>
<td>Precipitation of cubic ordered intermetallic precipitates from a disordered matrix</td>
</tr>
<tr>
<td>Hexagonal to orthorhombic transformations</td>
</tr>
<tr>
<td>Ferroelectric transformations</td>
</tr>
<tr>
<td>Phase transformations under an applied stress</td>
</tr>
<tr>
<td>Martensitic transformations in single and polycrystals</td>
</tr>
<tr>
<td><strong>Coarsening and Grain Growth</strong></td>
</tr>
<tr>
<td>Coarsening</td>
</tr>
<tr>
<td>Grain growth in a single-phase solid</td>
</tr>
</tbody>
</table>
2 Introduction to Phase-field Method of Microstructure Evolution

Grain growth in a two-phase solid  
(Chen and Fan 1996; Nestler 1999)

Anisotropic grain growth  
(Kazaryan, Wang et al. 2002)

Other Applications

Phase transformations in thin films  
(Leonard and Desai 1998; Leo and Johnson 2001; Li, Hu et al. 2001)

Surface-stress induced pattern formation  
(Lu and Suo 2001)

Spiral growth  
(Karma and Plapp 1998)

Crystal growth under stress  
(Muller and Grant 1999; Kassner, Misbah et al. 2001; Yeon, Cha et al. 2003)

Solute-dislocation interactions  
(Leonard and Desai 1998; Hu and Chen 2001)

Dislocation dynamics  

Crack propagation  

Electromigration  
(Mahadevan and Bradley 1999; Bhat, Bower et al. 2002)

2.6 Summary and Opportunities

Phase-field models have been applied to many different materials processes as shown in Table 2.1. They have played critical roles in the fundamental understanding of the underlying thermodynamic driving forces and the kinetic mechanisms leading to many fascinating morphological evolution phenomena, e.g. those induced by strain effects in coherent microstructures. Existing efforts have been focused on developing new phase-field models for new applications, establishing the relationships between sharp-interface or thin-interface descriptions and phase-field equations, studying microstructure evolution in relatively simple pure and binary systems. Some of the future efforts are expected to be focused on the following areas:

- Various novel applications of the phase-field method will continue to be explored, for example, to problems involving simultaneous long-range elastic and electric or magnetic dipole-dipole interactions, low-dimensional systems such as thin films and multiplayer structures, interactions between phase and defect microstructures such as random defects and dislocations, etc.

- There will also be increasing efforts in establishing schemes to obtain the phase-field parameters directly from more fundamental first-principles electronic structure or atomic calculations. For example, recently it was shown that it is possible to obtain all the necessary thermodynamic information for the input to a phase-field model from first-principles.
2.6 Summary and Opportunities

calculations combining with cluster expansions; it includes the bulk free energies of matrix and precipitate phases, the interfacial energy and its anisotropy, and the lattice mismatch (Vaithyanathan, Wolverton et al. 2002). 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 (Hoyt, Sadigh et al. 1999; Hoyt, Asta et al. 2001) for input to phase-field simulations of solidifications (Bragard, Karma et al. 2002).

- For applications to more complicated systems such as multicomponent alloys, it is essential to interface phase-field models to existing or future thermodynamic, kinetic, and crystallographic databases. It is possible to directly construct the free energy function of a phase-field model from existing databases using the CALPHAD method (Bottger, Grafe et al. 2000; Grafe, Bottger et al. 2000; Grafe, Bottger et al. 2000; Chen, Wolverton et al. 2001; Zhu, Liu et al. 2002; Kobayashi, Ode et al. 2003). The compositional dependence of atomic mobilities from databases can also be incorporated. 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. With independently assessed reliable databases, it will be possible to predict the microstructure evolution in complex multicomponent alloys using the phase-field method.

- As briefly mentioned above, three-dimensional phase-field simulations are very computational intensive, and thus it calls for more efficient numerical algorithms. Furthermore, many of the practical microstructures have their compositional or structural domains 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 or the model proposed by Kim et al. (Kim, Kim et al. 1999), 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.

- There have been a number of existing studies on nucleation and growth using the phase-field method (Roy, Rickman et al. 1998; Simmons, Shen et al. 2000; Granasy, Borzsonyi et al. 2002; Castro 2003). Clearly, significant effort is still required to establish a robust, physical, and quantitative approach to induce both homogenous 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 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.
Acknowledgements

The author is grateful for the financial support from NSF under the grant number DMR-0103354. His research in developing phase-field models for alloy microstructure evolution has also being supported by the NSF-ITR initiative under the grant number DMR-, NASA under the Ultra Efficient Engine Technology Program, as well as Ford Motor Company and ALCOA.

References


References


References


References


References


References


3 Cellular, Lattice Gas, and Boltzmann Automata

Dierk Raabe

The chapter presents concepts for the use of cellular, lattice gas, and Boltzmann automata in materials science. It concentrates on the simulation of recrystallization and materials–related fluid dynamics problems. Cellular automaton models of recrystallization are typically discrete in time, physical space, and orientation space. They can be defined on a 2D or 3D lattice considering different neighbor shells for the calculation of the driving forces and mobilities. The kinetic transformation rules are usually in the form of rate equations for sharp interface motion. The article reviews deterministic and probabilistic automaton formulations. The fluid dynamics simulations discussed in this chapter make use of a discrete automaton form of the Boltzmann transport equation. The method is used for simulating friction effects at lubricated rough surfaces during metal forming processes.

3.1 Cellular Automata

3.1.1 Introduction

Cellular automata are synchronous algorithms that describe the discrete spatial and temporal evolution of complex systems by applying local (or sometimes mid–range) deterministic or probabilistic transformation rules to lattice cells with local connectivity (von Neumann, 1963; Codd, 1968; Wolfram, 1986; Raabe, 1998; Wolfram, 2002).

The concept of cellular automata was introduced in the late 1940’s by von Neumann and Ulam for the simulation of self–reproducing Turing automata and population evolution (von Neumann, 1963; Toffoli and Margolus, 1987). In his early contributions von Neumann denoted these algorithms as cellular spaces. Other authors used notions like tessellation automata or iterative arrays. Early applications were particularly in the fields of fluid dynamics, biological processes, and reaction–diffusion systems (Wolfram, 1986). During the last decade cellular automata gained momentum for the field of microstructure simulation (Raabe, 1998, 2002).

Cellular automata do not have restrictions in the type of elementary entities or rules they use. They can map such different situations as the distribution of the values of state variables in a finite difference simulation, the colors in a blending algorithm, the elements of fuzzy sets, or growth and decay processes of cells. For instance, the Pascal triangle can be regarded as a one–dimensional (1D) cellular automaton where the value of a lattice point of a regular triangular lattice is calculated through the summation of the two numbers above it. In this case the entities of the automaton are integer numbers and the rule is a summation.
3.1 Cellular Automata

The space variable in cellular automata usually represents real space, but orientation space, momentum space, or wave vector space are used as well (Vichiniac, 1984; Raabe, 1998; Succi, 2001; Wolfram, 2002). Cellular automata can have arbitrary dimension. The lattice of the automaton maps the elementary system entities that are regarded as relevant to the model. The individual lattice points can represent continuum volume units, atomic particles, larger groups of atoms, cars, populations, lattice defects, or colors depending on the model. The state of each lattice point is characterized in terms of one or more generalized state variables. These can be dimensionless numbers, particle densities, lattice defect quantities, crystal orientation, particle velocities, blood pressure, animal species or any other quantity the model requires. The actual values of the state variables assume at each cell one out of a finite set of discrete states. The opening state of the automaton consists in the initial distribution of the values of the state variables for a chosen lattice.

The dynamical evolution of the automaton takes place through the application of deterministic or probabilistic transformation rules that act on the state of each lattice point. These rules determine the state of a cell as a function of its previous state and the state of its neighboring cells. The number of neighbor sites considered by the transformation rule when calculating a state switch determines the range of the interaction and the local evolution of the automaton. Cellular automata work in discrete time steps. After each time interval the values of the state variables are updated for all lattice points in synchrony. Owing to these features, cellular automata provide a discrete method of simulating the evolution of complex dynamical systems which contain similar components on the basis of simple local interactions. The cellular automaton method is, therefore, sometimes referred to as a counterpart to partial differential equations, which have the capability to describe dynamical systems in a continuous fashion. The meaning of discrete in this context is, that space, time, and properties of an automaton can assume only a finite, countable number of states. The basic rational of cellular automata is not to try to describe the evolution of complex systems by a global viewpoint, i.e. by using averaging differential equations, but by simulating such systems on the basis of the elementary dynamics of the interacting constituents following simple generic rules. In other words the cellular automaton approach pursues the goal to let the complexity of dynamic systems emerge by the interaction of simple individuals following trivial rules.

3.1.2 Formal Description and Classes of Cellular Automata

The local interaction of neighboring lattice points in a cellular automaton is specified through a set of transformation rules. While von Neumann’s original automata were designed with deterministic transformation rules probabilistic transformations are conceivable as well. The value of an arbitrary state variable $\xi$ assigned to a particular lattice site at a time $(t_0 + \Delta t)$ is determined by its present state $(t_0)$ (or its last few states $t_0, t_0 - \Delta t$, etc.) and the state of its neighbors (von Neumann, 1963). Using the last two time steps for the evolution of a 1D cellular automaton can be put formally by writing $\xi_j^{t_0 + \Delta t} = f(\xi_j^{t_0 - \Delta t}, \xi_{j-1}^{t_0 - \Delta t}, \xi_{j+1}^{t_0 - \Delta t}, \xi_{j-1}^{t_0}, \xi_{j+1}^{t_0})$ where $\xi_j^{t_0}$ is the value of the variable at time $t_0$ at node $j$. The positions $(j + 1)$ and $(j - 1)$ indicate the nodes in the neighborhood of position $j$ (in 1D). The function $f$ specifies the set of transformation rules. If the state of the node depends only on its nearest neighbors (NN) the array is referred to as von Neumann neighboring. If both the NN and the next-nearest
neighbors (NNN) determine the ensuing state of the node, the array is called Moore neighboring. The Extended Moore Neighborhood is equivalent to the Moore neighborhood, however, it reaches over the distance of the next adjacent cells, i.e. it considers two neighbor shells (Figure 3.1). Another approach of defining a local neighborhood is the Margolus rule which considers blocks of neighbor cells (e.g. 2 x 2) of a lattice at once. The cell neighboring affects the transformation rates and the evolving morphologies. For extended configurations (e.g. Moore) the transformation rule can in 1D and for interaction with the last two time steps be written as
\[ \xi_{j_0+\Delta t} = f(\xi_{j-n}, \xi_{j-n+1}, \ldots, \xi_{j-1}, \xi_{j+1}, \xi_{j+2}, \xi_{j+3}, \ldots, \xi_{j+n-1}, \xi_{j+n}) \]
where \( n \) indicates the range of the transformation rule in units of lattice cells.

Even for very simple automata there exists an enormous variety of possible transformation rules. If in a 1D Boolean cellular automaton with von Neumann neighboring and reference to the preceding time step each node can have one of two possible ground states, say \( \xi_j = 1 \) or \( \xi_j = 0 \), the transformation rule assumes the form
\[ \xi_{j_0+\Delta t} = f(\xi_{j-1}, \xi_j, \xi_{j+1}). \]
This simple Boolean configuration defines 2^8 possible transformation rules. One of them has the form

if \( (\xi_{j-1} = 1, \xi_j = 1, \xi_{j+1} = 1) \) then \( \xi_{j_0+\Delta t} = 0 \) (1, 1, 1) \( \sim \) 0
if \( (\xi_{j-1} = 1, \xi_j = 1, \xi_{j+1} = 0) \) then \( \xi_{j_0+\Delta t} = 1 \) (1, 1, 0) \( \sim \) 1
if \( (\xi_{j-1} = 1, \xi_j = 0, \xi_{j+1} = 1) \) then \( \xi_{j_0+\Delta t} = 0 \) (1, 0, 1) \( \sim \) 0
if \( (\xi_{j-1} = 1, \xi_j = 0, \xi_{j+1} = 0) \) then \( \xi_{j_0+\Delta t} = 1 \) (1, 0, 0) \( \sim \) 1
if \( (\xi_{j-1} = 0, \xi_j = 1, \xi_{j+1} = 1) \) then \( \xi_{j_0+\Delta t} = 1 \) (0, 1, 1) \( \sim \) 1
if \( (\xi_{j-1} = 0, \xi_j = 1, \xi_{j+1} = 0) \) then \( \xi_{j_0+\Delta t} = 0 \) (0, 1, 0) \( \sim \) 0
if \( (\xi_{j-1} = 0, \xi_j = 1, \xi_{j+1} = 0) \) then \( \xi_{j_0+\Delta t} = 1 \) (0, 0, 1) \( \sim \) 1
if \( (\xi_{j-1} = 0, \xi_j = 1, \xi_{j+1} = 0) \) then \( \xi_{j_0+\Delta t} = 0 \) (0, 0, 0) \( \sim \) 0

This particular transformation rule can be encoded by (01011010)_2. Its digital description is of course only valid for a given arrangement of the corresponding basis. This order is commonly chosen as a decimal row with decreasing value, i.e. (1, 1, 1) translates to 111 (one hundred eleven), (1, 1, 0) to 110 (one hundred ten), and so on. Transforming the binary code
into decimal numbers using

\[
\begin{array}{cccccccc}
2^7 & 2^6 & 2^5 & 2^4 & 2^3 & 2^2 & 2^1 & 2^0 \\
0 & 1 & 0 & 1 & 1 & 0 & 1 & 0 \\
\end{array}
\]

yields the decimal code number 90\textsubscript{10}. The digital coding system is commonly used for compactly describing transformation rules for cellular automata.

The number of rules can be calculated by \( k^{(kn)} \), where \( k \) is the number of states for the cell and \( n \) the number of neighbors including the core cell. For a 2D automaton with Moore neighborhood \((n = 9)\) and two possible cell states \((k = 2)\) a set of \(2^{29} = 262144\) different transformation rules exist.

If the state of a node is determined by the sum of the values of the neighbor sites, the model is referred to as totalistic cellular automaton. If the state of a node has a separate dependence on the state itself and on the sum of the values taken by the variables of the neighbors, the model is referred to as outer totalistic cellular automaton.

According to Wolfram (1986, 2002) cellular automata fall into four classes of behavior. Class 1 cellular automata evolve for almost any initial configuration after a finite number of time steps to a homogeneous and unique state from which they do not evolve further. Cellular automata in this class exhibit the maximal possible order both at the global and local scale. The geometrical analogy for this class is a limit point in the corresponding phase space. Class 2 cellular automata create short period patterns that repeat periodically, typically recurring after small periods, or are stable. Local and global order is exhibited in such automata. Class 2 automata can be interpreted as filters, which derive the essence from discrete data sets for a given set of transformation rules. In phase space such systems form a limit cycle. Class 3 cellular automata lead from almost all possible initial states to aperiodic chaotic patterns. The statistical properties of these patterns and the statistical properties of the starting patterns are almost identical at least after a sufficient period of time. The patterns created by class 3 automata are usually self–similar fractal arrays. After many time steps, the statistical properties of these patterns are typically the same for almost all initial configurations. Geometrically class 3 automata form strange attractors. Class 3 is the most frequent type of cellular automata. Class 4 cellular automata yield stable, periodic, and propagating structures which can persist over arbitrary lengths of time. Some class 4 automata dissolve after a finite steps of time, i.e. the state of all cells becomes zero. In some class 4 a small set of stable periodic figures can occur (such as for instance in Conway’s game of life). By properly arranging these propagating structures, final states with any cycle length may be obtained. Class 4 automata show a high degree of irreversibility in their time development. They usually reveal more complex behavior and very long transient lengths, having no direct analogue in the field of dynamical systems. The cellular automata in this class can exhibit significant local (not global) order.

3.1.3 Cellular Automata in Materials Science

One typical application of cellular automata for materials–related simulations consists in mapping the values of relevant state variables onto the points of a lattice and using the local finite difference formulations of the partial differential equations of the underlying model as transformation rules. The particular versatility of cellular automata for microstructure simulations
especially in the fields of recrystallization (Hesselbarth and Göbel, 1991; Pezzee and Dunand, 1994; Raabe, 2001; Janssens, 2003), dendritic growth (Brown et al., 1994; Spittle and Brown, 1995; Gandin and Rappaz, 1997; Gandin, 2001), grain growth (Liu et al., 1996; Geiger et al., 2001; Janssens et al., 2002), and phase transformation phenomena (Cortie, 1993; Spittle and Brown, 1994; Kumar et al., 1998) is due to its flexibility in considering a large variety of variables and transformations.

Discrete microstructure simulations which track kinetics and energy in a local fashion are of interest for two reasons. First, it is desirable to understand the dynamics of microstructure evolution on the basis of the interaction of large densities of lattice defects which are characterized by a wide spectrum of properties and interactions in heterogeneous materials. For instance, in the fields of recrystallization and grain growth the influence of local grain boundary characteristics, local driving forces, and local texture on the final microstructure is of particular interest. Second, from a practical point of view it is necessary to predict microstructure parameters such as grain size (Davies, 1995; Sheldon and Dunand, 1996; Davies, 1997) or texture (Raabe et al., 1996; Marx, 1997; Marx et al., 1998; Raabe, 1999) which determine the mechanical and physical properties of real materials subjected to industrial processes (Davies, 1999). The strength of cellular automata is that they combine the simplicity and scalability of a switching model (Raabe, 2000) with the physical stringency of a boundary dynamics model (Raabe, 2002).

Cellular automata can use atomic–scale or continuum–scale elements as underlying units. Since continuum formulations are not intrinsically scaled such automata must be physically calibrated by transformation rules which reflect the characteristic length and time scales of the model.

If local transformations in the same model must account for different activation energies and real time scales (e.g. competing recrystallization and recovery phenomena, bulk diffusion and grain boundary diffusion, different interface mobilities) it is required to achieve a correct common scaling of the system. This point is essential since the transformation behavior and the dynamical evolution of microstructure–oriented cellular automaton models are governed by noncoupled Boolean routines rather than by exact local solutions of coupled differential equations. The scaling problem becomes particularly important when simulating nonlinear systems (which applies for most microstructure–based cellular automata). During the simulation it can be useful to refine or coarsen the scale according to the evolving kinetics (time re–scaling) and spatial resolution (space re–scaling) of the system. Further aspects of meshing and scaling are discussed in the chapter of Janssens.

3.1.4 Recrystallization Simulations with Cellular Automata

3.1.4.1 Basic Constitutive Model of the Cellular Automaton

The recrystallization model for the present example is designed as a cellular automaton with a probabilistic transformation rule (Raabe, 1999, 2001, 2002). Independent variables are time $t$ and space $x$. Space is discretized into an array of equal cells (2D or 3D) each of which is characterized in terms of the dependent variables. These are scalar and configurational contributions to the driving force and the crystallographic orientation $\mathbf{g} = g(\varphi_1, \phi, \varphi_2)$, where $g$ is the rotation matrix and $\varphi_1, \phi, \varphi_2$ the Euler angles. The driving force is the negative change
3.1 Cellular Automata

in Gibbs enthalpy $G_t$ per transformed cell. The starting data, i.e. the local orientation and driving forces, can be provided by experiment, i.e. by orientation imaging microscopy (Lücken and Raabe, 2002) or by simulation, e.g. by a crystal plasticity finite element simulation (Raabe and Becker, 2000). Grains or subgrains are mapped as regions of identical crystal orientation, but the driving force may vary inside these areas.

The kinetics of the automaton emerge from changes in the state of the cells. They occur in accord with a transformation rule which determines the switching probability of each cell as a function of its previous state and the state of its neighbor cells. The rule reflects that the state of a non-recrystallized cell belonging to a deformed grain may change due to the expansion of a neighbor grain which grows according to the local driving force and boundary mobility. If such an expanding grain sweeps a non-recrystallized cell the stored dislocation energy of that cell drops to zero and a new orientation is assigned to it, namely, that of the expanding neighbor grain. To put this formally, the switching rule is cast in a probabilistic form of a linearized symmetric rate equation, which describes grain boundary motion in terms of isotropic diffusion processes perpendicular through a homogeneous planar grain boundary segment under the influence of a decrease in Gibbs energy,

$$\dot{x} = \nu D \lambda_{gb} n \exp \left( -\frac{\Delta G - \Delta G_t/2}{k_b T} \right) - \exp \left( -\frac{\Delta G + \Delta G_t/2}{k_b T} \right)$$

(3.1)

where $\dot{x}$ is the velocity of the interface, $\nu D$ the Debye frequency, $\lambda_{gb}$ the jump width through the interface, $c$ the intrinsic concentration of in-plane self diffusion carrier defects (e.g. grain boundary vacancies or shuffle sources), $n$ the normal of the grain boundary segment, $\Delta G$ the Gibbs enthalpy of interface motion, $\Delta G_t$ the transformation enthalpy, $k_b$ the Boltzmann constant, and $T$ the absolute temperature. Bold symbols indicate vectors. The Debye frequency is of the order of $10^{13} - 10^{14} s^{-1}$ and the jump width of the order of the Burgers vector $b$.

Inserting the enthalpy, entropy, and driving pressure, as well as linearization yields

$$\dot{x} = \nu_D \lambda_{gb} n m \exp \left( -\frac{Q_{gb}}{k_b T} \right)$$

(3.2)

where $p$ the driving force, $m$ the mobility, and $Q_{gb}$ the activation energy of grain boundary motion. Equations 3.1 and 3.2 provide a Turnbull–type kinetic picture of grain boundary segment motion, where the atomistic processes (including thermal fluctuations, i.e. random thermal backward and forward jumps) are statistically described in terms of the pre-exponential factor of the mobility and of the activation energy of grain boundary mobility.

For dealing with competing switches which may affect the same automaton cell the deterministic rate equation can be replaced by a probabilistic one which calculates switching probabilities (Raabe, 1999, 2002). For this purpose Equation (3.2) is splitted into a deterministic part, $\dot{x}_0$, which depends weakly on temperature, and a probabilistic part, $w$, which depends strongly on temperature:

$$\dot{x} = x_0 w = n \frac{k_B T m_0}{\Omega} \frac{p \Omega}{k_B T} \exp \left( -\frac{Q_{gb}}{k_B T} \right)$$

(3.3)

with $\dot{x}_0 = n \frac{k_B T m_0}{\Omega}$ and $w = \frac{p \Omega}{k_B T} \exp \left( -\frac{Q_{gb}}{k_B T} \right)$.
where $\Omega$ is the atomic volume. The reason for this form is the separation of the Boltzmann term from the linear temperature term which facilitates scaling and normalization as will be outlined below.

According to this expression non-vanishing switching probabilities occur for cells which reveal neighbors with different orientation and a driving force which points in their direction. The automaton considers the first, second (2D), and third (3D) neighbor shell for the calculation of the driving force acting on a cell. The local value of the switching probability depends on the crystallographic character of the boundary segment between such unlike cells.

### 3.1.4.2 Scaling and Normalization

Microstructure based cellular automata are usually applied to starting data which have a spatial resolution far above the atomic scale. This means that the lattice has a lateral scaling of $\lambda_m \gg b$ where $\lambda_m$ is the characteristic length scale of the cellular automaton. If a moving boundary segment sweeps a cell, the grain thus grows (or shrinks) by $\lambda_m^3$ rather than $b^3$. Since the real velocity of an interface must be independent of this artificial scaling introduced by the lattice, an increase in jump width must lead to a corresponding decrease in the grid attack frequency, i.e. to an increase of the characteristic time step, and vice versa. For obtaining a scale-independent boundary velocity, the grid frequency must be chosen in a way to ensure that the attempted switch of a cell of size $\lambda_m$ occurs with a frequency much below the atomic attack frequency which attempts to switch a cell of length $b$. This condition which is prescribed by a scaling length $\lambda_m$ leads to the equation

$$\dot{x} = \dot{x}_0 w = n (\lambda_m \nu) w \quad \text{with} \quad \nu = \frac{k_B T m_0}{\Omega \lambda_m}$$

where $\nu$ is the eigenfrequency of the chosen lattice characterized by the scaling length $\lambda_m$. The eigenfrequency represents the basic attack frequency for one particular grain boundary with constant mobility. In order to use a whole spectrum of mobilities and driving forces in one simulation it is necessary to normalize the eigenfrequency by a common grid attack frequency $\nu_0$,

$$\dot{x} = \dot{x}_0 w = n \lambda_m \nu_0 \left( \frac{\nu}{\nu_0} \right) w = \dot{x}_0 \left( \frac{\nu}{\nu_0} \right) w = \dot{x} \tilde{w}$$

(3.5)

The value of $\nu_0$ which is characteristic of the lattice can be calculated according to the condition that the maximum occurring switching probability cannot exceed 1

$$\tilde{w}^{\max} = \frac{m_0^{\max} p^{\max}}{\lambda_m \nu_0^{\min}} \exp \left( - \frac{Q_{gb}^{\min}}{k_B T} \right) \leq 1$$

(3.6)

where $m_0^{\max}$ is the maximum occurring pre-exponential factor of the mobility, $p^{\max}$ the maximum possible driving force, $\nu_0^{\min}$ the minimum allowed grid attack frequency, and $Q_{gb}^{\min}$ the minimum occurring activation energy. With $\tilde{w}^{\max} = 1$ one obtains the normalization frequency as a function of the upper bound input data.

$$\nu_0^{\min} = \frac{m_0^{\max} p^{\max}}{\lambda_m} \exp \left( - \frac{Q_{gb}^{\min}}{k_B T} \right) \leq 1$$

(3.7)
This frequency and the local values of the mobility and the driving force lead to

\[ \hat{\omega}_{\text{local}} = \frac{m_{\text{local}} p_{\text{local}}}{\lambda m \nu_0} \exp \left( -\frac{Q_{\text{gb}}}{k_B T} \right) = \frac{m_{\text{local}} p_{\text{local}}}{m_{\text{max}} p_{\text{max}}} = \frac{|\dot{x}_{\text{local}}|}{|\dot{x}_{\text{max}}|} \]

(3.8)

This expression is the central switching equation of the algorithm. One can interpret this equation also in terms of the local time \( t = \lambda m / \dot{x} \) which is required by a grain boundary with velocity \( \dot{x} \) to sweep an automaton cell of size \( \lambda m \).

Equation (3.8) shows that the local switching probability can be quantified by the ratio of the local and the maximum mobility \( m_{\text{local}} / m_{\text{max}} \) which are functions of the grain boundary character and by the ratio of the local and the maximum driving pressure \( p_{\text{local}} / p_{\text{max}} \). The probability of the fastest occurring boundary segment (characterized by \( m_{\text{local}} = m_{\text{max}} \), \( p_{\text{local}} = p_{\text{max}} \), and \( Q_{\text{gb}} = Q_{\text{min}} \)) to realize a cell switch is equal to 1. The Equation also shows that an increasing cell size leaves the switching probability unaltered, but the real time step elapsing during an attempted switch is changed. This relationship is obvious since the volume to be swept becomes larger which requires more time. The time constant of the simulation \( \Delta t \) is \( 1/\nu_{\text{min}} \).

While Equation (3.8) allows one to calculate the switching probability of a cell as a function of its previous state and the state of the neighbor cells the actual decision about a cell switch is made by a Monte Carlo step. The use of random numbers ensures that all cell switches are sampled according to their proper statistical weight, i.e. according to the local driving force and local mobility between cells. The simulation proceeds by calculating the individual local switching probabilities for each cell and evaluating them using a Monte Carlo algorithm. This means that for each cell the calculated switching probability is compared to a randomly generated number \( r \) which lies between 0 and 1. The switch is accepted if the random number is equal or smaller than the calculated switching probability. Otherwise the switch is rejected.

\[ \text{random number } r \text{ between 0 and 1} \begin{cases} \text{accept switch if } r \leq \left( \frac{m_{\text{local}} p_{\text{local}}}{m_{\text{max}} p_{\text{max}}} \right) \\ \text{reject switch if } r > \left( \frac{m_{\text{local}} p_{\text{local}}}{m_{\text{max}} p_{\text{max}}} \right) \end{cases} \]

(3.9)

Except for the probabilistic evaluation of the analytically calculated transformation probabilities, the approach is entirely deterministic. Thermal fluctuations are included via Turnbull’s rate equation. The use of realistic or experimental input data for the grain boundaries leads to predictions on a real time and space scale. The switching rule is scalable to any mesh and to any spectrum of boundary properties.

### 3.1.4.3 Kinetics and Topology of Primary Static Recrystallization

Figure 3.2 shows the kinetics for a number of 3D recrystallization simulations with site saturated nucleation conditions and an initial dislocation density of \( 10^{15} \text{ m}^{-2} \) in aluminum. The grid size was \( 10 \times 10 \times 10 (\mu \text{m})^3 \). The cell size was 0.1 \( \mu \text{m} \). All grain boundaries had the same mobility using an activation energy of the grain boundary mobility of 1.3 eV and a pre–exponential factor of the boundary mobility of \( m_0 = 6.210^{10} \text{ m}^3/(\text{N s}) \). Small angle grain boundaries had a mobility of zero. The temperature was 800 K. The time constant of
the simulation was 0.35 s. The different curves correspond to different initial numbers of nuclei (from 165 to 9624) which were statistically distributed in physical and orientation space. The curves show a typical Avrami shape. The logarithmic plot reveals Avrami exponents between 2.86 and 3.13 which is in good accord with the analytical value of 3.0 for site saturated conditions in 3D. The simulations with a very high initial density of nuclei reveal a more pronounced deviation of the Avrami exponent with values around 2.7 during the beginning of recrystallization which is due to lattice effects: While the Avrami equation assumes a vanishing volume for newly formed nuclei the cellular automaton has to assign one lattice point to each new nucleus. Figure 3.3 shows the effect of grain boundary mobility on growth selection. One grain boundary had a larger mobility than the others (activation energy of the mobility of 1.35 eV instead of 1.40 eV) and consequently grew much faster than the neighboring grains which finally ceased to grow. The grains in this simulation all grew into a heavily deformed single crystal.

Figure 3.2: Kinetics for different recrystallizing aluminum crystals. The different curves correspond to different initial numbers of nuclei.

Figure 3.3: Effect of grain boundary mobility on growth selection. All grains grow into a deformed single crystal. One grain boundary has a larger mobility than the others (activation energy of the mobility of 1.35 eV instead of 1.40 eV) and grows faster than the neighboring grains.
3.1 Cellular Automata

3.1.4.4 Coupling Cellular Automata with Finite Element Models

This section shows how the results obtained from a crystal plasticity finite element simulation can be used as a starting microstructure for a cellular automaton recrystallization simulation (Raabe and Becker, 2000; Raabe, 2000). The finite element model was used to simulate plane strain compression of aluminum to a logarithmic strain of 0.434. The values of the state variables (dislocation density, crystal orientation) were mapped on the regular lattice of a 2D cellular automaton. The original size of the specimen which provided the input microstructure to the crystal plasticity finite element simulation gave a lattice point spacing of $\lambda_m = 61.9 \mu m$. The maximum occurring driving force according to the stored dislocation density amounted to about 1 MPa. The temperature dependence of the shear modulus and of the Burgers vector was considered in the calculation of the driving force. The activation energy of the grain boundary mobility was 1.46 eV and its pre–exponential factor was $m_0 = 8.3 \times 10^{-3} \text{m}^3/(\text{Ns})$. Small angle grain boundaries were immobile. Together with the scaling length $\lambda_m$ these data were used for the calculation of the time step and of the local switching probabilities.

![Figure 3.4: 2D simulations of recrystallization in aluminum applied to crystal plasticity finite element data. The upper row show the texture and the lower row the dislocation density. The white areas indicate a fully recrystallized state. The black lines indicate misorientations above 15° and the thin gray lines between 5° and 15°. a) 0% recrystallized; b) 22% recrystallized; c) 48% recrystallized. (800 K)](image)

Nucleation was treated in a phenomenological fashion using the kinetic and thermodynamic instability criteria from classical recrystallization theory. The kinetic instability criterion means that nucleation includes the formation of a large angle grain boundary. The thermodynamic criterion means that the stored energy changes across the grain boundary providing a driving force into deformed grains. Nucleation in this simulation is performed in accord with these two criteria. The used nucleation model does not create any new orientations. At the beginning of the simulation the thermodynamic criterion, i.e. the local value of the dislocation density was first checked for all lattice points. If the dislocation density was larger than a critical value the cell was spontaneously recrystallized without orientation change, i.e. a dislocation density of zero was assigned to it and the original orientation was preserved. In the next step the growth algorithm was started according to the above equations, i.e. the kinetic conditions for nucleation were checked by calculating the misorientations among all spontaneously recrystallized cells (preserving their original crystal orientation) and their neighborhood. If any pair of cells revealed a misorientation above 15°, the cell flip of the unrecrystallized cell was calculated according to its actual transformation probability. For a successful cell flip the orientation of the recrystallized neighbor cell was assigned to the flipped cell (Raabe, 2001,
3 Cellular, Lattice Gas, and Boltzmann Automata

2002). Figure 3.4 shows simulated microstructures for site saturated spontaneous nucleation in all cells with a dislocation density larger than 70% of the maximum value.

3.2 Cellular Automata for Fluid Dynamics

3.2.1 Introduction

Fluid dynamics offers many challenges in modern materials science. The fundamentals of fluid motion were already in 1823 studied by Navier and Stokes. Their formulation accounts for the conservation of mass, momentum, and energy, and the requirement that these quantities be conserved locally. Although the Navier–Stokes equation acts as a well–confirmed basis for predicting fluid behavior it has still not been possible to completely resolve some basic questions of fluid dynamics such as fluid behavior under complicated boundary conditions or turbulence. These problems are typically associated with the highly non–linear character of the Navier–Stokes equation under such circumstances. Introducing instead cellular automaton approaches which use the flow vectors of simple coarse grained fluid portions offers a means of rendering the conservation and translation laws underlying the Navier–Stokes equation into the form of discrete operators which work on a regular lattice and use simple local rules. Even Navier himself had actually reasoned about the local dynamics of small discrete portions of fluid elements.

The procedures of motion and collision of such discrete units of fluid is taken literally by fluid dynamics cellular automata. In these simplified models, fluid portions which assume a certain velocity reside on a lattice in 2D or 3D. In each time step they move according to their current momentum vector. If two particles happen to end up on the same lattice site, they collide and change their velocities according to some fixed collision rule. The only restriction is that collisions have to conserve the particle number and they must conserve momentum and energy, too. These vector automata are referred to as lattice gas automata (including special forms such the HPP and FHP lattice gas automata as well as Boltzmann lattice automata).

Materials science often requires profound knowledge about fluids such as in the realms of liquid materials processing and tribology. Prominent examples are in the fields of melting, casting, two–phase fluid flows, solidification, fluid percolation of cellular structures, liquid polymer processing, lubrication in metal forming, friction, or the processing and infiltration of metallic foams.

The introduction of robust cellular automaton methods designed to tackle such highly nonlinear and complex issues has created a new field which could be termed discrete computational materials fluid dynamics. The following sections will give introduction to this new and exiting science.

3.2.2 The HPP and FHP Lattice Gas Cellular Automata

Lattice–gas models are cellular automata for fluid dynamics. They are discrete in both space and time, assume discrete states, and are able to mimic complex dynamical behavior on the basis of simple local translation, reaction, and diffusion rules. These rules simulate – in a simplified and averaging fashion – some of the microscopical effects occurring in a real fluid. This
means that lattice–gas automata take a microscopic view of continuum fluid mechanics. In the macroscopic view, one analyses the pressure, density, viscosity, and velocity of the fluid flow. In the microscopic view taken by an automaton approach, discrete sets of pseudo–particles move with different speeds in different directions on a fixed lattice. The total effect of these pseudo–particles is the pressure and temperature which can be measured in the macroscopic view. It must be noted though that lattice–gas automata themselves are averaging methods, i.e. the fluid portions which they use as elementary constituents are averaged particles which do not perform individual Newtonian dynamics.

**Figure 3.5:** 2D lattice gas cellular automaton on a hexagonal grid. The diagram shows two consecutive particle arrangements without collisions. The momentum (magnitude of particle velocities), the flight directions, and the mass (particle number) are preserved. In the next step, collisions are carried out in accord with a set of collision rules. Collisions occur for particle groups (a,c,d), (e,f), and (g,h).

The main difference to the automata discussed in the preceding sections is that lattice gas automata describe the state of a node (or bond) in terms of discrete moving particles, i.e. the state variable of a lattice gas cellular automaton is a momentum vector. The particles are usually characterized by a certain velocity, mass, and interaction energy. The number of particles is conserved during the simulation. The particles move from site to site on their regular lattice and obey simple conservative scattering rules when they collide. Lattice–gas cellular automata show that simple and highly idealized microscopic dynamics can reproduce certain aspects of isotropic macroscopic hydrodynamics. They allow to perform simulations with many more particles than molecular dynamics models in which one has to take into account the continuous values of positions, velocities and interaction potentials with finite precision. In microstructure simulation lattice gas cellular automata have gained momentum especially for addressing materials–related Navier–Stokes-type problems, such as plain fluid dynamics, diffusive phenomena, fluid phase separation, hydrodynamic interfaces, multiphase flow, flow through porous media, phase separation, and interface dynamics.

Lattice gas cellular automata were originally suggested by Hardy, de Pazzis and Pomeau (thus HPP–model) for the simulation of long–term behavior of complex reaction–diffusion systems (Hardy et al., 1973). The HPP lattice gas model is based on a discrete square lattice
whose nodes can be occupied by particles. All particles have the same unit mass and unit velocity. Elementary motion of the particles consists in moving them from one lattice node to their nearest neighbor in one discrete unit of time according to their given unit momentum vector. The HPP model imposes as an exclusion principle that no two particles may sit simultaneously on the same node if their direction is identical. For the square lattice originally suggested by the HPP model, this implies that there can be at most four particles per node. This occupation principle, originally meant to permit simple descriptions for computers, has the direct consequence that the equilibrium–distribution of the particles follows a Fermi-Dirac distribution. Interactions between particles arriving at the same node coming from different directions take place in the form of local instantaneous collisions. The elastic collision rules conserve both mass and momentum.

The actual evolution of the system from one time step to the next takes place in four successive stages: The first step is the advection step or propagation step. It consists in moving all particles from their nodes to their nearest neighbor nodes in the directions of their individual velocity vectors. The second step is the collision step. In this step particles arriving at the same node may exchange momentum if it is compatible with the imposed invariance–rules. In the original HPP-model only two out of the sixteen \((2^4)\) possible pre–collision configurations lead to different post–collision configurations. The other fourteen are transparent collisions which is due to the fact that single particles can not be discerned from each other when they collide. The third step is the bounce–back step. It imposes appropriate boundary conditions for those particles which hit an obstacle. A typical condition is an elastic bounce–back process. The fourth step is the update step. It consists in synchronously mapping all new particle coordinates and velocity vectors obtained from the preceding steps on the lattice and updating the time by one unit.

A limitation of the original HPP model are the few degrees of freedom it offers which is due to the four–fold symmetry of the underlying lattice. This entails artificial anisotropy at the macroscopical scale for situations where the Navier–Stokes equation yields isotropic solutions.
It can be shown that this drawback is resolved when using a lattice with triangular cells and hexagonal symmetry as suggested by Frisch, Hasslacher and Pomeau in their FHP–model (Frisch et al., 1986). The basic design of the FHP–lattice gas automaton is similar to the HPP–model, but the lattice has hexagonal symmetry and offers thus a larger variety of possible collision rules. In a grid with a planar hexagonal symmetry each node can be occupied by up to six particles which can have six different velocity vectors (Figure 3.5). Usually, five different variants of the FHP–model are in use which differ by the number and the type of collisions included in their collision-tables (Frisch et al., 1987). Solutions obtained by use of the FHP automaton reveal a good agreement with numerical solutions of the Navier–Stokes equation, both in laminar and reasonably turbulent flows. When addressing diffusion problems it must be noted though that the transport coefficients of the lattice gas depend in a non–trivial fashion on the choice of the collision rules (Figure 3.6) The transition from the HPP model with a cubic lattice to the FHP model with a hexagonal one came along with the problem of finding a adequate lattice structure for 3D simulations. All known 3D lattice structures did not provide isotropic solutions to fluid dynamics. The problem was finally resolved by using a lattice which was based on the projection of a slide of a lattice formed by four-dimensional cubes (tesseracts) into 3D. Models based on this mesh are referred to as FCHC–models (Face Centered Hyper–Cubic lattice).

More recently thermal variants of the lattice gas cellular automaton have been developed (Ernst, 1992). These models use a discrete spectrum of absolute velocities different from one and an energy conservation law taking this aspect into account. Modified versions of lattice gas automata also consider collision reactions which lead to new species. The post–collision products and configurations arising from all possible different two–, three–, four–, five–, or even six–particle collisions, including those between ordinary particles and reaction products, are usually recorded in a database before the simulation. The information in this list thus reflects the underlying microscopical physical and chemical characteristics of all possible particle interactions.

Owing to the discrete treatment of the pseudo–particles and the discreteness of the collision rules lattice gas automata reveal some intrinsic flaws. First, solutions of lattice gas automata are not Galilean invariant. A Galilean transformation is a change to another inertial reference frame moving with constant velocity. Such transformation should not affect the properties of the flow as it occurs for lattice gas solutions. Second, the fluid pressure does not have the right dependence on the speed of sound in the fluid. Third, the fluctuations in the fluid flow are too large, i.e. the solutions are very noisy, though coarse–graining methods can smooth the results. Fourth, lattice gas automaton simulations give correct solutions only for low Reynolds numbers, i.e. only very viscous fluid flows can be simulated correctly. Fifth, there are unphysical conserved quantities in addition to the physical conserved quantities (mass, momentum, energy, angular momentum). Finally, the collision rules become exponentially complex in many models, including the simplest realistic 3D models.

3.2.3 The Lattice Boltzmann Automaton

3.2.3.1 Introduction

The lattice Boltzmann automaton approach has evolved from the lattice gas models to overcome the shortcomings discussed above (Lallemand and Luo, 2000). The main rationale be-
hind the introduction of the lattice Boltzmann automaton is to incorporate the physical nature of fluids from a more statistical standpoint into hydrodynamics solutions. Compared to conventional lattice gas automata the basic idea of the lattice Boltzmann automaton is to use distribution functions instead of single pseudo–particles. Thus, the particle densities are not boolean, but real numbers (Chopard and Droz, 1998). The underlying equation for the lattice Boltzmann technique is a space–, momentum–, and time–discretized version of the Boltzmann equation which is valid for the description of an ensemble averaged particle density distribution function (Succi, 2001). The kinetic lattice Boltzmann method differs from continuum–based conventional Navier–Stokes solvers which consider homogenized state variables such as pressure, density, and flow velocity. It also differs from earlier versions of lattice gas automata which are based on a local discrete mapping of pseudo–particles and their discrete collision and discrete re-distribution rules rather than on the local evolution of their density distribution function.

3.2.3.2 The Kinetic Boltzmann Equation

The continuum theory of gases and fluids can be viewed at different lateral scales such as the Navier–Stokes scale or the Boltzmann scale above the discrete molecular level (Higuera and Jimenez, 1989; Qian et al., 1992). The Boltzmann equation is a formulation which originates from kinetic theory. This means that the solution of the Newtonian equations of motion for a set of k discrete atoms or rigid molecules with position vectors $\mathbf{x}_k(t)$ and velocities $\dot{\mathbf{x}}_k(t)$ as practiced in the field of molecular dynamics is in the kinetic theory of Boltzmann replaced by a position– and velocity–probability distribution function $f(\mathbf{x}, \mathbf{v}, t)$ which expresses the probability to observe a particle $i$ with position vector $\mathbf{x}_i$ and velocity vector $\dot{\mathbf{x}}_i$ at time $t$ (Gladrow, 1998). The evolution of this probability distribution function obeys the Boltzmann transport Equation (Succi, 2001).

$$\frac{\partial f}{\partial t} + \dot{\mathbf{x}} \cdot \frac{\partial f}{\partial \mathbf{x}} + \mathbf{F} \cdot \frac{\partial f}{\partial \mathbf{v}} = \frac{d f}{d t} \bigg|_{\text{coll}}$$

(3.10)

where the collision term $df/dt$ is determined by the two–body scattering cross section. $\mathbf{F}$ is a body force per mass unit. From this kinetic formulation the Navier–Stokes and continuity equations for fluid flow can be recovered by averaging over small but macroscopic regions of space and over times much longer than the mean free time of molecules.

3.2.3.3 The Lattice Boltzmann Cellular Automaton Method

The discrete counterpart of the Boltzmann transport equation is

$$f_i(\mathbf{x} + \dot{\mathbf{x}}_i \Delta t, t + \Delta t) - f_i(\mathbf{x}, t) = -\Omega_i(f) + \mathbf{F} \Delta t \quad i = 1, \ldots, b$$

(3.11)

where $f_i(\mathbf{x}, t)$ represents a set of $b$ discrete probability density distribution functions to find a particle population with the corresponding set of discrete speeds $\dot{\mathbf{x}}_i$ resident at some discrete cell position $\mathbf{x}$ at time $t$. $\Omega_i(f)$ is a discrete version of the collision operator. The lattice discretization which renders the kinetic Boltzmann equation a lattice equation enters by choosing discrete and appropriate units for $t$ and $x$. Equation (3.11) amounts to replacing the discrete
particles with definite velocities and their discrete collisions as underlying to the lattice gas model by a probability distribution. This means that the boolean occupation state (0 or 1) of a cell in a lattice gas model is in the Boltzmann automaton replaced by a real number for describing the fraction to which each cell is filled. The left hand side of Equation (3.11) represents the unconstrained streaming of the particle populations. The right hand side represents a statistical formulation of the intermolecular collisions and of an external momentum source. Equation (3.11) states that the particle density distribution will in a discrete unit of time be convected (or propagated) to a neighboring site $x + \dot{x}_i \Delta t$. Hence $\dot{x}_i$ is a lattice or link vector and the model is characterized by a finite set of velocities $\dot{x}_i$.

The collision operator, $\Omega_i (f)$, which was boolean in the lattice gas method can be viewed as a term which quantifies the relaxation of the probability density function $f_i (x, t)$ towards its local equilibrium distribution $f_i^{eq} (x, t)$ at some typical relaxation time step $1/\omega$, i.e.

$$\Omega_i (f) = \omega \Delta t \left[ f_i - f_i^{eq} \right]$$  \hspace{1cm} (3.12)

where the local equilibrium distribution amounts to a Maxwellian with local fluid density, speed and temperature. In most current lattice Boltzmann approaches the collision operator, Equation (3.12), is usually formulated as a single step relaxation operation according to the approach of Bhatnagar, Gross, and Krook (BGK) (Bhatnagar et al., 1992).

$$\Omega_i (f) = - \frac{f_i - f_i^{eq}}{\tau}$$  \hspace{1cm} (3.13)

where $\tau$ is the relaxation time.

The mass density $\rho (x, t)$ and the momentum density $j (x, t)$ can be calculated from the distribution function by using

$$\rho (x, t) = \sum_i f_i (x, t) \quad \text{and} \quad j (x, t) = \sum_i \dot{x}_i f_i (x, t)$$  \hspace{1cm} (3.14)

The dynamics of lattice Boltzmann automata is characterized by two basic dynamic stages. The first one is the propagation stage as expressed by the left–hand side of equation 3.11. The
second one is the collision stage (right hand side of Equation (3.11)) which covers the local non–boolean rearrangements of the particle distribution function. The boundary conditions in lattice Boltzmann simulations, when fluid portions hit upon obstacles, can be formulated as a non–slip or as a slip condition. In the first case the velocity field is strictly zero on the obstacle surface. In the second case only the perpendicular component of the velocity field on the obstacle surface is zero. Two examples of the application of the lattice Boltzmann method are given in Figures 3.7 and 3.8.

3.3 Conclusions and Outlook

A review was given on the fundamentals and some applications of cellular and Boltzmann automata in the field of materials research. Special attention was placed on detailing the fundamentals of mapping rate formulations for interfaces, driving forces, and the motion of fluid portions on cellular grids. Applications were discussed from the fields of recrystallization and fluid dynamics in porous media.

The future of the cellular automaton method in the field of mesoscale materials science lies most likely in the discrete simulation of equilibrium and non-equilibrium phase transformation phenomena and fluid dynamics. The particular advantage of automata in this context is their versatility with respect to the constitutive ingredients, to the consideration of local effects, and to the modification of the grid structure and the interaction rules. In the field of phase transformation simulations the constitutive ingredients are the thermodynamic input data and the kinetic coefficients. Both sets of input data are increasingly available from theory.
and experiment rendering cellular automaton simulations more and more realistic. The second advantage, i.e. the incorporation of local effects will improve our insight into cluster effects, such as arising from the spatial competition of expanding neighboring spheres already in the incipient stages of transformations. The third advantage, i.e. the flexibility of automata with respect to the grid structure and the interaction rules is probably the most important aspect for future applications. By introducing more global interaction rules (in addition to the local rules) and long-range or even statistical elements in addition to the local rules for the state update might establish cellular automata as a class of methods to solve some of the intricate scale problems that are often encountered in the materials sciences. It is conceivable that for certain mesoscale problems such as the simulation of transformation phenomena in heterogeneous materials in dimensions far beyond the grain scale cellular automata can occupy a role between the discrete atomistic approaches and statistical Avrami–type approaches. Very similar positive aspects apply for fluid dynamics materials simulations.

The major problem of the cellular automaton method in the field of transformation simulations, both for solids and liquids, is the absence of robust approaches for the treatment of nucleation phenomena. Although basic assumptions about nucleation sites, nucleation rates, and nucleation textures can often be included on a phenomenological basis as a function of the local values of the state variables, physically based phenomenological concepts such as available to a certain extent in the Ginzburg–Landau framework for the case of spinodal decomposition are not yet available for automata. It might hence be beneficial in future work to combine Ginzburg–Landau–type phase field approaches with the cellular automaton method. For instance the (spinodal) nucleation phase could then be treated with a phase-field method and the resulting microstructure could be further examined using a cellular automaton simulation.

References


References

4 The Monte Carlo Method

Anthony D. Rollett and Priya Manohar

4.1 Introduction

This chapter is aimed at describing the Monte Carlo method for the simulation of grain growth and recrystallization. It has also been extended to phase transformations and hybrid versions (Monte Carlo coupled with Cellular Automaton) of the model can also accommodate diffusion. If reading this chapter inspires you to program your own version of the algorithm and try to solve some problems, then we will have succeeded! The method is simple to implement and it is fairly straightforward to apply variable material properties such as anisotropic grain boundary energy and mobility. There are, however, some important limitations of the method that must be kept in mind. These limitations include an inherent lattice anisotropy that manifests itself in various ways. For many purposes, however, if you pay attention to what has been found to previous work, the model is robust and highly efficient from a computational perspective. In many circumstances, it is best to use the model to gain insight into a physical system and then obtain a new theoretical understanding, in preference to interpreting the results as being directly representative of a particular material. Please also keep in mind that the “Monte Carlo Method” described herein is a small subset of the broader use of Monte Carlo methods for which an excellent overview can be found in the book by Landau and Binder (2000).

4.2 History of the Monte Carlo Method

This section describes the history and development of the MC method. The basic features of the model are described as needed.

The Monte Carlo method as known in the materials community is an adaptation of a method used primarily to study the statistical physics of phase equilibria (Landau and Binder 2000). The name “Monte Carlo” was coined by Metropolis (inspired by Ulam’s interest in poker) during the Manhattan Project of World War II, because of the similarity of statistical simulation to games of chance, and because Monte Carlo, the capital of Monaco was a center for gambling (http://csep1.phy.ornl.gov/mc/node1.html). Monte Carlo now refers to any method that utilizes sequences of random numbers to perform statistical simulation. The main requirement to use Monte Carlo method for simulation of a physical system is that it must be possible to describe the system in terms of probability density function (PDF), also called partition function \( Z \). Once the PDF or \( Z \) for a system is known, then the simulation begins by random “sampling” from the PDF, and subsequently determining the desired properties of
4.2 History of the Monte Carlo Method

the sample by conducting some kind of a “trial”. There must be a rule available, based on some reasonable mathematical and/or physical theory, to decide the outcome of such a trial. Many trials are conducted and outcomes of all of these trials are recorded. The final step in the Monte Carlo (MC) method is that the behavior of the overall system is obtained by computing the average of outcomes of the trials conducted.

MC methods are used in many different ways e.g. as a technique of integration of a function, as a way to model stochastic (random) processes, as tool to calculate properties of state such as $E$, $T$, $P$, and $V$, and as a model to simulate a system of interacting particles e.g. ferromagnetic materials. In materials science, however, the MC method has been primarily applied to simulate microstructural evolution where equilibrium occurs on a local basis at best (e.g. at triple junctions). In general, the system is far from equilibrium and we attempt to study the kinetics of the processes that lead to equilibrium as a function of time e.g. grain growth or recrystallization. Although the model has proven to be useful for many different problems, it is important to understand that its ability to simulate physical behavior at the continuum (or mesoscopic) level is heuristic. One notable exception to this remark is recent work that has demonstrated rigorously that interfacial velocity is linearly related to the curvature of the interface (Holm 2002).

4.2.1 Ising and Potts Models

The genesis of the method lies in solid state physics community and the development of models for ferromagnetic materials. The Ising model (1925) represents a magnetized material as a collection of spins where only two states are possible, namely up or down. Potts (Potts 1952) later generalized the Ising model and allowed for $Q$ states for each particle in the system, hence the term “$Q$-state Ising model.” It is the Potts model that has been used most extensively to simulate mesoscopic (where the length scale is of the order of the grain size) behavior of materials such as recrystallization, grain growth and texture evolution. While we will describe the algorithms for solving the Ising model, as the Ising model is the simpler of the two, it should be noted that the same algorithms are equally applicable for solving the Potts model with some modifications to accommodate the $Q$ states.

In both models, neighboring spins interact with each other through a contribution to the system energy: if the spins are the same, no interaction energy is contributed (ground state) whereas a difference in spin leads to a (positive) contribution to the system energy. The system is symmetric in the sense that the minimum energy condition is reached when either all spins point up or all point down. If there are $N$ interacting particles and each particle has two states (“up” or “down” in Ising model) then the system as a whole has $2^N$ possible states in which it can exist. The problem is to be able to determine the behavior of such a system at a temperature $T$ and predict the properties of the system at equilibrium configuration. Consider a system that is in contact with a thermal reservoir with infinite heat capacity so that energy exchange between the system and the reservoir occurs at constant $T$. The precise nature of the reservoir is not very relevant and it can be viewed as consisting of an infinitely large number of copies of the system that we set out to study. For the reservoir as well as for the system, $V$, $N$ and $T$ are fixed and $E$ can vary between 0 and $\infty$. The assembly of systems within the reservoir is referred to as a (macro)canonical ensemble (Pathria 1972). The question now is: what is the probability $P_i$ that the system is in state $i$, with energy $E_i$ at any time $t$? The
Monte Carlo approach consists of generating a series of possible macrostates $i, j, \ldots$ such that the probability $P_i$ that the system is in state $i$ within the state space, is given by an appropriate PDF. Such a distribution is called a canonical distribution given according to:

$$P_i = \frac{1}{Z} e^{-\frac{E_i}{k_B T}},$$  \hspace{1cm} (4.1)$$

where $k_B$ is Boltzmann’s constant, $T$ is the absolute temperature and $Z$ is the partition function given as follows:

$$Z = \sum_i e^{-\frac{E_i}{k_B T}}.$$  \hspace{1cm} (4.2)$$

The next step is to determine equilibrium properties of the canonical ensemble such as energy and magnetization. The energy associated with each state depends on the exchange energy of the particles and interaction of the particles with the external magnetic field. However, in the absence of an external field, system energy is dependent only on the spin exchange energy, $E$:

$$E_i = J \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{z} \left(1 - \delta_{S_i S_j}\right),$$  \hspace{1cm} (4.3)$$

where $J$ is the energy associated with a dissimilar pair of spins, $\delta$ is the Kronecker delta, $S_i$ is the spin on the $i^{th}$ site ($S = \{0, 1\}$), $z$ is the coordination number of each grid point, $N$ is the number of points in the grid and the factor of one-half compensates for counting each pair of spins twice.

For a given temperature $T$ and spin exchange energy $J$, the system will approach an equilibrium configuration with a consequent system energy $E$ and an order parameter $m$ ($m$ is the expectation value of the magnetization of the system at temperature $T$). In the Ising model, $m$ is determined from the following implicit equation:

$$m = \tanh[\beta(zJm + H)],$$  \hspace{1cm} (4.4)$$

where $\beta = (1/k_B T)$, $z$ is the number of nearest neighbor spins and the quantity $H$ is proportional to the external magnetic field, but has units of energy. Equation (4.4) must be solved numerically to obtain the actual value of $m$. It is clear from Equations (4.1)–(4.4) that the behavior and properties of a system in the Ising model depend critically on the values of $J$ and $T$.

It is noted here that the Ising spin model does not determine the dynamics of evolution i.e. there is no reference to how long it takes for the system to approach the equilibrium canonical distribution. The dynamics of the Ising model were subsequently determined by Kawasaki (1972) for a conserved order parameter spin transition and by Glauber (1963) for a non-conserved order parameter approach. The two approaches vary in the way the system’s structure evolution is handled as shown schematically in Figure 4.1.

In Figure 4.1 a simple Ising one-dimensional model is shown schematically. In Figure 4.1a, the second and third spins simply exchange their spin orientations so that the total number of Up and Down spins is not changed during the transition. On the other hand, in
4.2 History of the Monte Carlo Method

Figure 4.1: Schematic representation of (a) order parameter conserved, and (b) order parameter not conserved spin transitions.

Figure 4.1b, the second site reverses its spin to line up in the direction of spins of its neighbors on either side. In this case, the order parameter is not conserved, as there are now 3 Up spins and 1 Down spin in the system. Essentially all applications of the Monte Carlo method to microstructural problems have used the latter, non-conserved Glauber dynamics: one crystal in a polycrystal can obviously grow at the expense of its neighbor such that the volume of a particular orientation is not conserved. Conserved Kawasaki dynamics are useful in magnetism problems where extensions of the Ising model continue to be studied.

For a system with large numbers of particles, it is extremely complex to analytically solve Equations (4.1) – (4.4) above. To deal with this problem, Metropolis et al. (1953) proposed an algorithm to solve these equations using Monte Carlo method. The Metropolis algorithm for the solution of the Ising model using the Monte Carlo method is described in the following section.

4.2.2 Metropolis Algorithm

At this point it is useful to introduce the procedure for changing the state of a MC model. The simulation begins by initializing an array of lattice sites, with nearest neighbor interactions that depend upon the state (or spin) of the neighboring sites. The Metropolis algorithm is then enforced on such an array. The key steps of the Metropolis algorithms are as given below (Landau and Binder 2000):

1. Choose a site $i$ at random
2. Calculate the energy change $\Delta E$ (using Equation (4.3)) associated with changing the spin at the $i^{th}$ site
3. Generate a random number $r$ such that $0 < r < 1$
4. If $r < \exp(-\Delta E/k_B T)$, flip the spin
5. Increment time regardless of whether a site changes its spin or not
6. Go to 2 until sufficient data is gathered.

There are several important points to be noted here to elucidate the procedure. First of all, the changes are made on an individual site basis where the choice of site is random in time (Step 1) in contrast to other models where simultaneous updating of all sites is the norm (finite difference, finite element, phase field, some types of cellular automata). Secondly, the $\Delta E$ mentioned in Step 2 depends upon the states of the neighbors and the number of neighbors...
depends on the assumed lattice structure as discussed in more detail in Section 4.3.4. In Step 4, the transition from one state to another at any given site is calculated by a transition probability given as $\exp(-\Delta E/k_BT)$. This is in accordance to Equation 4.1. There are different schemes to determine the transition probability as discussed below. The transition probability is then compared to a random number that is generated in Step 3. This is a necessary rule to simulate the roll of a dice or the “chance” involved in making the decision to change the state. Finally, the time mentioned in Step 5 is not the wall-clock time of a program run, but the simulation time in some arbitrary units. One iteration of the Metropolis algorithm represents $1/N$ time increments. On average, $N$ iterations are required for each site in the lattice to have a chance to change its state and therefore, 1 unit of system simulation time elapses after $N$ iterations. This unit of simulation time is called Monte Carlo Step (MCS) and represents an integer time increment. An alternative time accounting method is discussed in Section 4.2.3 that uses a continuous time increment.

The particular expression given here is the Metropolis method that is the one most commonly used in materials simulations; see (Landau and Binder 2000).

\[ p(\Delta E) = \begin{cases} 
1 & \text{if } \Delta E \leq 0 \\
\exp(-\Delta E/k_BT) & \text{if } \Delta E > 0 
\end{cases} \] (4.5)

There is also the symmetric method that uses a hyperbolic tangent function. In simulations of domain coarsening, this method yields similar results to the Metropolis method.

\[ p(\Delta E) = \tanh(-\Delta E/kT) \] (4.6)

To gain an intuitive understanding of the transition kinetics, the influence of $\Delta E$ and $T$ on $p(\Delta E)$ is shown in Table 4.1.

The major inefficiency associated with the Metropolis algorithm is that during the late stages of evolution (sparse systems), or for low temperatures, the transition probability approaches 0 at most sites so that the system evolves very slowly and many reorientation attempts are wasted. This characteristic increases computation times. This weakness is substantially mitigated at low temperatures by the $n$-fold way algorithm described in the following section.

### 4.2.3 n-fold Way Algorithm

The Monte Carlo method is very inefficient when applied in its basic form to large data sets because of the sparseness of the problem to be worked. Once much coarsening has occurred in grain growth, or recrystallization is nearly complete, most sites in the lattice are surrounded by sites of the same orientation, i.e. they are in the bulk of a grain. Therefore the probability of their changing orientation is very low and expending computational effort there is wasted. The crucial contribution of Bortz et al. (1975) was to propose a method for eliminating the need to compute unsuccessful changes in orientation. The $n$-fold way algorithm speeds up simulations by eliminating Steps 3 and 4 in the Metropolis algorithm given above. The essential concept is that for a given state of the system, the spin-flip transition probability for each of the lattice sites can be calculated before choosing a site to flip. The spin-flip transition probabilities in a system will have certain specific values depending on the configuration of surroundings and
4.2 History of the Monte Carlo Method

Table 4.1: Influence of $\Delta E$ and $T$ on kinetics and equilibrium state of a system.

<table>
<thead>
<tr>
<th>$T \rightarrow 0$</th>
<th>$\Delta E \geq 0 \Delta E &lt; 0$</th>
<th>$P(\Delta E) \rightarrow 0$</th>
<th>$P(\Delta E) \rightarrow 1$</th>
<th>System is in equilibrium in ground state, all spins point in the direction of $H$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zero Temperature</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$T \rightarrow \infty$</td>
<td>$\Delta E$</td>
<td>$P(\Delta E) \rightarrow 1$</td>
<td>System is in completely disordered state as entropy dominates and spins are randomly oriented w.r.t. $H$</td>
<td></td>
</tr>
<tr>
<td>Infinite Temperature</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$T &lt; T_c$</td>
<td>$\Delta E \leq 0 \Delta E &gt; 0$</td>
<td>$P(\Delta E) \geq 1$</td>
<td>$P(\Delta E) &lt; 1^*$</td>
<td>$T$ is expressed in terms of the ratio $J/k_B$. Variation in $T$ leads to first order (discontinuous) or second order (continuous) transitions</td>
</tr>
<tr>
<td>$T_c$ is a critical temperature below which system orders spontaneously, $T_c = 2.269185 \ J/k_B$ for 2D, square, Ising model</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Finite Temperature</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$T &gt; T_c$</td>
<td>$\Delta E \leq 0 \Delta E &gt; 0$</td>
<td>$P(\Delta E) \geq 1$, $P(\Delta E) &lt; 1$, but greater than $^*$ above</td>
<td>– as above –</td>
<td></td>
</tr>
</tbody>
</table>

choice of lattice type. The energy associated with site $i$ is given according to the following modification of Equation (4.3):

$$E_i = \frac{J}{2} \sum_{j=1}^{z} (1 - \delta_{S_i S_j}) + H (S_i) \tag{4.7}$$

Consider a square lattice Ising model with four nearest neighbors ($z = 4$) and $J = |H| = 1$, and $k_B T = 0.4$. Based on the orientation of the reference spin and those of its neighbors and the direction of the external field there are 10 distinct values of transition probabilities. The transition probabilities are calculated using the Metropolis scheme, Equations (4.5) and (4.7). The different possible values are termed classes in the n-fold way terminology. Three such classes and their corresponding transition probabilities are shown in Figure 4.2. Three examples of different classes and their initial orientation of the external field and spins are shown in Figure 4.2. For Class 1, the external field is in the opposite direction to the spins while the four neighboring spins are in the same direction as the central reference spin leading to $E_{old} = 1$. If the central reference spin were to flip its spin direction, then it would align itself in the direction of the external field, but would be anti-parallel to all four of its neighbors so the new $E_{new}$ would be 4. The net change for this transition in energy ($\Delta E$) would be $4 - 1 = +3$. The probability of this transition is then determined according to Equations (4.5) and (4.7) as shown in Figure 4.2, together with two additional examples. A similar analysis is required for all ten classes of spin configurations.
The next step in the \( n \)-fold way algorithm involves sorting of the lattice sites into these spin-flip transition probability classes. The sum of spin-flip probabilities for the entire system \( Q_n \) is calculated according to

\[
Q_n = \sum_{j=1}^{n} n_j p_j .
\]

where \( n = 1, 2, 3, \ldots, 10 \) are the number of classes of spin-flip transition probabilities, \( n_j \) is the number of sites in \( j^{th} \) class and \( p_j \) is the spin-flip probability for that class. For a 2D isotropic system with \( z = 4 \), there are 10 classes. If anisotropy is introduced between the X- and Y- axes, then there are 18 classes (Novotny 1995). To decide which spin to flip in the \( n \)-fold way algorithm, a random number \( r(0, 1) \) is generated and the \( k^{th} \) class is chosen based on satisfying the following condition:

\[
Q_{k-1} \leq rQ_n < Q_k .
\]

This procedure thus chooses classes according to their weight – classes with high \( p \) are chosen more often, classes with low \( p \) are chosen less often, while those classes with probability \( = 0 \) will not be chosen at all. Thus sites internal to a domain or a grain that will not flip (transition probability \( = 0 \)), are simply never chosen to flip in the \( n \)-fold way, thus avoiding the rejections involved (Metropolis Step 4) in the conventional MC method. Within the chosen class \( k \), the particular site to be flipped is selected randomly from the \( n_k \) sites that belong to that class. Finally, the spin at the chosen site is flipped with probability \( 1 \) and this eliminates Steps 3 and 4 in the Metropolis algorithm. The class of the chosen spin and those of its nearest neighbors is then updated. The \( n \)-fold way algorithm may be described in the following manner (Novotny 1995):

1. Generate a random number and increment the time by an appropriate amount.

2. Choose a class \( k \) that satisfies the condition given in Equation (4.9)
3. Generate a random number to choose one of the sites from class \( k \)

4. Flip the spin at the chosen site with probability 1

5. Update the class of the chosen spin and all of its nearest neighbors

6. Determine \( Q_n \)

7. Go to 1 until sufficient data is gathered

The time increment in the \( n \)-fold way algorithm is explained in detail by Novotny based on the concept of absorbing Markov chains. In effect, the time increment, \( \Delta t \), in the \( n \)-fold way algorithm is correlated to the probability that the given system configuration will change to a different configuration during the time increment:

\[
\Delta t = -\frac{\ln r}{Q_n}. \quad (4.10)
\]

Equation (4.10) is based on the assumption that the successful re-orientation of a site is described by an exponential probability distribution, so that successive evolution steps are Poisson events.

It is useful to see exactly how the \( n \)-fold way speeds up the computation time. Consider class 1 as shown in Figure 4.2. The transition probability for this class is calculated as \( 6.7 \times 10^{-3} \). Hence, on average, about \( 1/p(1) = 148 \) spin flip trials are required for a spin in class 1 to flip successfully using the Metropolis algorithm. For sparse systems, a large proportion of sites have a low transition probability and therefore Metropolis trials require long times in these cases for successful flips whereas the \( n \)-fold way flips those sites in one attempt. One drawback of the \( n \)-fold way algorithm is that it requires more memory than the standard MC algorithm, because it needs to store the class tables and the list of sites that belong to each class. Also, some computation time is spent in updating the class tables after every iteration. A practical consideration is that, at high lattice temperatures, the advantage of the \( n \)-fold way disappears because all sites approach similar activity levels. However, these restrictions are now almost not relevant since the memory size of desktop computers is commonly 2 GB and parallel computing is making inroads into computational materials science. For most grain growth and recrystallization problems, the overall performance of the \( n \)-fold way algorithm is much better compared to the conventional MC algorithm because the density of boundary sites is low and typically decreases during a simulation.

This approach was adapted by Sahni et al. (1983) to the \( Q \)-state Potts model to take account of variable activity in the system where \( Q \) is the number of possible spin numbers that an individual grid point can take. The system activity, \( A \), is defined as the sum of the separate probabilities for each site for each possible re-orientation over all distinct spin numbers in the system

\[
A = \sum_{i=1}^{N} \sum_{j=1}^{Q-1} p_j (S_i \rightarrow S'_i) \quad (4.11)
\]
Here the subscript $j$ indicates the counter for each possible new spin value, of which there are $(Q - 1)$ possibilities at each site (since changing to the same spin is excluded). Based on this definition, the time increment associated with each successful re-orientation is as follows:

$$\Delta t = -\frac{(Q - 1)}{A} \ln r.$$  \hspace{1cm} (4.12)

This particular approach was, however, limited to zero temperature and uniform grain boundary energy (i.e. a single value of $J$ in Equation (4.3)). An important generalization of the $n$-fold way was made by Hassold and Holm (1993) who pointed out that, although finite temperature introduces a finite probability that a given site will switch to any of the $(Q - 1)$ other possible orientations, most of the switching probabilities are known analytically. Thus the additional work of computing the activities at each site due to finite temperature is minimized and a practicable algorithm was generated. For 2D grain growth, for example, Holm found that the conventional Monte Carlo method was more efficient only when the mean grain radius was less than 3 and computation efficiency increased monotonically as the simulation progressed (Holm, Glazier et al. 1991). Even at high fractions of the critical temperature, the $n$-fold way was more efficient for long simulation times.

It has been noted that the time scale depends on the number of distinct orientations or spins in the system (Equation (4.12)). In reality, however, the rate of grain growth does not depend on the number of different orientations present initially. Therefore it has been suggested that the time increment should be adjusted so as to remove the dependence on $Q$:

$$\Delta t = -\frac{\ln r}{A}.$$  \hspace{1cm} (4.13)

### 4.3 Description of the Monte Carlo Method for Grain Growth & Recrystallization

#### 4.3.1 Discretization of Microstructure

The Monte Carlo also uses a discretized representation of microstructure; however, site interactions are energetically controlled. A continuum microstructure is mapped onto a two-dimensional (2D) or three-dimensional (3D) lattice. Each lattice site is assigned a number, $S_i$, which corresponds to the orientation of the grain in which it is embedded. Lattice sites that are adjacent to sites having different grain orientations are regarded as being separated by a grain boundary, whereas a site surrounded by sites with the same orientation is in the grain interior.

Each site contributes bulk energy, $H(S_i)$, to the system; in recrystallization modeling, $H(S_i)$ is the energy stored at site $i$ during deformation, analogous to dislocation densities in real crystals. In addition, each unlike pair of nearest neighbors contributes a unit of grain boundary free energy $J$ to the system as already described. Summing bulk and surface energy contributions, the total energy of the system is calculated via the Hamiltonian specified in Equation (4.7). In static recrystallization simulations, the stored energy per site is assumed to be positive for initially unrecrystallized material and zero for recrystallized material. When
4.3 Description of the Monte Carlo Method for Grain Growth & Recrystallization

Figure 4.3: Diagram of the square 2D lattice, showing orientation numbers at each site and grain boundaries drawn between sites with unlike orientations. The circled site with spin value 9 has three like neighbors when 1st and 2nd nearest neighbors are counted, and five unlike neighbors. For isotropic grain boundary energy, flipping this site from spin = 9 to spin = 4 would leave the number of unlike nearest neighbors unchanged; thus $\Delta E = 0$ and the flip probability is one.

4.3.2 Evolution of the Microstructure

The evolution of the structure is modeled by picking a site and a new orientation at random from the set of allowable values. The change in total system energy $\Delta E$ for reorienting the site is computed, and the reorientation is implemented with the transition probability, $p$ as discussed in Section 4.2.2. It is important to note the difference between the meaning of temperature in the context of the Monte Carlo model and the physical parameter relevant to recrystallization. In the simulation, temperature governs the degree of disorder in the lattice, and below some critical temperature $T_c$, dependent on the lattice type, the system orders spontaneously. Only second order effects are observed for variations in simulation temperature (Hassold, Holm et al. 1990) on the kinetics of grain growth. Because of this lack of sensitivity to lattice temperature, much simulation work with model has been performed at zero lattice temperature. The consequence of this is to simplify the transition probabilities as follows.

$$p(\Delta E) = \begin{cases} 1 & \text{if } \Delta E \leq 0 \\ 0 & \text{if } \Delta E > 0 \end{cases}$$

(4.14)

The usual procedure in Monte Carlo simulations is to start with a completely random structure that corresponds to a structure obtained at infinite temperature. The structure may then be evolved based on the manner in which it is cooled. One approach is to quench the structure to a temperature below $T_c$ and observe the evolution as the system approaches the equilibrium ordered state. This is the so-called Quenched Potts model. When the quenching temperature is low, i.e. $T$ approaches 0, then the evolution rate is low and it would take long time for
the system to approach equilibrium. Therefore, an alternative approach termed *simulated annealing* is used to evolve the structure. In this case, the temperature of the system is lowered slowly and the structure is allowed to attain equilibrium at each temperature step. The time steps needed to ensure slow cooling are proportional to $\exp(1/T)$ such that, as the temperature decreases, the number of time steps needed increases exponentially. In such cases, a fast *simulated annealing* procedure may be used, see Plischke and Bergersen (1994) for more details.

One Monte Carlo time step (1 MCS) is typically defined as $N$ reorientation attempts, i.e. each site is given an opportunity to change orientation. The number of Monte Carlo Steps is assumed to be proportional to physical time. Recently, it has been pointed out that this definition results in a dependence on the $Q$ value, i.e. the number of orientations in the system. The higher the $Q$ value, the slower the rate at which boundaries will move. This is clearer in the definition of the time step for $n$-fold way in Section 4.2.3 (Equation (4.12)).

### 4.3.3 Inert Particles

Particles are introduced into the simulation as sites which have an orientation different from any of the grains and which cannot be reoriented during the course of the simulation. An individual particle may consist of a cluster of contiguous sites. In 2D, single site particles are effective but in 3D it is important to use a particle size that is comparable (or bigger than) the interaction distance implicit in Equation (4.3). The particles do not react, dissolve or grow themselves and hence are called *inert* rather than *second phase particles or precipitates*. This assumption results in an equality of the particle-matrix interfacial energy and the grain boundary energy, which is reasonable for particles that are incoherent with respect to the matrix. Just as the grain boundary energy can be made a function of the boundary type, so the particle-matrix energy can be varied. Also, the particles cannot move through the lattice, which means that grain boundary drag of particles is not permitted (Ashby and Centamore 1968). Results for particle pinning effects are discussed in Sections 4.9.3, 4.9.4.

### 4.3.4 Lattices

Implicit so far in the discussion has been the connectivity of the points that represent the discretized microstructure. It turns out that the lattice can have a strong effect on the results of the simulation. A survey of lattice types for both two and three dimensions is available in the thesis by Holm (1992). The grain boundary energy per unit length is anisotropic with respect to the boundary orientation in the lattice. This anisotropy can be characterized by a Wulff shape, which is directly related to the coordination number and symmetry of the lattice. Tables 2 and 3 list the lattice types with their geometries and the anisotropy of the Wulff plot. The number in parentheses after the lattice type denotes the number of shells of neighbors, such that square (1,2) means a square lattice with first and second nearest neighbors. The lattice type cubic (2*) denotes a simple cubic lattice with first, second, third nearest neighbors and neighboring points located at [222].

The characteristic of many lattices that has been ignored by several authors is the tendency towards self-pinning for grain growth (explained later) in simulations performed at zero temperature. The last column in each table shows which lattices can sustain grain growth without
4.3 Description of the Monte Carlo Method for Grain Growth & Recrystallization

Table 4.2: Listing of 2D lattice types with geometries and anisotropies.

<table>
<thead>
<tr>
<th>Lattice Type</th>
<th>Wulff Shape</th>
<th>Coordination Number</th>
<th>Anisotropy</th>
<th>Grain Growth</th>
</tr>
</thead>
<tbody>
<tr>
<td>Square (4.1)</td>
<td>Square</td>
<td>4</td>
<td>1.414</td>
<td>Inhibited</td>
</tr>
<tr>
<td>Triangular</td>
<td>Hexagon</td>
<td>6</td>
<td>1.154</td>
<td>Normal</td>
</tr>
<tr>
<td>Square (1,2)</td>
<td>Octagon</td>
<td>8</td>
<td>1.116</td>
<td>Normal</td>
</tr>
<tr>
<td>Triangular (1,2)</td>
<td>Dodecagon</td>
<td>18</td>
<td>1.057</td>
<td>Normal</td>
</tr>
</tbody>
</table>

Figure 4.4: Diagram of the nearest neighbor relationships around a central point labeled 'A' in:
(a) triangular lattice with first nearest neighbors; (b) triangular lattice with 1st and 2nd nearest neighbors; (c) square lattice with 1st and 2nd neighbors.

Table 4.3: Listing of 3D lattice types with geometries.

<table>
<thead>
<tr>
<th>Lattice Type</th>
<th>Wulff Shape</th>
<th>Coordination Number</th>
<th>Grain Growth</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cubic (4.1)</td>
<td>Cube</td>
<td>6</td>
<td>Inhibited</td>
</tr>
<tr>
<td>Cubic (1,2)</td>
<td>18-hedron</td>
<td>18</td>
<td>Inhibited</td>
</tr>
<tr>
<td>Cubic (1,2,3)</td>
<td>26-hedron</td>
<td>26</td>
<td>Normal</td>
</tr>
<tr>
<td>Cubic (2*)</td>
<td>98-hedron</td>
<td>124</td>
<td>Normal</td>
</tr>
<tr>
<td>fcc (4.1)</td>
<td>Rhomboid Dodecahedron</td>
<td>12</td>
<td>Inhibited</td>
</tr>
<tr>
<td>fcc (1,2)</td>
<td>18-hedron</td>
<td>18</td>
<td>Inhibited</td>
</tr>
<tr>
<td>hcp (4.1)</td>
<td>Trapezoidal Dodecahedron</td>
<td>12</td>
<td>Inhibited</td>
</tr>
</tbody>
</table>

self-pinning and are therefore suitable for studies of microstructural evolution. For example, in three dimensions, both of the close packed lattices, face centered cubic and hexagonal, cannot sustain coarsening to long times. The reason for this is a combination of high lattice
anisotropy tending to favor grain facets that lie on high symmetry planes, and a tendency for kinks or steps in boundaries to anneal out with time. If the microstructural features that allow boundaries to move are lost, then it is not surprising that self-pinning occurs. Finite temperature can be used, however, to maintain a population of kinks and steps thereby allowing grain growth to proceed. In general it is advisable to use a finite temperature in order to avoid faceting and loss of the kinks that allow interface motion to occur, regardless of the particular lattice used.

If the grain boundary energy \( J \) is uniform for all boundary segments in a recrystallization simulation, the triangular lattice is suitable for the simulation of recrystallization in two dimensions (Srolovitz, Grest et al. 1986), and that the cubic (1,2,3) lattice is suitable in three dimensions (Anderson, Grest et al. 1989). For other types of simulations, careful examination for lattice effects must be made even for these lattices. Figure 4.3 shows a 2D triangular lattice with boundaries drawn between regions of uniform orientation. Despite the availability of 3D simulation methods, no substantial results have been reported for 3D Monte Carlo simulation of recrystallization.

4.3.5 Boundary Conditions

It is common to use periodic boundary conditions on the spatial domain. This has been likened to modeling a circle or Möbius strip of points in 1D, or a set of points on a torus in 2D. The advantage of using periodic boundary conditions is that it avoids the singularity of edges in a finite domain. The effect of such boundary conditions is illustrated in Figure 4.5. A point on the edge of the domain is connected to points on the opposite edge. The equivalent effect can be obtained by copying the domain along all the edges. Care is needed if a triangular lattice with an oblique two index-addressing scheme is used in order to avoid an implicit shear from connecting the top and bottom edges of the lattice.

Periodic boundary conditions are simple to implement and a natural choice when the initial state of the lattice is artificial. If, however, a microstructure has been measured experimentally, e.g. by orientation imaging microscopy, and is used to initialize the lattice then periodic boundary conditions are obviously inappropriate (Baudin, Paillard et al. 1999; Cheong, Hilinski et al. 2003). The typical choice is to make each edge a mirror, which has the effect of allowing boundaries to terminate at an edge but be able to slide along the edge as dictated by the driving forces that act on them.

4.3.6 Parallelization of the Monte Carlo Algorithm

Since parallelization of the algorithm is largely concerned with boundary conditions, it is appropriate to review the efforts that have been made in connection with the Monte Carlo model. Two main concepts are discussed here: one is based on dividing the simulation domain into subdomains, and the second is based on the checkerboard approach (Holm, 1992). Lubachevsky (1987, 1988) proposed parallelization of the Ising model based on an asynchronous scheme. Korniss et al. (1999) developed a parallel MC model for studying magnetic domains using Lubachevsky approach, described in more detail below. In complete contrast to the subdomain approach, Miodownik (2000) used the checkerboard approach in which a subset of the pixels is operated on, chosen such that in each time-step, each active pixel (or
4.3 Description of the Monte Carlo Method for Grain Growth & Recrystallization

Figure 4.5: Periodic boundary conditions illustrated for a 2D lattice with 1st and 2nd nearest neighbors. The site in solid shading has 5 of its neighbors with similar index values and 3 of its neighbors on the left hand side of the lattice. The arrow indicates the connection from one side of the lattice to the other. The concept is easily extended into 3D.

The main ideas involved in the asynchronous parallelization of the \( n \)-fold way algorithm for Ising-like systems are as follows. The first step in parallelization involves dividing the simulation domain into suitably sized pieces for each processor to work with. One such example of domain decomposition is shown in Figure 4.6 below.

Each processor works on its piece of simulation domain with its local simulation time. When a processor chooses a site that is situated on a processor boundary or a corner, special care must be taken to ensure that the simulation trajectory is not corrupted. This is achieved by allowing a border or corner site update only if its local simulation time is less than or equal to the local times of all the corresponding processors that carry its neighboring sites. For example, in Figure 4.6, if the processor 0 has chosen the diamond-patterned site to update, then this update will be evaluated only when the local simulation time of processor 0 becomes less than or equal to the local simulation time of processors 1, 3 and 4 that carry the neighboring sites marked with circles. Until this happen, processor 0 must wait. Lubachevsky and Korniss et al. defined a special class for boundary and corner sites to handle PE boundaries. The
The Monte Carlo Method

Figure 4.6: Domain decomposition in parallel processing. In the example shown in the figure, a global simulation domain size of $12 \times 12$ sites is divided up into 9 processors, each carrying a local simulation domain size of $4 \times 4$ sites. Note that the sites situated at a processor corner (e.g. horizontal lines) have some of their neighbors located on three other processors (gray), sites at a processor boundary (vertical lines) have some of their neighbors residing in one other processor while sites within the kernel of the local domain (solid) have all of their neighbors within the same processor. The numbers shown are the processing element (PE) or processor ID numbers.

Activity of all of the PE boundary and corner sites was always set to 1. Assigning a high weight to PE boundary sites leads to PE boundary sites being picked more often and this compensates for the waiting at the PE boundaries. This procedure ensures comparable evolution kinetics of the kernel and boundary sites. The time increment in the parallel $n$-fold way is as follows:

$$\Delta t = \frac{-1}{N_b + \sum_{j=1}^{n_j} n_j p_j} \ln r, \quad (4.15)$$

where $N_b$ is the total number of sites on PE boundaries.

The parallel $n$-fold way algorithm based on the asynchronous approach is then stated as follows:

1. Select a class based on Equation (4.13). Note that in parallel, asynchronous $n$-fold way, PE boundary sites have a class of their own as explained above.

2. (a) If the chosen site is in kernel, flip it with probability 1 and go to Step 3.

(b) If the chosen site is in the boundary class, then wait until the local simulation time of this update becomes less than or equal to the local simulation times of the neighboring PEs. When this condition is satisfied, evaluate the environment, compute the
transition probability and flip with *Metropolis probability*, Equation (4.5). Go to Step 3.

3. Update tabulation of spin classes in kernel.

4. Determine time of next update based on Equation (4.15).

5. Go to 1 until sufficient data is gathered.

### 4.4 Nucleation in Recrystallization

Nucleation presents several challenges at the level of mesoscopic simulation of microstructural evolution. It has long been known (Martin *et al.* 1997) that homogeneous nucleation is impossible and that for all practical purposes new grains arise from the existing deformation microstructure. The subgrain structure and other relevant features of the deformation microstructure have, however, a characteristic length scale on the order of 1 $\mu$m or less (Miodownik *et al.*, 1999). This is more than an order of magnitude smaller than the typical grain size after recrystallization (10 to 100 $\mu$m) and so it is impracticable to model the nucleation process in detail. New grains must be introduced *deus ex machina* according to empirical rules for their location and orientation, as described below. There are two exceptions to these remarks. One is work by Radhakrishnan *et al.* who have studied combined models of plastic deformation and recrystallization (Radhakrishnan *et al.* 1998; Sarma *et al.* 2002). The second is work by Holm *et al.* (2003) on modeling the process subgrain coarsening in detail which has led to a theory of nucleation of new grains, discussed in more detail below in Section 4.9.1 on abnormal grain growth.

From an algorithmic point of view, the fact that recrystallization introduces new grains means that it is not practicable to work with a fixed list of orientation values as can be done in grain growth, wherein the maximum numbers of grains occurs at time zero. Therefore it is convenient to introduce each new grain with a unique orientation value. Alternatively, the range of orientation values can be partitioned into recrystallized and non-recrystallized sets. Nucleation is accomplished by introducing new (recrystallized) grains into the simulation with orientation numbers chosen from the appropriate set. The computational efficiency of the continuous time method is unaffected by this characteristic of recrystallization simulations.

Nucleation of recrystallized grains is modeled by adding small embryos to the material at random positions at the beginning of the simulation (i.e. site saturated nucleation) (Srolovitz, Grest *et al.* 1986). The stored energy is set to zero at each site belonging to the embryo. Adding embryos at regular intervals during the simulation simulates continuous nucleation. In both cases the effective nucleation rate decreases with time because, at long times, most of the available space has been recrystallized and has zero stored energy; embryos placed in recrystallized material will shrink and vanish. Dynamic recrystallization is modeled by adding stored energy to each lattice point continuously (Rollett, Luton *et al.* 1992). This means that it is possible to distinguish between recrystallized and unre crystallized material after the first cycle of recrystallization is complete. The process of work hardening starts anew with each new grain, which means that the stored energy at any given point is related to the length of time that has elapsed since the nucleation event associated with that grain.
Embryos have orientations that differ from those of all other grains and particles as discussed above. If the bulk stored energy $H$ is too small, the embryos are sub-critical and shrink away. The value of $H$ required for embryo growth depends on its surroundings and on the lattice being used. Above some critical $H/J$, an isolated embryo is super-critical and can grow as a new grain nucleus. Below that $H/J$, the embryo must be adjacent to an existing grain boundary in order to become a nucleus; its growth then occurs preferentially along the prior grain boundaries. In the 2D triangular lattice, homogeneous nucleation cannot occur for $H/J < 2$. Embryos of two lattice sites can grow when $2 \leq H/J < 4$. Single-site embryos can grow when $H/J \geq 4$.

### Table 4.4: Critical Embryo Sizes in the MC model for Recrystallization (Holm 1996).

<table>
<thead>
<tr>
<th>Stored Energy: Boundary Energy</th>
<th>Critical Size (lattice sites)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>2D, Triangular Lattice:</strong></td>
<td></td>
</tr>
<tr>
<td>$H/J &lt; 2$</td>
<td>(Very large)</td>
</tr>
<tr>
<td>$2 &lt; H/J &lt; 4$</td>
<td>2</td>
</tr>
<tr>
<td>$4 &lt; H/J &lt; 6$</td>
<td>1</td>
</tr>
<tr>
<td>$H/J &gt; 6$</td>
<td>(Any embryo grows)</td>
</tr>
<tr>
<td><strong>2D, Square Lattice,</strong></td>
<td></td>
</tr>
<tr>
<td>with 2\textsuperscript{nd} nearest neighbors:</td>
<td></td>
</tr>
<tr>
<td>$H/J &lt; 1$</td>
<td>(Very large)</td>
</tr>
<tr>
<td>$1 &lt; H/J &lt; 2$</td>
<td>3</td>
</tr>
<tr>
<td>$2 &lt; H/J &lt; 8$</td>
<td>1</td>
</tr>
<tr>
<td>$H/J &gt; 8$</td>
<td>(Any embryo grows)</td>
</tr>
<tr>
<td><strong>3D, Simple Cubic lattice,</strong></td>
<td></td>
</tr>
<tr>
<td>with 2\textsuperscript{nd} and 3\textsuperscript{rd} nearest neighbors</td>
<td></td>
</tr>
<tr>
<td>$H/J &lt; 3$</td>
<td>(Very large)</td>
</tr>
<tr>
<td>$3 &lt; H/J &lt; 5$</td>
<td>5</td>
</tr>
<tr>
<td>$5 &lt; H/J &lt; 8$</td>
<td>3</td>
</tr>
<tr>
<td>$8 &lt; H/J &lt; 26$</td>
<td>1</td>
</tr>
<tr>
<td>$H/J &gt; 26$</td>
<td>(Any embryo grows)</td>
</tr>
</tbody>
</table>

### 4.5 Initialization of MC Simulations

The 2D MC recrystallization simulations reviewed here were typically initialized with a microstructure obtained from grain growth simulations (Anderson et al. 1984; Srolovitz et al. 1984). Grain growth for a period of $10^3$ MCS in a domain size of $200 \times 200$ sites yields a microstructure with approximately 1000 grains and a mean grain area of approximately 40 sites. Whether new grains placed in the structure survive and grow will depend on their spatial location because prior boundaries act as heterogeneous nucleation sites (Srolovitz et al. 1988). If second phase (inert) particles are required, single-site “particles” are randomly placed within
the microstructure to obtain a certain area fraction. The typical method is to assign particle sites an orientation number that is not permitted to change during a simulation. The stored energy is initialized to the desired value for each site. It is also reasonable to start with no prior grain structure, which is equivalent to having a single crystal. The survival and growth of new grains is then independent of spatial location in the lattice.

Another interesting aspect of initialization is how to represent experimentally measured microstructures. Several authors have recently reported using experimental orientation maps to specify the initial state of the model (Baudin et al., 1999; Cheong, et al., 2003). These maps are currently obtained from automated electron back-scatter diffraction (EBSD) characterization of metallographic specimens in a scanning electron microscope (SEM). The transfer is exceptionally simple because the size of the maps is well within the capacity of modern 2D MC codes. It is also noteworthy that Demirel et al. (2003) made a direct comparison of experimentally characterized grain growth with simulations albeit using a finite element code rather than an MC model. They were able to demonstrate that it was essential to include the anisotropy of the grain boundary properties in order to obtain a good agreement between simulation and experiment.

In three dimensions, the challenges of devising accurate representations of microstructure are considerable. Miodownik et al. (1999) have used a simulated annealing procedure (see Section 4.2.3) to match a misorientation distribution to a measured distribution. More recently Saylor et al. have extended this concept to the specification of 3D initial structures that reproduce measured grain shape, orientation and misorientation distributions, again based on EBSD measurements (Saylor et al., 2003). Figure 4.7 shows an example of a 3D microstructure generated by matching EBSD orientation maps on orthogonal sections of a recrystallized aluminum 1050 alloy. This statistical approach opens up the possibility of being able to perform simulations of microstructural evolution that can be compared directly with experimental measurements.

4.6 Verification of the Monte Carlo Model

An important aspect of any model is to verify that it behaves as expected. This is not a trivial issue for this model because most of the basic features governing grain growth are not imposed on the model. For example, there is nothing in the formulation given above to guarantee that the migration rate of boundaries is proportional to the driving force, or that local equilibrium is maintained at triple junctions between the surface tensions.

The most basic verification that the model works as desired is to simulate a single isolated grain shrinking under the action of the curvature of its boundary. Mullins’ (1956) analysis shows that, for velocity proportional to mean curvature, the rate of change of area, $\frac{dA}{dt}$, of any grain is constant and equal to the integral of the turning angle around the perimeter of the grain. In the case of a grain with $n$ vertices, this leads to the expression:

$$\frac{dA}{dt} = -2\pi M \gamma (n - 6),$$

(4.16)

where $M$ is the mobility, and $\gamma$ is the grain boundary energy. For an isolated grain with one side, the rate of area change is constant. The reason for 6-sided grains being neutral is that,
for isotropic grain boundaries, the turning angle at a three-fold vertex (triple junction) must be $60^\circ$.

To test this relationship for the case of $n = 1$ is simple because the multi-state model collapses to the Ising model with one grain isolated within another grain. In either 2D or 3D, the result is the same, i.e. that $dA/dt$ is constant (and negative). Recently, Holm has verified (Holm 2002) that the rate of collapse of the isolated grain is exactly equivalent to motion by curvature for the case of a 2D square lattice with first nearest neighbors. This is also important because it verifies that the model satisfies the relationship

\[ V = M P , \]  

(4.17)

where $V$ is the boundary velocity, $M$ is the mobility, $P$ is the driving pressure, taken here to be the product of boundary energy and curvature, $P = \gamma \kappa$.

A second test of the model is to examine grain growth kinetics in a polycrystal. Classical theory states that the average grain area, $\langle A \rangle$, should vary linearly with time:

\[ \langle A \rangle - \langle A \rangle_{t=0} = kt . \]  

(4.18)
This is justified on the basis of the observed self-similarity of the microstructures, which simply means that snapshots taken at different times can only be distinguished if either a time or a length scale is provided. More specifically, self-similarity means that the average curvature in the system scales with the inverse grain size. This permits a simple differential equation to be solved as follows:

\[
\frac{d\langle R \rangle}{dt} = k \langle \kappa \rangle = k \langle R \rangle^{-1} \\
\frac{d\langle R \rangle}{\langle R \rangle} = k \, dt.
\]

Therefore,

\[
\langle R \rangle^2 = kt + C. \tag{4.19}
\]

Thus the predicted exponent on area is one (or two for grain diameter). Provided that self-pinning does not occur (through poor choice of lattice) the exponent obtained from MC simulations is very close to one. There has been considerable attention paid to this point because the earliest results suggested an exponent slightly less than one. Various authors have ascribed this discrepancy to either a too small simulation domain, incorrect choice of switching algorithm (Radhakrishnan and Zacharia 1995), or to improper regression analysis. Nevertheless, the MC model is sufficiently robust that, provided a reasonable choice of lattice, domain size, and method of regression analysis is made, then the expected theoretical grain growth kinetics will be obtained. A more complex issue in grain growth is that of grain size distributions which space does not permit us to review in detail here. Suffice it to say that the distributions observed in 2D MC simulations correspond closely to experimental observations (Srolovitz et al. 1984) and that, as of the time of writing, the theoretical basis for size distributions is still an active topic of discussion in the literature.

Another way of looking at the connection between polycrystal coarsening and motion by mean curvature is to examine the link between topology and size. There is a strong correlation between topological class and size that is known in both the experimental literature (Feltham 1957; Aboav and Langdon 1969) and for the MC model (Srolovitz et al. 1984). The nearly linear relationship between size and number of sides of a grain provides a direct link between the \((n - 6)\) rule, Equation (4.16), and the mean field equation for growth rate of an individual grain, below. It is this connection that Hillert exploited to derive his seminal theory of grain growth based on the Lifshitz–Slyozov–Wagner theory of coarsening.

\[
\frac{dR}{dt} = -k \left( \frac{1}{R_{\text{critical}}} - \frac{1}{R} \right). \tag{4.20}
\]

In recrystallization, there are two basic verifications of the model. The first is that the driving forces of curvature and stored energy can be set in opposition to one another. In the context of the isolated grain, this means that the Gibbs-Thomson effect can be verified: in other words, for a given grain boundary energy (set by \(J\)) and stored energy (set by \(H\)) there should be a grain size that neither shrinks nor grows. This is indeed the case for the MC model as has been verified (Rollett and Raabe 2001). Figure 4.8 shows the result of simulating the behavior of a single grain with an initial radius of eight sites and a range of stored energies associated with the sites surrounding the single grain. For a large enough stored energy, the grain radius...
grows with constant velocity as expected, Equation (4.17). For small stored energies, the grain shrinks under capillary pressure. At some intermediate stored energy the size is metastable: fluctuations on either side will lead to growth or shrinkage.

The second verification is to perform simulations of recrystallization and compare the kinetics against the Kolmogorov-Johnson-Mehl-Avrami (KJMA) theory. This has been done by several investigators (Srolovitz et al. 1986) and extensions of the model have been made in order to explore experimentally observed deviations from the expected kinetics (Rollett et al. 1989). The slopes observed in KJMA plots correspond to the expected theoretical values for a variety of nucleation and growth morphologies.

### 4.7 Scaling of Simulated Grain Size to Physical Grain Size

Most computer simulations require significant amounts of computer time so it common practice to minimize the size of the lattice that is used. Therefore it is useful to analyze the relationship between grain size in the Monte Carlo recrystallization model, and physical grain...
sizes (Rollett et al. 1992). As of the time of writing, increases in computer power have made it feasible to perform simulations in a few hours on desktop computers with more than a million grid points over simulation times of many millions of MCS.

We will assume that the ratio of stored energy to grain boundary energy per unit volume can be equated in the model and in real, physical recrystallization process. The physical process has two characteristic driving pressures for boundary migration, \( P_{\text{store}} \) and \( P_{\text{gr}} \). The stored dislocation content typically yields a \( P_{\text{store}} \) of 10 MPa. For grain growth driven by grain boundary curvature, \( P_{\text{gr}} = \gamma/\langle R \rangle \), where \( \gamma \) is the grain boundary energy per unit area, and \( \langle R \rangle \) is the mean grain radius. In the 2D triangular lattice, the stored energy per unit area is given by \( P_{\text{store}} = H/[3s^2 \sin(60^\circ)] \) where \( s \) is the unit boundary length on the lattice. The stored energy due to boundary curvature is

\[
P_{\text{gr}} = \frac{\gamma_{\text{model}}}{\langle R \rangle_{\text{model}}} = \frac{J}{s\langle R \rangle_{\text{model}}}. \tag{4.21}
\]

The initial grain size in the Monte Carlo simulations is typically about 6 \( s \). Thus, for the model

\[
P_{\text{store}} \approx P_{\text{gr}} = \frac{H\langle R \rangle_{\text{model}}}{3sJ \sin 60^\circ} = 2.5 \frac{H}{J}. \tag{4.22}
\]

Using a typical value for the grain boundary energy, \( \gamma = 0.5 \text{ Jm}^{-2} \), for physical systems

\[
P_{\text{store}} = 20\langle R \rangle \text{ m}^{-1}. \tag{4.23}
\]

Equating the energy densities for the model and physical systems and rearranging gives

\[
\langle R \rangle = 0.125H/J \mu \text{m}. \tag{4.24}
\]

Then for a typical simulation with \( H/J = 2 \), we can estimate \( \langle R \rangle = 0.25 \mu \text{m} \), which is a small but not unphysical grain size. Clearly, however, it would be preferable to simulate recrystallization with lattices with linear dimensions an order of magnitude larger than is currently typical.

### 4.8 Recrystallization Kinetics in the Monte Carlo model

A characteristic of the Monte Carlo model of recrystallization is that a finite recrystallized volume is introduced at the beginning of simulations when site saturated nucleation conditions apply. This becomes apparent in a KJMA plot as a curvature at early times. A simple correction may be made for the finite initial fraction transformed by adding a constant to the measure of time. The correction is of course heuristic because it depends on the results themselves.

A limitation of the Monte Carlo model is that the growth rate of recrystallized grains is not linearly related to the stored energy density. As with nucleation, each integer increment of \( H/J \) leads to a discrete change in the number of sites that can change orientation with a neutral or negative change in system energy.
4.9 Results of Simulation of Recrystallization by Monte Carlo Method

The section provides a review of a number of areas in which useful results have been obtained from recrystallization simulation using the Monte Carlo method. Note that all the results discussed here are for two-dimensional simulations only.

4.9.1 Abnormal Grain Growth

The early stages of recrystallization are labeled as nucleation even though no new phase appears in the material. At the level of the dislocation structure, however, the existing heterogeneities of the deformed structure coarsen in the process known as polygonization. The heterogeneities exist at several length scales from cells to shear bands to prior grain boundaries. Many observations have been made which suggest that individual subgrains acquire a growth rate advantage over their neighbors and become identifiable as new grains. This growth advantage can result from a difference in mobility between the boundary of the new grain and the boundaries in the surrounding material. This process of competitive growth has been observed in the Monte Carlo model in both 2D (Rollett et al. 1989) and 3D (Grest et al. 1990) simulations. By altering the rate at which sites are sampled for reorientation, the mobility of specific grain boundaries can be varied. Small ratios in mobility between one grain and another lead to marked abnormal grain growth behavior (Rollett and Mullins 1996), which may correspond to the early growth of new grains in recrystallization. Holm et al. (2003) have recently derived a theory for the frequency of abnormal grains as a function of orientation spread and the characteristic angle at which the grain boundary mobility transitions from low to high values. The key feature of the simulations carried out in this work was that coarsening in microstructures based on a single texture component with a realistic spread in orientation of the subgrain structure will occasionally exhibit abnormal grain growth. It turns out that if a particular grain has the topology required for growth (more than six sides, for example) and happens to be at the edge of the orientation distribution such that its perimeter possesses a high misorientation, it will grow much faster than the average size of the matrix. Figure 4.9 shows a snapshot of the microstructure from such a simulation after MCS. High angle grain boundaries are drawn in solid black whereas low angle boundaries are white. Grains that have grown to sizes significantly larger than the average size clearly tend to be surrounded by high angle, mobile boundaries.

4.9.2 Static Recrystallization

A key observation in grain growth is that an isolated grain will shrink and eventually vanish in response to tendency to minimize grain boundary area. When an isolated grain can eliminate stored energy by growing, however, it does so, provided that the grain is not smaller than a critical size. In a polycrystalline structure with $0 < H/J < 1$, any recrystallized grain will grow at the expense of its neighbors because of the bias imposed on the unit step of the grain growth process. That is to say, the motion of kinks along a boundary is reversible when the change in energy associated with a step is zero, as is often the case for the 2D
4.9 Results of Simulation of Recrystallization by Monte Carlo Method

Figure 4.9: Microstructure showing abnormal grain growth in a subgrain structure whose spread about the cube orientation was $8^\circ$. High angle grain boundaries ($>10^\circ$) are drawn as black lines and low angle boundaries ($>1^\circ$) are white. Large grains are associated with highly misoriented boundaries.

The kinetics of recrystallization have been found to reproduce those anticipated from theoretical analysis very closely. For example, in 2D simulations (Srolovitz et al. 1986) site saturated nucleation conditions with a high enough stored energy density ($H/J > 2$), a KJMA plot of the fraction recrystallized versus time show a slope of 2 at long times; continuous nucleation gives a slope of 3. Both results are as predicted from classical KJMA analysis. For low stored energy densities ($H/J < 2$), nucleation is heterogeneous in the sense that embryos must be adjacent to existing boundaries in order to survive and grow. In this case the kinetics show significant deviations from the classical KJMA pattern (Srolovitz, Grest et al. 1988). The key observation here is that the growth of new grains at low stored energies is highly dependent on the prior structure. For $H/J \leq 1$, the recrystallization front grows out from triangular lattice; however, such energy neutral steps become irreversible when biased by the elimination of stored energy at each step. Clearly, the Monte Carlo simulation of the motion of recrystallization fronts is much closer to a deterministic model (e.g. cellular automaton) than for grain growth.
a triple point and is concave with respect to the unrecrystallized side. The net effect is that growth is very slow in early stages of recrystallization.

Work by Martin (1994) with the Monte Carlo method has examined the effect of spatially non-random distributions of nuclei under site-saturated conditions. This work used a variant of the $n$-fold way in which only flips to nearest neighbor orientations were considered. Also, a stored energy (as a scalar contribution to the system energy associated with each unrecrystallized site) was not explicitly considered. Instead, once a recrystallized grain was inserted, flips are allowed from unrecrystallized to recrystallized, but not vice versa, thereby guaranteeing growth of the recrystallized regions. The main result was to demonstrate the importance of the distribution of nuclei. Nucleation on a square lattice gave narrow size distributions whereas (heterogeneous) nucleation on a coarse prior grain structure gave wide distributions and non-compact grain shapes. Subsequent grain growth broadens the size distribution and, as expected, grain shapes become compact.

4.9.3 Grain Growth in the Presence of Particles

Miodownik et al. (2000) applied a parallel version of the 3D model with the primary aim of being able to perform simulations on particle pinning with large domain sizes. For the problem of particle pinning that they studied, this was of crucial importance because they needed to use particles larger than one voxel (a unit volume element) in order to avoid thermally activated unpinning. Their result was instructive: in contrast to earlier 2D studies (Srolovitz et al. 1984; Doherty et al. 1990) that suggested that particles pinned grain boundaries more effectively than predicted by the Zener-Smith theory (Zener 1948), the new 3D results showed that the classical theory was indeed applicable. The difference between 2D and 3D appears to be that in 2D, particles can remove curvature because a boundary can effectively pivot about a particle. Thus the limiting grain size is similar to the nearest neighbor spacing, $\Delta_2$. In 3D, however, the effect of a particle on a boundary is more local and so the limiting grain size is similar to the mean free path between particles.

4.9.4 Recrystallization in the Presence of Particles

Recrystallization with inert particles present is easily modeled by assigning sites orientation values that cannot be changed. Although only the effects of single site particles have been studied (Rollett et al. 1992) in 2D simulations (whereas particle shape and size has been examined for grain growth (Hassold et al. 1990)), the results appear to be general, at least for small particles. Large particles in physical systems have the effect of stimulating nucleation, an effect that has not been addressed by microstructural simulations. Another worthwhile extension of computer modeling of recrystallization would be three-dimensional simulations because the interaction of boundaries with particles is quite different than in two dimensions.

During recrystallization simulation with sufficient stored energy ($H/J \geq 3$ in the triangular lattice) the recrystallization front can readily bypass particles regardless of particle size or area fraction. Under these circumstances the recrystallization growth kinetics are unaffected by particles. The overall kinetics is accelerated slightly, however, by heterogeneous nucleation on particles.
4.9 Results of Simulation of Recrystallization by Monte Carlo Method

Figure 4.10: Microstructures of a single recrystallizing grain growing into a matrix containing inert second phase particles. The last picture illustrates the effect of periodic boundary conditions as the new grain wraps around the edges of the simulation.

At intermediate stored energies \(1 \leq H/J \leq 3\) nearly all recrystallization boundaries can move past single-site particles. Boundaries with a very high particle density will stop moving. However, the irreversible propagation of grain boundary kinks allows most boundaries to achieve a configuration from which two kinks can join to move past single-site particles. The presence of prior grain boundaries further enhances recrystallized boundary motion. In these systems, boundaries intersect particles at random, and recrystallization kinetics are substantially unaffected by particles, as shown in Figure 4.10, where a single recrystallized grain shows unrestricted growth for \(H/J = 1\). Larger particles may inhibit recrystallization in this stored energy regime, however, since two boundary kinks cannot join directly.

At low stored energies \((H/J < 1)\) grain boundary energy governs boundary motion, and the recrystallization front is strongly pinned by particles which leads to a much higher than random density of particles on the recrystallization front. In these circumstances recrystallization is strongly inhibited which usually results in incomplete recrystallization, as shown in Figure 4.13. When only small particle fractions are present, however, recrystallization may go to completion because pinning does not occur until after the transformation is complete.

In addition, prior grain boundaries may still enhance motion of recrystallized grain boundaries, so that the recrystallized grains can grow (at low particle fractions) much larger than the deformed matrix grains. This is sufficient to drive the recrystallizing grains past some particles, but only if the matrix grain size is much smaller than the interparticle spacing. In other words, recrystallization at low stored energy and at very low particle fractions is similar to abnormal grain growth.
Grain boundaries undergoing curvature driven growth (both in the deformed state and after complete recrystallization) rapidly acquire a higher than random density of particles which then inhibits grain growth. If the recrystallized grain size is smaller than the critical grain size, grain growth continues until pinning occurs and the microstructure is a normal grain growth microstructure. However, when the recrystallized grain size is large compared to the critical grain size, particle pinning occurs almost immediately following the completion of recrystallization thus preserving the non-compact grain shapes and sharply peaked grain size distribution that are characteristic of randomly distributed nuclei.

Analysis of recrystallization in particle containing materials suggests that there are two limiting values of particle drag; a low (Zener) value with a random density of particles and a much higher value if particles have become highly correlated with the recrystallization front (as in grain growth). The simulation results show both these behaviors, depending on the $H/J$ ratio. Figure 4.11 plots the density of particles on boundaries for two types of boundaries; recrystallization fronts have a lower, near random spatial distribution of particles; general boundaries exhibit a higher than random density, suggesting that they are more strongly pinned. Experimental studies appear to show only the lower particle drag (Zener) as studied by Ashby et al. (1969). A note of caution about the comparison is that boundary-particle interaction is more complex in three dimensions than in two.

### 4.9.5 Texture Development

Tavernier and Szpunar extended the 2D Monte Carlo model of Anderson et al. (1984) to account for the different boundary energies and mobilities expected for different texture components. They chose low carbon steel sheet as their model material and developed parameters to account for eight different texture components. They found that the $<111>$ class of texture component tended to become dominant after a period of grain growth when the simulations were run with high boundary energies between the $<111>$ component grains and other texture components. They also simulated the recrystallization process by making provision for the variation of stored energy and boundary mobility, depending on whether a given volume of material is recrystallized or not. Unfortunately many details of the key parameters were omitted from their paper so that it is difficult to judge the success of their efforts. They also point out, reasonably enough, that there is very little experimental data available for the variation of grain boundary energy and mobility over the general range of misorientations. This lack is currently being addressed by both simulation of grain boundary properties (Upmanyu et al. 1998) and by experimental measurement (Yang et al. 2001; Saylor et al. 2002). More recently, Hinz & Szpunar have used the MC model to investigate the role of special boundary types (coincident site lattice boundaries) on texture development in electrical steels, especially with respect to the Goss texture, \{110\}\langle001\rangle.

Holm et al. (2001) have investigated the development of grain boundary character as quantified by the misorientation distribution during grain growth. They find that low energy boundaries become strongly preferred during coarsening in 2D structures where the range of orientation is limited to only one degree of freedom. In 3D coarsening with 3-parameter orientations, however, even if the special boundaries based on CSL relationships are assigned low energies, the grain boundary character does not reflect the energy distribution because of the geometric constraints.
Recently Rollett (2004) has investigated the development of the cube texture component in fcc metals. The investigation was confined to grain growth with no stored energy driving forces. The initial microstructure was based on a microstructure from simulations of isotropic grain growth. A rolling texture was imposed on the structure except for a small fraction of the grains that were assigned orientations close to the cube component. A combination of a Read-Shockley model of grain boundary energy for low angle boundaries and a broad maximum in mobility around a misorientation of $40^\circ <111>$ with very low mobility at low misorientation angles was used to describe the grain boundary properties. For a sufficiently large ratio of maximum to average mobility, the cube component was observed to increase markedly during grain growth.
Incorporating texture into the Monte Carlo model is straightforward although, as always, symmetry (both crystal and sample-based) must be dealt with carefully. The most efficient way to proceed is to calculate all the grain boundary properties that could be required during a simulation at the beginning and store them in a look-up table. Each spin number of the $Q$ possible spins represents an individual orientation. The required lookup table is then $Q(Q-1)/2$ in size for each property in order to account for all possible boundary types in the system.

### 4.9.6 Texture

Although texture is a substantial topic in itself, it is perhaps helpful to the reader to provide a basic introduction while providing references to the standard works. Texture has to do with how one quantifies the relationship between the crystal axes in a particular grain of a polycrystal and a set of axes associated with the external shape of the polycrystalline body. Most often the term texture is used in conjunction with x-ray pole figures which are the most efficient experimental means of quantifying the average texture of a (polycrystalline) material. Most fundamental is the understanding that crystallographic orientation requires specification of a rotation. This rotation is most often used as an axis transformation in order to express properties known in crystal axes into properties in the frame of the material. In the context of Monte Carlo simulations, orientations are used to determine the properties of grain boundaries. The second vital piece of information is that rotations can be expressed in a wide variety of mathematical and not-so-mathematical parameterizations, all of which have three independent parameters. The standard symbol for orientation is "$g$" but other symbols are used, especially for the less well-known Rodrigues-Frank vectors and quaternions. The table below provides a summary of commonly used approaches. The entries are ordered by familiarity to materials scientists. Specification of a plane and direction by Miller indices is most intuitive but numerically awkward! Euler angles are common because of the convenience series expansion methods based on generalized spherical harmonics. Serious computational work, however, uses quaternions for speed and simplicity. Axis-angle descriptions are the most intuitive description of grain boundaries because of the intimate connection to crystal geometry and Rodrigues-Frank vectors have some very attractive features for both representation and for certain types of computation. A few conversion formulae are given in the table in order to provide some clarification of the meaning of the parameters. There are formulae available in the standard texts to convert between any pair of representations (Bunge 1982; Kocks et al. 1998).

- **Texture Component:**
  Specifies alignment of a plane normal specified by Miller indices, (hkl), parallel to sample direction 3 (ND) and a crystal direction, [uvw], with sample direction 1 (RD). Each normal and direction is, in effect, a unit vector and they must be perpendicular, yielding three independent parameters.
  Parameters: (hkl)[uvw]

- **Euler Angles:**
  The Euler angles specify a triple of rotations (transformations) about the Z, X and Z directions. Many variants of Euler angles are known and we present here the most
4.9 Results of Simulation of Recrystallization by Monte Carlo Method

Parameters: $g = g(\phi_1, \Phi, \phi_2)$

Conversion Formula:

$$a_{ij} = \begin{pmatrix}
\cos \varphi_1 \cos \varphi_2 & \sin \varphi_1 \cos \varphi_2 & \sin \varphi_2 \sin \Phi \\
-\sin \varphi_1 \sin \varphi_2 \cos \Phi & + \cos \varphi_1 \sin \varphi_2 \cos \Phi & \\
-\cos \varphi_1 \sin \varphi_2 & -\sin \varphi_1 \sin \varphi_2 & \cos \varphi_2 \sin \Phi \\
\sin \varphi_1 \sin \Phi & -\cos \varphi_1 \sin \Phi & \cos \Phi \\
\end{pmatrix}$$

- **(Orthogonal) Matrix:**
  The coefficients of an axis transformation are defined by: $a_{ij} = \hat{e}_i \cdot \hat{e}_j$, where the $e$ are the unit basis vectors in the primed and unprimed coordinate systems. All columns and rows are unit vectors which reduces the number of independent coefficients to 3. The unit vector, $b$, is derived from the Miller indices, $[uvw]$, of the crystal direction parallel to sample direction 1. Similarly the unit vector $n$ is derived from $(hkl)$ and the unit vector $t$ is the cross-product $n \times b$.

Parameters: $g = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$

Conversion Formula: $a = \begin{pmatrix} b_1 & t_1 & n_1 \\ b_2 & t_2 & n_2 \\ b_3 & t_3 & n_3 \end{pmatrix} \equiv Crystal \begin{pmatrix} s_{Sample} \\ b_1 & t_1 & n_1 \\ b_2 & t_2 & n_2 \\ b_3 & t_3 & n_3 \end{pmatrix}$

- **Axis-Angle:**
  The rotation axis is specified by a unit vector, $n$, and the rotation angle by $\theta$. For grain boundaries, the rotation axis is often specified in crystallographic terms with a set of Miller indices.

Parameters: $g = g(\theta, n)$

Conversion Formula: $a_{ij} = \delta_{ij} \cos \theta + n_i n_j (1 - \cos \theta) + \sum_{k=1,3} \varepsilon_{ijk} n_k \sin \theta$

- **Rodrigues-Frank Vector:**
  The RF-vector is the rotation axis but scaled by the tangent of the semi-angle. In this space, all rotations that share a common rotation axis lie on a straight line.

Parameters: $\rho = \{\rho_1, \rho_2, \rho_3\} = \tan(\theta/2)n$
Conversion Formula:

\[
\begin{align*}
\rho_1 & := \frac{\tan \left( \frac{\Phi}{2} \right) \cdot \sin \left( \frac{\phi_1 - \phi_2}{2} \right)}{\cos \left( \frac{\phi_1 + \phi_2}{2} \right)} \\
\rho_2 & := \frac{\tan \left( \frac{\Phi}{2} \right) \cdot \cos \left( \frac{\phi_1 - \phi_2}{2} \right)}{\cos \left( \frac{\phi_1 + \phi_2}{2} \right)} \\
\rho_3 & := \tan \left( \frac{\phi_1 + \phi_2}{2} \right)
\end{align*}
\]

- Quaternion:
  The quaternion is closely related to the Rodrigues-Frank vector. The \( n_i \) are the components. For rotations, the quaternion is always of unit length. Be aware that the quaternions representing \( \theta \) and \( 2\pi - \theta \) are the negative of each other but represent the same orientation (2-to-1 mapping).

Parameters: \( q = \{q_1, q_2, q_3, q_4\} = \{\sin(\theta/2)n_1, \sin(\theta/2)n_2, \sin(\theta/2)n_3, \cos(\theta/2)\} \)

Conversion Formula:

\[
\begin{align*}
q_1 & := \sin \left( \frac{\Phi}{2} \right) \cdot \cos \left( \frac{\phi_1 - \phi_2}{2} \right) \\
q_2 & := \sin \left( \frac{\Phi}{2} \right) \cdot \sin \left( \frac{\phi_1 - \phi_2}{2} \right) \\
q_3 & := \cos \left( \frac{\Phi}{2} \right) \cdot \sin \left( \frac{\phi_1 + \phi_2}{2} \right) \\
q_4 & := \cos \left( \frac{\Phi}{2} \right) \cdot \cos \left( \frac{\phi_1 + \phi_2}{2} \right)
\end{align*}
\]

The next step in understanding and using texture is to become familiar with the characteristic preferred orientations of the particular material and processing history of interest. This is far too broad a subject for treatment here and the reader is referred to Kocks, Tomé and Wenk (1998) and Randle and Engler (2002) for detailed information and analysis.

From a mathematical point of view, the next step is to realize that rotations can be combined together. The simplest method is matrix multiplication: any pair of orthogonal matrices can be (matrix) multiplied together to yield another rotation. Note that group theory is very useful in this context. Less intuitive but just as useful are the methods of combining (or ‘composing’) two Rodrigues vectors or quaternions. Two Rodrigues vectors combine to form a third as follows where \( \rho_2 \) follows after \( \rho_1 \):

\[
(\rho_1, \rho_2) = \left\{\rho_1 + \rho_2 - \rho_1 \times \rho_2\right\}/\{1 - \rho_1 \rho_2\}. \tag{4.25}
\]

The algebraic form for combining quaternions is given as, where \( q_B \) follows \( q_A \):

\[
q_C = q_A \cdot q_B
\]
θ is the misorientation angle, the smallest possible rotation angle can be identified. The formula for determining the smallest which crystal axes can be labeled. The following expression summarizes the way in which ically equivalent descriptions of any misorientation because of the multiplicity of ways in

\[ q_{C1} = qa_4q_b4 + qa_4q_b2 - qa_2q_b3 + qa_3q_b2 \]
\[ q_{C2} = qa_2q_b4 + qa_4q_b2 - qa_3q_b1 + qa_1q_b3 \]
\[ q_{C3} = qa_3q_b4 + qa_4q_b3 - qa_1q_b2 + qa_2q_b1 \]
\[ q_{C4} = qa_4q_b4 - qa_1q_b1 - qa_2q_b2 - qa_3q_b3 \] (4.26)

Now we can introduce the concept of misorientation, or the difference in orientation be-

\[ \Delta g = g_2 \cdot g_1^T \] (4.27)

Similarly for quaternions, the expression is:

\[ q_c = q_A \cdot q_B^{-1} \]
\[ q_{C1} = qa_1q_b4 - qa_4q_b1 + qa_2q_b3 - qa_3q_b2 \]
\[ q_{C2} = qa_2q_b4 - qa_4q_b2 + qa_3q_b1 - qa_1q_b3 \]
\[ q_{C3} = qa_3q_b4 - qa_4q_b3 + qa_1q_b2 - qa_2q_b1 \]
\[ q_{C4} = qa_4q_b4 + qa_1q_b1 + qa_2q_b2 + qa_3q_b3 \] (4.28)

Finally we must summarize the effect of crystal symmetry because there are many phys-

\[ \theta^* = \min \left\{ \cos^{-1} \left( \frac{(O^{(i)}g_bg_a^{-1}O^{(j)} - 1)}{2} \right), \cos^{-1} \left( \frac{(O^{(k)}g_bg_b^{-1}O^{(l)} - 1)}{2} \right) \right\}, \]
\[ \{i, j, k, l = 1, \ldots, n\} \] (4.29)

Inspection of this expression will reveal that it can also be used to specify the misorientation in a unique way. In fact, one can chose specific symmetry operators in such a way as to always locate the misorientation axis in a particular asymmetric unit such as the standard stereographic triangle (SST) for cubic materials. Finally, we note that the properties of grain
boundaries (and other interfaces) very often depend also on the boundary normal in addition to the misorientation. This adds another two parameters to the three already required to describe misorientation and boundary normals require additional computational effort in the MC model (and which has not yet been implemented as of the time of writing).

4.9.7 Dynamic Recrystallization

Dynamic recrystallization has been successfully modeled with the 2D Monte Carlo model (Rollett et al. 1992). In its simplest form, stored energy at each point of the lattice is increased at a fixed rate and embryonic new grains are continually added at a constant rate. The basic result is that temporal oscillations are observed both in the grain size and in the stored energy, which is analogous to the oscillations in the flow curve, as shown in Figure 4.12. These oscillations are observed over almost the entire range of recrystallization parameters examined (energy storage rate, nucleation rate, initial grain size) and damp out over time periods that decrease with increasing storage rate and increasing nucleation rate. The oscillations in both grain size and stored energy have the same period but are out of phase by approximately one quarter of a period. Examination of the simulated dynamic recrystallization microstructures which were formed under the same conditions but with different initial grain sizes shows that the evolution of the microstructure may be divided into three distinct stages: an initial microstructure dependent transient stage, an initial microstructure independent transient stage, and a steady state stage. While necklace nucleation was observed in these simulations under some circumstances, it is apparent that it is not a necessary condition for grain refinement in dynamic recrystallization. This phenomenon is associated with refinement of the relatively coarse initial microstructure and over-damped oscillations in the flow stress. Therefore, even when there are no obvious oscillations in the flow curve and no necklace nucleation is observed, dynamic recrystallization cannot be precluded.

It should be noted that the geological community had applied the Monte Carlo model to dynamic recrystallization but had only examined the microstructural aspects. Jessel (1988) described an adaptation of the Monte Carlo model for the simulation of deformation of quartzite. Although the model is similar to the combined grain growth and recrystallization model, the Jessel work used only differences in stored energy between sites to evaluate transition probabilities for orientation changes. In this form, the model is essentially a cellular automaton (CA) model. Also, the only results given were for microstructural evolution, with no attempt to investigate stress-strain relationships. The simulations were used to investigate the development of texture (fabric in geological terms), based on the Taylor model (strain compatibility enforced on all grains) with some degree of success.

Peczak and coworkers have investigated several aspects of the correspondence between this type of simulation and have added several refinements to the model, see for example (Peczak 1995), in order to allow detailed comparisons with experimental data. They have modified the nucleation process such that the probability of an embryo appearing in a region of high stored energy is higher than for low stored energies. This has the consequence that the effective nucleation rate increases with time as stored energy is added to the system and is also position dependent during the simulation. They have also modified the rate of stored energy addition from the original constant addition rate to correspond to a Voce-type equation (i.e. an exponential work hardening curve) with a saturation (asymptotic) flow stress expressed in
4.10 Summary

The principal methods for modeling the phenomenon of recrystallization have been reviewed. In addition to the geometrical and analytical models, several methods of modeling microstructural evolution during recrystallization are briefly described. Many of the latter examples illustrate to need to be able to model the local environment of grains. In view of the frequency with which the Monte Carlo model has been used in the literature, that method has been re-

Figure 4.12: Stress plotted versus strain for a range of rates of addition of stored energy in a Monte Carlo model, Rollett (1992). Note the oscillations observed in the simulated flow curves whose period varies with the rate of addition of stored energy.

terms of temperature-compensated strain rate (Zener-Holloman parameter). By making these modifications they have been able to reproduce many of the characteristics of the phenomenon of dynamic recrystallization. For example, the transition from multiple peaks to a single peak in the flow curve occurs when the ratio of the initial to the steady state grain size goes over 2.3. This is in agreement with Sakai’s experimental observations on the effect of initial grain size on dynamic recrystallization in OHFC copper (Sakai 1995). The relationship between the steady state grain size and flow stress in the model shows grain size decreasing with increasing stress as expected; the slope is \( \sim 1 \), Figure 4.13, which is close to the experimental value of 0.7 (Derby 1991).
viewed in more depth. The characteristics of the various lattice types, the Hamiltonians for the system energy, the transition probabilities, and the continuous time method were briefly treated. Key results from the simulation literature have been reviewed and summarized.

Acknowledgements

Numerous discussions over the years with D.J. Srolovitz, Elizabeth Holm, A. Karma, Y. Brechet, B. Radhakrishnan, S. Cheong, K. Okuda and others are gratefully acknowledged. This work was supported primarily by the MRSEC program of the National Science Foundation under Award Number DMR-0079996. Additional support was provided by the Computational Materials Science Network, a program of the Office of Science, US Department of Energy and by the Programming, Environment and Training Program under the auspices of the High Performance Computing Modernization Office of the US Department of Defense.

References

References


References


5 Crystal Plasticity

Paul R. Dawson

5.1 Introduction

Many materials of engineering and scientific interest exhibit a crystalline structure. Commonly, these materials are both polycrystalline and polyphase. The individual crystals, or grains, are domains with continuous lattice orientation and vary widely in size, with $10^{-7}$ to $10^{-3}$ m being a typical range. Recent attention to materials with nanoscale grains have pushed the range to even smaller dimensions in engineering applications. The crystalline structure of materials is a key element in a model of the mechanical response, as individual crystals are anisotropic in both the elastic and plastic behaviors. For the plastic response, slip is intimately linked to crystallographic planes and directions, making crystal strength inherently anisotropic. If one were to consider only volumes of material that contain large numbers of grains whose lattice orientations are uniformly distributed, the strength of the aggregate would exhibit little if any anisotropy. However, as a consequence of prior thermomechanical processes, most materials display crystallographic texture, meaning that the lattice orientation distribution is patterned rather than being uniform. In such cases, the aggregate strength displays anisotropic behavior that is a convolution of the strength at the level of single crystals together with the distribution function for the crystal lattice orientations.

In this chapter we outline the essential elements of a constitutive representation for the elastoplastic response of polycrystalline solids that is based on plastic flow by restricted slip within crystals. We include the model for slip which depends explicitly on the crystal lattice orientation. To this we add a representation of the spatial distribution of lattice orientations in polycrystals, known as the orientation distribution function (ODF). Methods to average the responses over all of the crystals are then discussed, leading to various approaches for representing the anisotropic strength predicted by the theory. This is followed by a brief description of the implementation anisotropic strength in finite element formulations. We show several examples of the the methodologies for large strain applications and finally summarize.

5.2 Theoretical Background

5.2.1 Mechanical Response of Single Crystals

Inelastic deformations often occur in a regime of temperature and strain rate that results in slip dominated plastic deformations. In this case, we can consider slip on a limited number of slip systems as the principal mechanism for plastic flow and neglect contributions from diffusional...
5.2 Theoretical Background

mechanisms or from mechanical twinning. Strain rate and temperature play important roles in determining which slip systems are activated and the degree to which each contributes to the total deformation. Further, evolution rates of both slip system strength and lattice orientation are affected by the temperature and strain rate through the kinetics associated with the hardening and recovery processes Kocks et al. (1998). There is opportunity to embed the relevant relations within the crystal descriptions for these critical aspects of the behavior while remaining within the framework of plasticity based on slip-dominated behavior.

5.2.1 Kinematic Decomposition

The equations for deformation in a single crystal form the building blocks for the polycrystal behavior. The kinematics begin with mapping of the motion $\mathbf{x} = \mathbf{x}(\mathbf{X}, t)$, where $\mathbf{x}$ are the current coordinates of a point in a crystal, $\mathbf{X}$ are the reference coordinates of the same point, and $t$ is time. The deformation gradient $F$ is obtained from the motion as:

$$F = \frac{\partial \mathbf{x}}{\partial \mathbf{X}}$$

and can be used to formulate measures of the strain and rotation. The velocity gradient $L$ is given as:

$$L = \text{grad} \mathbf{u} = \dot{F} F^{-1}$$

where grad is the gradient operator\(^1\) and $\mathbf{u}$ is the velocity. The velocity gradient may be decomposed into its symmetric and skew parts as:

$$D = \text{sym} L \quad \text{and} \quad W = \text{skw} L$$

to give the deformation rate and spin, respectively. We begin with a schematic picture of the elastoplastic deformation in a single crystal defined by a multiplicative decomposition of the full deformation gradient into elastic and plastic contributions Peirce et al. (1983); Asaro (1983). Slip is actually not such a contrived sequence, but the decomposition provides a framework to develop expressions for the simultaneous rates of elastic and plastic deformation. We adopt a particular version of this decomposition given by Dawson and Marin (1998):

$$F = V^* R^* F^p$$

where $F$ is the deformation gradient, $V^*$ is the elastic stretch, $R^*$ is the lattice rotation, $F^p$ is the plastic deformation gradient coming from slip. This idealization of the deformation is shown schematically in Figure 5.1. The slip planes and slip directions are designated by $\mathbf{\hat{m}}^\alpha$ and $\mathbf{\hat{b}}^\alpha$, respectively. Different crystal types typically exhibit different families of the most active slip systems; ones commonly assumed for fcc metals are shown in Figure 5.2. A combination of slip plane and slip direction is used to define the Schmid tensor:

$$\hat{T}^\alpha = \mathbf{\hat{b}}^\alpha \otimes \mathbf{\hat{m}}^\alpha$$

\(^1\) The gradient is taken with respect to the current coordinates unless otherwise stated.
which is partitioned into its symmetric and skew parts as:

\[
\hat{P}^\alpha = \text{sym} \hat{T}^\alpha \quad \text{and} \quad \hat{Q}^\alpha = \text{skw} \hat{T}^\alpha.
\] (5.6)

From differentiation of Equation (5.4) with time, the combined elastic and plastic deformation can be written in rate form to give for the volumetric part of the deformation:

\[
tr(\dot{D}^c) = tr(\dot{\epsilon}^*)
\] (5.7)

for the deviatoric part of the deformation:

\[
\dot{D}^c = \dot{\epsilon}^* + \dot{D}^p + \epsilon^* \dot{W}^p - \dot{W}^p \epsilon^*
\] (5.8)

and for the spin:

\[
\dot{W}^c = \dot{W}^p + \epsilon^* \dot{D}^p - \dot{D}^p \epsilon^*
\] (5.9)

where \(D\) is the deformation rate and \(W\) is the spin. A superscript \(p\) indicates the part associated with slip and a prime indicates a deviatoric quantity. We have made use of the elastic strains being small to allow us to neglect nonlinear terms of \(\epsilon^*\). This is justified because the elastic moduli of common metallic alloys and minerals are typically several orders of magnitude greater than the associated strengths:

\[
V^* = I + \epsilon^* \quad \text{with} \quad |\epsilon^*| \ll 1
\] (5.10)

### 5.2.1.2 Elastic and Plastic Relations

The constitutive relations for the elastic and plastic behaviors can be written separately. The elastic response is a linear relation between the elastic strain and the Kirchhoff stress, \(\tau\):

\[
\tau = \mathcal{L} \epsilon^* \quad \text{or} \quad \epsilon^* = \mathcal{L}^{-1} \tau
\] (5.11)

where \(\tau = \det(I + \epsilon^*)\sigma\) and \(\mathcal{L}\) reflects the elastic anisotropy appropriate to the crystal symmetry Nye (1957); Kelly and Groves (1970). The plastic response is highly nonlinear owing to the low rate sensitivity exhibited by most structural alloys at low homologous temperature Kocks et al. (1998). These relations are written in the unloaded configuration \(\hat{B}\) depicted in Figure 5.1 by writing the plastic velocity gradient, \(\hat{L}^p = \hat{D}^p + \hat{W}^p\), in this configuration and then introducing the contributions from the individual slip systems:

\[
\begin{align*}
\dot{D}^p &= \sum_\alpha \dot{\gamma}^\alpha \hat{P}^\alpha \\
\dot{W}^p &= \dot{\mathcal{R}} \mathcal{R}^T + \sum_\alpha \dot{\gamma}^\alpha \hat{Q}^\alpha
\end{align*}
\] (5.12)

where \(\dot{\gamma}^\alpha\) is the rate of shearing on the \(\alpha\) slip system. According to a Schmid assumption, a kinetic relation is written between the slip system rate of shearing and the resolved shear stress on the corresponding system, \(\tau^\alpha\), in general as:

\[
\dot{\gamma}^\alpha = \dot{\gamma}^\alpha(\tau^\alpha, g^\alpha)
\] (5.13)
5.2 Theoretical Background

Figure 5.1: Elastoplastic decomposition showing the contributions from plastic slip, rotation and elastic stretch.

where $g^\alpha$ is the strength of the $\alpha$ slip system (commonly referred to as the critical resolved shear stress). The resolved shear stress is the component of the deviatoric crystal stress acting on the slip plane and in the slip direction:

$$
\tau^\alpha = \hat{b}^\alpha \cdot \tau'^c \hat{m}^\alpha = \text{tr}(\tau'^c \hat{P}^\alpha)
$$

A power law approximation to Equation (5.13) is frequently invoked for a limited range of strain rates about a reference value. We refer the reader to Kocks et al. (1998) for a thorough discussion of this power law kinetic equation, especially with regard to its limitations in modeling polyslip. Efficient and accurate methods for integrating Equations (5.4)–(5.14) (or alternative but closely related ones) have been developed by a number of researchers. Peirce, Asaro and Needleman Peirce et al. (1983) presented a methodology in 1983, which has been followed using a variety of strategies by others Rashid and Nemat-Nasser (1990); Maniatty et al. (1992); Kalidindi et al. (1992); Cuitino and Ortiz (1992); Steinmann and Stein (1996); Marin and Dawson (1998).

5.2.1.3 Single Crystal Plastic Stiffness

Equations (5.12), (5.13), and (5.14) can be combined to give a relation between the crystal deviatoric stress and the plastic deformation rate:

$$
\hat{D}^p = \mathcal{M}(r, \tau'^c) \tau'^c
$$

where $\mathcal{M}(r, \tau'^c)$ is the crystal plastic compliance. Equation (5.15) may be inverted only if there exists a sufficient number of independent slip systems. For an incompressible plastic deformation there are five independent components of the deformation rate (the five deviatoric components). In a local coordinate system aligned with the slip system (axes aligned
with the slip direction, slip plane normal and a vector perpendicular to these), a slip system shearing is described by a single component of deformation rate, and thus contributes only one independent number to the five that comprise the crystal deformation rate. Due to crystal symmetry, it often is the case that several slip systems contribute to the same component of the crystal deformation rate, and thus are linearly dependent. For $M^c$ to be invertible, the slip systems must be arranged spatially such that there is at least one slip system contributing to each component of the crystal deformation rate. In such a case, Equation (5.15) can be inverted, giving

$$\tau' = C^c(r, \tau'^c) \hat{D}^p$$

(5.16)

If this is not the case, the $M^c$ is said to be rank deficient. Face-centered cubic crystals such as copper and aluminum frequently are modeled with the set of twelve \{111\} $< 110 >$ slip systems, which possesses five independent deformation modes. Hexagonal close packed crystals, such as $\alpha$-titanium or magnesium, are at times modeled with a combination of basal and prismatic slip systems that do not permit extension along the crystal c-axis. This combination of slip systems possesses only four independent deformation modes, making $M^c$ rank deficient (singular).

### 5.2.1.4 Evolution of Crystal Strength and Lattice Orientation

The evolution of slip system strength is governed by an empirical relationship that is chosen to match observed behavior of the material of interest. A very general form of the hardening equation is written as:

$$\dot{g}^\alpha = \sum_\beta h^{\alpha\beta} \dot{\gamma}^\beta$$

(5.17)

where $h^{\alpha\beta}$ is a matrix of hardening coefficients that depend on the slip system shearing rates, the temperature, and the slip system strengths. A commonly used equation that is motivated by
5.2 Theoretical Background

A Voce hardening law Kocks (1976) and that specifies the same strength for all slip systems is:

\[
\dot{g} = \Theta_0 \left( \frac{g_s - g}{g_s - g_0} \right) |\dot{\gamma}| \tag{5.18}
\]

where \(|\dot{\gamma}|\) is the net slip system shearing rate and \(\Theta_0, g_0,\) and \(g_s\) are parameters. Often \(g_s\) represents a saturation strength that itself depends on strain rate and temperature. Equation (5.18) is numerically integrated.

Each orientation can be updated by writing Equation (5.12) to give the rate of lattice reorientation \(\dot{R}^*\):

\[
\dot{R}^* = (\hat{W}^p - \sum_{\alpha} \dot{\gamma}^\alpha \hat{Q}^\alpha) \cdot R^* \tag{5.19}
\]

and then integrating with a standard method for ordinary differential equations.

5.2.2 Lattice Orientation Distributions for Polycrystals

Metal parts usually are polycrystalline with crystal dimensions usually much smaller than the characteristic dimensions of the part. As discussed in Section 5.1, a macroscopic sample of material is considered to be sufficiently large to contain enough grains to define an orientation distribution. To do this, we first specify the orientation of an individual grain with an angle-axis representation given by Frank (1988); Becker and Panchanadeeswaran (1989); Morawiec (1990); Neumann (1991); Kumar and Dawson (1998). (Alternatively, a set of Euler angles accomplishes the same objective.) The orientation distribution function (ODF), \(A(r)\), delivers the volume fraction of a sample having lattice orientation within the volume in orientation space \(\Omega_f\):

\[
\nu_f = \int_{\Omega_f} A(r) \, d\Omega \tag{5.20}
\]

The ODF must be initialized to reflect the state of the material at the beginning of a simulation; a number of methodologies exist for this purpose Kocks et al. (1998). An example of an ODF is shown in Figure 5.3 for rolled stainless steel.

Knowing \(A(r)\) at any time permits the computation of various average quantities at the macroscopic level. For example, the volume averaged Cauchy stress is given by:

\[
\sigma = \int_{\Omega} A(r) \sigma^c(r) \, d\Omega = \frac{1}{B} \int_B \sigma^c(x) \, dB \tag{5.21}
\]

and the volume averaged deformation rate is:

\[
D = \int_{\Omega} A(r) D^c(r) \, d\Omega = \frac{1}{B} \int_B D^c(x) \, dB \tag{5.22}
\]

where the superscript \(c\) refers to the value in an individual crystal and \(B\) is the physical volume of the polycrystal. It is common practice to approximate the ODF with a discrete sample of crystals, each with an orientation randomly drawn from the ODF. In some instances these are
given a weighting, designated here by $w^c$. A critical aspect of the use of polycrystal plasticity theory is the direct use of the crystallographic texture as one of the measures of state for the material. Crystal orientations may be measured by one of several diffraction techniques (e.g. electron back-scattered diffraction (EBSD), X-ray diffraction (XRD) and neutron diffraction (ND)) and used to determine the orientation distribution function.

### 5.2.2.1 Evolution of Texture

The microstructural state of a metal evolves over the course of an imposed deformation. The macroscopic evidence of this can be the elevation of the flow stress, which has its microscopic origins in both the texture evolution and in the dislocation structure (density). Within the framework provided by polycrystal plasticity, that portion of the increase in apparent strength stemming from the changes in texture is referred to as geometric hardening, and is quantified by the Taylor factor. That portion due to changes in the slip system strength, $g$, is referred to as material hardening.

To evolve the texture with deformation, the ODF must be updated appropriately. This may be accomplished by integrating the evolution equation for the ODF Clément (1982); Kumar and Dawson (1995, 2000) using:

$$\frac{\partial A}{\partial t} + \nabla A \cdot \mathbf{v} + A \nabla \cdot \mathbf{v} = 0$$

In Equation (5.23) the rate of change of the ODF depends on the reorientation velocity $\mathbf{v}$, which in turn is determined from the slip system activity. For Rodrigues’ space, in particular,
5.2 Theoretical Background

the velocity developed under spin $\omega$ is:

$$v = \frac{1}{2} (\omega + (\omega \cdot r) r + \omega \times r)$$ (5.24)

in which the spin $\omega$ is computed from the slip system shearing rates as:

$$\omega = \text{vect} \left( \hat{W}_p - \sum_{\alpha} \dot{\gamma}_\alpha \hat{Q}_\alpha \right)$$ (5.25)

and $\hat{W}_p$ is the crystal spin given in Equation (5.12).

5.2.3 Mechanical Response of Polycrystals

The formulation described here is one in which the crystals are much smaller than the finite elements. Any element could contain a population of crystals sufficiently large in number to reliably define a crystallographic orientation distribution function. In this case, elements are macroscopic in nature with properties averaged over microscopic distributions. The macroscopic motion defined by the element motions must be linked with the response of crystals assumed to exist within it. Linking assumptions perform this function (these are also referred to as mean field assumptions). A commonly used example is the isostrain assumption attributed to Taylor (1938). In this section we summarize the equations for single crystal behavior, and review procedures for averaging the single crystal responses to obtain the behavior of a polycrystalline aggregate. The desired result is a description of the anisotropic strength written as a relation between the average stress and deformation.

We now proceed to define the material response for a polycrystalline aggregate. Besides the single crystal behavior outlined in the preceding section, we also need to specify rules for the crystal interactions. As the methodologies for using these models in finite element codes that are predominantly ones that are “deformation driven,” we will describe the interaction rules in terms of a partitioning rules: formulae that divide a macroscopic deformation among the crystals of the associated aggregate. The stress at the macroscopic level is the orientational average of the crystal stresses.

5.2.3.1 Partitioning Rules

One of the most challenging aspects of polycrystal plasticity is the task of partitioning the macroscopic strain rate among the crystals that contribute to the orientation distribution for that sample of the material (that point at which the macroscopic strain rate is defined). This has been an issue dating back to the work of Taylor (1938) and Sachs (1928). We will consider only a few possibilities here, and refer the reader to Kocks et al. (1998) for a more complete discussion.

The extended Taylor hypothesis is that all crystals exhibit the same velocity gradient:

$$L^c = L$$ (5.26)

Identical velocity gradients over the history of a deformation will render identical deformation gradients. In essence, all crystals deform and spin in the same way. Clearly, Equation (5.22)
is satisfied by the more restrictive case of Equation (5.26). The stress is evaluated from Equation (5.21). The linking assumption given by Equation (5.26) is not only quite simple, but also quite realistic for many materials when subjected to strains well in excess of the elastic limit. In particular, the behavior of crystals with a high degree of symmetry.

Lower symmetry crystal types do not obey this simple rule as well, and other formulations are used in lieu of the Taylor hypothesis. For example, one other possibility to link the crystal responses to the macroscopic behavior is to make the assumption that all crystals experience the average stress Chastel and Dawson (1994); Prantil et al. (1995):

$$\sigma^{ce} = \sigma'$$  \hspace{1cm} (5.27)

In this case Equation (5.21) is satisfied as well. Equation (5.22) can be extended to include the spin, taking the form:

$$L = \int_{\Omega} A(r) L'(r) \, d\Omega = \frac{1}{B} \int_{B} L'(x) \, dB$$  \hspace{1cm} (5.28)

This assumption has been applied to the modeling of minerals Dawson and Wenk (2000) and to metal alloys having one phase with hexagonal close-packed (hcp) crystal structure Dawson and Marin (1998), but has not been used to any significant degree for materials comprised of either body-centered cubic (bcc) or face-centered cubic (fcc) crystals.

Another class of theories that are heavily used for bulk deformation processes are those referred to as “self-consistent” theories Chin et al. (1969); Tome et al. (1991). These build on a fundamental solution for an anisotropic inclusion in a homogeneous effective medium. The effective medium is obtained from the average taken over all inclusions Kocks et al. (1998). A variety of other partitioning rules have been reported for linking the macroscopic stress and motion to the corresponding quantities at the crystal scale. Relaxed constraints models require that the macroscopic value of only some of the deformation rate components are imposed on all crystals Honneff and Mecking (1978); Mathur et al. (1990) and are intended for polycrystals with flattened or elongated grains. The constrained hybrid model Parks and Ahzi (1990) was constructed to model the deformation of crystals whose slip systems are inadequate to allow an arbitrary deformation (five independent, non-zero components of the deformation rate). Increases in computational power associated with scalable parallel capabilities (computers and programs) have provided the means to build aggregates of crystals where each crystal is resolved by one or more elements Beaudoin et al. (1995a); Sarma and Dawson (1996); Mika and Dawson (1999).

### 5.2.3.2 Polycrystal Plastic Stiffness

Given the texture (the ODF), the single crystal compliance, and a rule for partitioning the macroscopic deformation among the constituent crystals, it is possible to construct the macroscopic plastic stiffness, $C^\tau$, and macroscopic plastic compliance, $M^\tau$, as:

$$\tau' = C^\tau(A(r), \dot{\bar{D}}^P) \dot{\bar{D}}^P \quad \text{and} \quad \dot{\bar{D}}^P = M^\tau(A(r), \dot{\bar{D}}^P) \tau'$$  \hspace{1cm} (5.29)

This is a linearization about the current loading point, as specified by the deviatoric stress active in each crystal. For example, a form of the Taylor assumption to the plastic deformation
rate (in contrast to the total deformation rate), Equations (5.16), (5.26), and (5.21) can be combined to give the plastic stiffness in the form:

$$C^a = \int_\Omega A(r)C^c(r, D^e) \, d\Omega = \int_\Omega A(r)[M^c(r, D^e)]^{-1} \, d\Omega$$

(5.30)

as an orientation average. Similarly, if an equilibrium assumption (lower bound) is invoked, Equations (5.16), (5.27), and (5.22) can be combined to give the plastic compliance as:

$$M^a = \int_\Omega A(r)M^c(r, D^e) \, d\Omega.$$  

(5.31)

In Equation (5.29), $C^a$ is a secant stiffness for the macroscopic plastic response obtained by inverting either $M^c$ or $M^a$ according to Equation (5.30) or (5.31). The inverse of $M^c$ will exist provided that every crystal has five independent slip systems. For the inverse of orientational average of the compliances, $M^a$, to exist there need be only five independent slip systems for the entire aggregate. Thus, individual crystals may have singular compliances, but when averaged over the orientation distribution, the resultant compliance will be invertible provided there totals five independent slip systems.

That $C^a$ depends on the direction of deviatoric stress (or equivalently on the direction of the deformation rate) is an expression of the anisotropic strength. For a prescribed rate of work, this results in a surface in stress space. (One can view this as a plastic potential Houette (1994); Kocks et al. (1998).) An important attribute of the flow surface is that the strain rate is normal to the surface. This permits approximation of the flow surface as the inner envelope of planes that are perpendicular to the strain rate direction and containing the stress. This surface will be everywhere convex.

### 5.3 Macroscopic Criteria for Anisotropic Strength

#### 5.3.1 Generalities

Mathematical descriptions of strength anisotropy typically embody criteria that define flow surfaces or potentials at the continuum scale. The criteria emerged as extensions of those proposed by Tresca (1864), Huber (1904), von Mises (1913), and Hencky (1924) and recover one of these in the limit of isotropic response. For metals that exhibit no inelastic volumetric deformation, a yield criterion is a convex surface in deviatoric stress space:

$$f(\sigma^e) = 0$$

(5.32)

where $\sigma$ is the Cauchy stress and the prime indicates its deviatoric part Gurtin (1981); Khan and Huang (1995). In the rate independent limit: should the deviatoric stress lie on the surface, plastic deformations are possible; should the deviatoric stress lie inside the surface, plastic deformation will not occur; the deviatoric stress cannot lie exterior to the surface. Rate dependence implies the existence of families of surfaces Kocks et al. (1998) parameterized by a variable such as the Zener-Holloman parameter or the Fisher factor, $F$:

$$f(\sigma^e, F) = 0.$$  

(5.33)
With this representation of the inelastic behavior, plastic flow occurs at any stress level, but because of the highly nonlinear dependence of the flow surface on the rate parameter, the plastic deformation rate drops off rapidly as the stress diminishes. The plastic deformation strain rate may be derived from the yield criterion if the criterion is constructed as a potential function. In this case, the plastic deformation rate is:

\[ D_p' = \frac{\partial f}{\partial \sigma'} \]

and \( f \) is proportional to the plastic rate of work. Further, \( f \) is a dual potential Hill (1987); Houtte et al. (1989) such that:

\[ \sigma' = \frac{\partial f}{\partial D_p'} \]  

The relation between the direction of straining and the deviatoric stress (e.g., the flow law) is a critical element of the complete model of plastic flow. Many aspects of the mechanical behavior depend strongly on this relation, such as the contraction ratio (or R-value\(^2\)). A material’s formability depends on the stability of homogeneous deformation, which may be directly linked to the flow law. Accurately quantifying the relation between the applied stress and the resulting plastic deformation rate thus is a primary, if not the primary, reason for constructing models of the anisotropic strength.

Anisotropy implies that the surface given by \( f \) cannot be a function solely of the invariants of the stress. Rather, the parameters in the criteria are linked to a specific coordinate basis attached to the material. There are recurrent difficulties encountered with anisotropic strength descriptions, including: the difficulty in obtaining both the correct curvature and size Lademo et al. (1999); Wu et al. (2003); the difficulty in evaluating parameters from mechanical testing alone, especially parameters associated with shear components of the stress in plane stress implementations, or for general stress states (five deviatoric stress dimensions rather than just the three that are needed in plane stress); and the extension of criteria to include the evolution of strength with deformation. Piecewise surfaces offer alternatives to analytical expressions in being able to better capture the surface geometry. The improvement comes at the expense of added complexity in the algorithms that compute properties from the surface (such as the elastoplastic stiffness matrix used in finite element codes). This added complexity, and its associated cost, must be weighed against the increase in fidelity of the representation.

The isochoric deformation rate and deviatoric stress both exist in five-dimensional spaces. The reduction from six to five components to form bases for the spaces is not unique and many choices work equally well. Here we make a common choice in which the components of the deviatoric deformation rate tensor, \( D' \), and the deviatoric Cauchy stress\(^3\), \( \sigma' \), are given by Kocks et al. (1998):

\[
(d_1, d_2, d_3, d_4, d_5) = \left( (D'_{22} - D'_{11}), \sqrt{3} D'_{33}, 2D'_{23}, 2D'_{31}, 2D'_{12} \right) / \sqrt{2} \]  

\[
(s_1, s_2, s_3, s_4, s_5) = \left( (\sigma'_{22} - \sigma'_{11}), \sqrt{3} \sigma'_{33}, 2 \sigma'_{23}, 2 \sigma'_{31}, 2 \sigma'_{12} \right) / \sqrt{2}.
\]

\(^2\) The R-value is defined as the ratio of the width to thickness strain increments in sheet metal under tensile loading.

\(^3\) Alternatively, we could use the Kirchhoff stress, \( \tau' \), obtained by scaling the Cauchy stress by the determinant of the elastic stretch.
5.3 Macroscopic Criteria for Anisotropic Strength

The set of reference deformation rates defines a four dimensional surface within five dimensional deviatoric deformation rate space.

5.3.2 Yield Surfaces Defined by Expansions

Instead of representing the strength as a surface in stress space, it is possible to express it as a surface in strain rate space, following from recognition that the plastic rate of work is a dual potential Hill (1987); Houtte et al. (1989). A variety of functional forms have been proposed, including ones at a macroscopic scale Arminjon and Bacroix (1991); Barlat et al. (1993). Here we cite ones developed by Arminjon and coworkers Arminjon and Bacroix (1991); Arminjon and Imbault (1994); Arminjon et al. (1994) in which the mapping is related to the texture. The plastic potential takes the form:

\[ f(D^p) = \sum_{ijkl} v_{ijkl} \frac{D^p_i D^p_j D^p_k D^p_l}{\|D^p\|^3} = \sum_K v_K \frac{\Upsilon_K(D^p)}{\|D^p\|^3} \]  

where there are 22 \( \Upsilon_K \) terms for the restricted case of orthotropic sample symmetry, that are each fourth-degree in the components for the plastic deformation rate as shown by Bacroix and Gilormini Barlow and Gilormini (1995). The coefficients \( v_K \) scale the individual contributions to \( \Upsilon \) and are determined from the harmonic texture coefficients \( C_i \):

\[ v_K = \sum_{i=1}^{13} \sum_{j=1}^{13} C_i \beta_{ij}^{(k)} \]  

Here \( \beta_{ij}^{(k)} \) are coefficients that are computed from the plastic response of crystals independent of the texture. This is done by minimizing the difference between the plastic rate of work computed from the potential function and the plastic rate from crystal plasticity for a suite of deformation modes. Bacroix and Gilormini Barlow and Gilormini (1995) indicate that 13 terms are sufficient in the expansion to provide adequate quantification of the orientation distribution function. The \( \beta_{ij}^{(k)} \) coefficients are subsequently weighted by the texture coefficients to obtain the final scaling for the potential. The evolution of the yield surface with straining is handled by updating the texture coefficients \( C_i \) according to crystal plasticity, and then re-computing the \( v_K \) coefficients Arminjon and Imbault (1994). A comparison between surfaces defined by the expansion and one defined by the isostrain hypothesis (Taylor model) is shown in Figure 5.4.

An approach that is similar to the fourth-degree plastic potential outlined in the previous section has been developed by Van Houtte and coworkers Houtte et al. (1989); Houtte (1994); Hiwatashi et al. (1997). In this case the plastic potential is expanded in terms that are of sixth-degree in components of the deformation rate. The plastic potential is written as:

\[ f(\hat{D}^p) = \tau^e D_{\nu M} G(\hat{D}^p) \]  

where \( G \) is an expansion in terms of the normalized deformation rate as:

\[ G(\tilde{D}^p) = G(\tilde{d}^p) = G_{ijklmn} \tilde{d}_i \tilde{d}_j \tilde{d}_k \tilde{d}_l \tilde{d}_m \tilde{d}_n \]  

We’ve used somewhat different terminology than the cited authors to be consistent within this paper.
Note that $\tilde{d}^p$ is a five-dimensional vector form of $\tilde{D}^p$. The expansion $G$ is computed from the texture by determining the coefficients $G_{ijklmn}$ that best fit a surface defined by the average scalar Taylor factor $\bar{M}$:

$$G(\tilde{D}^p) \approx \sqrt{\frac{2}{3} \bar{M}(\tilde{D}^p)}.$$  

(5.42)

For crystals having a single slip system strength, $g$, a tensorial crystal Taylor factor $M^c$ corresponding to a straining direction given by $\tilde{D}^p$ is defined such that:

$$\sigma^c = M^c(\tilde{D}^p)g$$

(5.43)

which can be projected onto the straining direction to give a scalar crystal Taylor factor $M^c$ as:

$$M^c = \frac{\text{tr}(\sigma^c D^c)}{g}$$

(5.44)

An average scalar Taylor factor is computed using the orientational average given by:

$$\bar{M} = \int_{\Omega} M^c(r) A(r) d\Omega$$

(5.45)

This representation has been applied to a number of materials, including both aluminum and steel alloys. Recent efforts have included the incorporation of state variables to describe the effects of dislocations structure Hiwatashi et al. (1997).

### 5.3.3 Yield Surfaces Defined by Hyperplanes

Maudlin and coworkers Maudlin et al. (1995, 1996) have developed a methodology that utilizes a flow surface defined by hyperplanes in deviatoric stress space. The shape of the flow
5.3 Macroscopic Criteria for Anisotropic Strength

surface is defined by the inner envelope of planes:

\[ f_\chi \equiv \zeta_{ij} \sigma_{ij} - Y^\chi = 0, \quad \chi = 1, 2, \ldots, m \]  \hspace{1cm} (5.46)

and then scaled to the appropriate size according to:

\[ \sigma_{ij} = \tilde{\sigma}_{ij} \frac{Y}{\bar{M}} \quad \text{and} \quad Y^\chi = \tilde{Y}^\chi \frac{Y}{\bar{M}} \]  \hspace{1cm} (5.47)

Here \( \bar{M} \) is the average Taylor factor Kocks et al. (1998), \( Y \) is an effective (reference) flow stress that establishes the overall size of the yield surface, \( Y^\chi \) is the minimum distance from the origin of deviatoric stress space to the \( \chi \) hyperplane, and \( \zeta_{ij} \) are the coefficients of the normal vector to the \( \chi \) hyperplane. The hyperplane data is determined by probing of the flow surface using crystal plasticity, together with the measured crystallographic texture for a material, to find the flow strength in a number of stress directions. This set of points is used in a tessellation algorithm to determine a piecewise surface defined by Equation (5.46). A hyperplane surface is shown in Figure 5.5 for a strongly anisotropic zirconium material. Also needed is the flow law, which is obtained by taking \( f_\chi \) as a plastic potential and differentiating with respect to the stress:

\[ \dot{\epsilon}^p_{ij} = \sum_{\chi=1}^{m_{act}} \dot{\lambda}^\chi \zeta_{ij}^\chi \]  \hspace{1cm} (5.48)

where the linearity of \( f_\chi \) with respect to \( \sigma_{ij} \) has been exploited, \( \dot{\lambda}^\chi \) are referred to as the loading scalars, and \( m_{act} \) is the number of active hyperplanes, as explained below.

Maudlin and coworkers implemented the methodology in a formulation for dynamic analyses (an explicit formulation). The algorithm consists of five tasks that are performed for each time step. At each step, the algorithm first determines if the stress at the beginning of the step satisfies the equation of any of the hyperplanes (these are called admissible planes). If this is so, it chooses the most active admissible hyperplanes based on those having the largest

---

**Figure 5.5:** Hyperplane representation of a yield surface. Left diagram shows the surface projected onto the \( \tau_1, \tau_2, \tau_5 \) hyperplane; right diagram shows the surface projected onto the \( \tau_3, \tau_4, \tau_5 \) hyperplane; Reproduced from Tome et al. (2001).
loading scalars. The algorithm then determines if the material is being unloaded for any of the admissible hyperplanes (those admissible planes that are not unloading are called candidate planes). Using the candidate planes, it determines the proportionality scalars from the consistency condition and back substitutes into the hyperplane equations to determine the stress. Finally, a corrective step is performed to reduce numerical errors. The flow surface has been evaluated, for example, for clock-rolled zirconium plate that demonstrates a high degree of mechanical anisotropy owing to its strong basal texture.

5.3.4 Isoparametric Flow Surface

Piecewise flow surfaces in either stress or deformation rate space offer alternatives to analytical expansions for representing anisotropic strength. An anisotropic flow surface in stress space gives the flow stress directly via its coordinates. For associated flow, the deformation rate is normal to the surface. It also is possible to define a surface in deformation rate space corresponding to a reference deformation rate of fixed magnitude. All of the possible deformation modes exist as points on the hypersphere. The flow stress is represented as a field over the surface. The two representations possess comparable information but in different formats. We proceed using the deformation rate space representation, introducing an isoparametric mapping for both the surface and the data on it.

The coordinates of points on the hypersphere are defined by the mapping:

\[ d_i = [N_D(\zeta)] \{D_i\} \] (5.49)

where \( N_D(\zeta) \) represents a set of continuous piecewise interpolation functions and \( \zeta \) are local coordinates on a four dimensional patch used in the interpolation. \( \{D_i\} \) are the nodal point values in the five dimensional space. We employ low-order Lagrange elements for this purpose:

\[ N_D^{ijkl} = N^i(\zeta_1) N^j(\zeta_2) N^k(\zeta_3) N^l(\zeta_4) \]

\[ i, j, k, l = 1, 2 \] (5.50)

where for each dimension

\[ N^1(\varphi) = 1 - \varphi \quad \text{and} \quad N^2(\varphi) = \varphi \quad \text{with} \quad 0 \leq \varphi \leq 1. \] (5.51)

The components of the deviatoric stress are represented over the reference deformation rate sphere using the isoparametric representation:

\[ s_i = [N_D(\zeta)] \{S_i\} \] (5.52)

where \( \{S_i\} \) are the nodal point values for the \( i^{th} \) component of \( (s) \). With this representation the flow stress is defined for all possible values of the reference deformation rate, where the reference deformation rate has a fixed magnitude. The interpolation functions used give the representation a number of useful properties: the patches cover the hypersphere once and only once; fields over the hypersphere possess \( C^0 \) continuity – the flow stress is everywhere continuous over the hypersphere of deformation modes; the minimum and maximum values always are nodal values – a search of nodal values will always find the extreme values; and, coefficients in the parameterization correspond to nodal values, which in turn are the flow
stress values at the deformation rate associated with the nodal point. The flow surface is shown on subspaces in Figure 5.6 for a rolled stainless alloy where values are computed using an isostrain linking hypothesis.

We assume that the flow stress for rates other than the reference deformation rate can be derived from the flow stress at the reference deformation rate according to a kinetic law for the plastic flow. For the model being used here, the magnitude of the stress is related to the magnitude of the deformation rate through a power law:

\[
\frac{s}{s_{\text{ref}}} = \left( \frac{d}{d_{\text{ref}}} \right)^m
\]  \hspace{1cm} (5.53)

where \( s \) and \( d \) are the norms of \( \{s\} \) and \( \{d\} \), respectively. The scaling given in Equation (5.53) follows directly from a power law kinetic relation for slip:

\[
\frac{\dot{\gamma}^\alpha}{\dot{\gamma}_{\text{ref}}^\alpha} = \left( \frac{\sigma^\alpha}{\tau_{\text{ref}}} \right)^{1/m}
\]  \hspace{1cm} (5.54)

when averaged over an aggregate of crystals.

The sensitivity of stress to changes in the direction of the deformation rate, at deformation rate magnitude and fixed state, \( \frac{\partial s}{\partial d_j} \), is a critical anisotropic characteristic of a material. Where this sensitivity is high, small changes in the straining direction can induce large changes in the stress. Conversely, where the sensitivity is low, the stress changes little as the deformation rate changes rapidly. The latter corresponds to a vertex of a flow surface in stress space while the former corresponds to a facet. This sensitivity is central to the robustness of numerical
procedures that iterate on the nonlinear material response. Computing the sensitivity is readily accomplished from the surface through the use of isoparametric mappings:

\[
\frac{\partial S_i}{\partial d_j} = \left[ \frac{\partial N_D}{\partial \xi} \frac{\partial \xi}{\partial d_j} \right] \{S_i\} \tag{5.55}
\]

The values of \( \frac{\partial N_D}{\partial \xi} \) are the same for all patches and thus need be computed but once. The values of \( \frac{\partial \xi}{\partial d_j} \) differ for every patch, but depend only on coordinates of the sphere which remain fixed as long as the same reference deformation rate is used. The gradient given by \( \frac{\partial S_i}{\partial d_j} \) does not give any information about the changes in stress with changes in the magnitude of the deformation rate. The derivative in the radial direction can be determined from the scaling law.

### 5.3.5 Direct Polycrystal Plasticity Implementation

It is also possible to embed a representation of the texture within every finite element and compute the stress and stiffness as needed using the equations for the single crystal response and the linking hypothesis. The most straightforward approach is to utilize the stiffness matrix computed in Equation (5.30) or (5.31) directly in the equilibrium residual given by Equation (5.65). Provided that the texture is not too strong, implying that there are not sharp corners on the flow surface, this will work adequately. A second approach offers advantages for more extreme anisotropy, and for loading situations such as arise in sheet forming. This approach is a hybrid methodology that involves two principal differences from the more standard finite element formulations. The first difference is that residuals invoke domain decomposition in which traction equilibrium is enforced between subdomains of the body. The second is the use of prescribed interpolation for the stress. In this case, a residual is formed from Equation (5.29) based on the average stiffness from Equation (5.30) or (5.31):

\[
\int_B \tau^c \cdot (C^a D' - \tau^e) \, dB = 0. \tag{5.56}
\]

Here, \( \tau^c \) are the weights associated with this residual. Trial functions are introduced in this residual for the stress (deviatoric and spherical portions) with approximating functions:

\[
\tau^e = [N^\sigma]\{\Xi\} \quad \text{and} \quad p = [N^p]\{\Pi\}, \tag{5.57}
\]

where \( \{\Xi\} \) and \( \{\Pi\} \) are nodal values for \( \tau^e \) and \( p \), respectively, and \([N^\sigma]\) and \([N^p]\) are the associated interpolation functions. The stress interpolation is constructed such that equilibrium within an element is satisfied a priori. The residual is written in matrix form using the trial functions and solved for the stress degrees of freedom:

\[
\{\Xi\} = [H]^{-1} [R]^T \{U\} \tag{5.58}
\]

where \([H]\) and \([R]\) are given by:

\[
[R] = \int_B [N^w]^T [N^\sigma] \, dB \tag{5.59}
\]
and

\[ [H] = \int_B [N^\sigma]^T [M^\sigma] [N^\sigma] dB \quad (5.60) \]

and \([N^u']\) gives the derivatives of the velocity interpolation functions. The direct approach delivers a stiffness that depends on the deformation rate and material state (texture and slip crystal strengths) at the corresponding macroscopic point. While the flow surface does not appear in its totality, the stiffness matrix gives equivalent information at the active deformation rate at each point where it is evaluated.

### 5.4 Numerical Implementations

#### 5.4.1 Balance Laws

At the global level, the simulation of a deforming material is based on balance of momentum, conservation of mass, and conservation of energy coupled with the appropriate kinematic and constitutive relations. The momentum and mass expressions are used to formulate the solution for the workpiece motion, while the energy equation is used to determine the temperature field. The focus here is on the mechanical response, and although the thermal and mechanical responses often are tightly coupled, only the mechanical formulation will be discussed.

Balance of linear momentum, neglecting body forces, may be written in local form as:

\[ \text{div } \sigma = \rho \frac{Du}{Dt} \quad (5.61) \]

and conservation of mass as:

\[ \frac{D\rho}{Dt} + \rho \text{div } u = 0 \quad (5.62) \]

where \(\rho\) is the density, and \(\text{div}\) indicates the divergence. Essential and natural boundary conditions, respectively, are given by:

\[ u = \bar{u} \quad \text{on } \partial B_u \quad (5.63) \]

for the velocity vector \(u\), and

\[ \sigma n = t = \bar{t} \quad \text{on } \partial B_t \quad (5.64) \]

for the traction vector \(t\) where \(\partial B(= \partial B_u \cup \partial B_t)\) is the boundary of the body \(B\).

#### 5.4.2 Finite Element Formulations

Finite element formulations develop solutions from the weak forms of the governing equations. For elastoplastic deformations of solids, residuals leading to the weak forms may be constructed from the balance laws presented in the previous section and the constitutive equations, with the latter being used principally in hybrid formulations. Our discussion here begins
with the former as these appear in virtually all formulations. For linear momentum, the weak form resulting from Equation (5.61) is:

$$\int_B \operatorname{tr}(\sigma' \operatorname{grad} \varphi) \, dV - \int_B p \operatorname{div} \varphi \, dV = \int_{\partial B} \bar{t} \cdot \varphi \, dS - \int_B \rho \frac{Du}{Dt} \cdot \varphi \, dV \tag{5.65}$$

where $\varphi$ is a vector weighting function for the velocity field $u$. For conservation of mass, the weak form resulting from Equation (5.62) is:

$$\int_B \left( \frac{D\rho}{Dt} + \rho \operatorname{div} u \right) \phi \, dV = 0 \tag{5.66}$$

where $\phi$ is a scalar weighting function. Equation (5.66) explicitly appears in formulations which neglect the elastic response and treat the motion as incompressible (with constant density). In formulations that include elasticity, conservation of mass is implicit in the volumetric response of the elasticity relations in that the density may be recovered from Equation (5.62) once the motion is determined from Equation (5.65).

The evolution of the state description must be performed in concert with the computation of the motion. Here, the state description of interest is that used to quantify the anisotropic strength. As presented in the following sections, this might involve updating the lattice orientations and slip system strengths, the yield surface description, or parameterization of a plastic potential. In addition, the stress (or equivalently, the elastic strains) must be advanced over each time step.

The specific choices made in the finite element implementation stem from the features retained in the above equations. A few pertinent points are whether or not the application involve dynamic loading, whether or not the deformations are large, and whether or not the elastic response may be neglected. The inclusion of dynamic terms often favors the use of explicit time integration methods, especially in relation to resolving stress waves. Structural dynamics applications, while not necessarily focused on stress wave propagation applications, often utilize formulations that embody explicit integration of the governing equations because of the nature of contact conditions. The methods associated with these applications have become quite robust, and have been employed in applications that are essentially static (in that the inertial effects are insignificant). Handling the geometric changes associated with large deformations is complex, and motivates a number of different approaches. First, there is the issue of kinematics, which can be correctly handled using Lagrangian, Eulerian, or updated Lagrangian reference frames Belytschko et al. (2000); Bonet and Wood (1997). Work conjugate stress and deformation measures are known for the different reference frames. A more subtle issue is associated with specification of constitutive equations that render responses that do not differ for different observers (observers whose relative positions differ by time-varying translation and rotation). Two features of the constitutive equations assure frame indifference are the use of “hyper” (involving a potential function) rather than “hypo” (cast in rate form) elasticity relations and the use of rotated stress variables. The decision to include the elastic response is largely governed by the loading and by the information sought from the simulation. Loosely stated, only applications involving large strains generated by monotonic loading are candidates for analysis using a viscoplastic approximation. When the loading is monotonic the stress over an increment of strain may be nearly constant, changing slowly as a consequence of hardening and shape changes.
Tasks common to all implementations involve the computation of the stress and/or stiffness. In the residual given by Equation (5.65) this enters in the replacement of the stress by an expression involving the velocity field. This expression is, of course, developed from the constitutive equations. For implicit formulations, the expression embodies the tangent stiffness Belytschko et al. (2000). The actual tangent stiffness that appears in modern formulations is adapted from the continuum form to account for the discretized space associated with the finite element approximations. This form often is called the consistent tangent stiffness, with the consistent descriptor referring to the gradients being taken with respect to the spaces employed in the finite element discretization. For explicit formulations, similar information enters in the methodology to integrate the constitutive equations. This is commonly referred to as the stress update algorithm, and for elastoplastic behaviors, the radial return method frequently is used. In either case, information regarding the stress, the plastic straining direction, and the rate of change of state variables (from strain hardening) must be computed.

5.5 Applications

We illustrate with three analyses: the first examines the explosive forming of a zirconium plate and uses a piecewise yield surface representation; the second uses a direct implementation of crystal plasticity to investigate the influence of crystallographic texture on formability of an aluminum alloy under various loading modes; the last examines lattice strains under macroscopically inhomogeneous deformation of a stainless steel alloy using a multiscale discretization. These provide a small sample of the potential applications of crystal plasticity as the means of evaluating strength anisotropy.

5.5.1 Application to Explosive Forming

A striking example of the influence of anisotropic strength has been reported by Maudlin and coworkers Tome et al. (2001). Explosive loading of a zirconium plate has been simulated that causes the plate to deform into a conical shape. The initial strength of the plate is strongly anisotropic owing to a pronounced basal texture in the plate normal direction. To contrast the differences in response predicted by isotropic and anisotropic models of the strength, simulations were performed with an isotropic analytical (von Mises) model and with an anisotropic piecewise representation of the yield surface.

The initial geometry of the explosive punch problem is simply that of a flat plate. The punching operation was simulated in three dimensions with the finite element EPIC code Johnson et al. (1997) using a Mie-Gruneisen equation-of-state and the rate-dependent mechanical threshold strength (MTS) flow stress model Follansbee and Kocks (1988). Loading is applied axissymmetrically to a flat, circular disc having a 50 mm radius and a 2.75 mm thickness. The disc is backed by a cylinder of high explosive 100 mm in length. The high explosive is point initiated on the centerline of the cylinder at the end opposite to the disc.

The punched shapes after 200 µs are presented in Figures 5.7 and 5.8; additional deformations after this time are purely elastic. Shown on the deformed meshes are contours of the effective (von Mises) plastic strain and a measure of the rigid body rotation given by an
angle-axis convention (a Rodrigues vector.) The isotropic results (Figure 5.7) show the expected overall axisymmetry with strains computed for the final shape ranging from 80% at the apex to 300% close to the backend. The rigid body rotation ranges from a few degrees at the apex to maximum of $56^\circ$ roughly mid-length. The uniform distribution of Rodrigues vectors emphasizes the axisymmetry of the deformation. The anisotropic results (Figure 5.8) based on the piecewise yield surface show a complete loss of any geometric symmetry and larger plastic strains in the final shape. Strains range from 60% at the apex to 470% localized at the 12 o’clock azimuth. The rigid body rotation ranges from a few degrees at the apex to maximum of $80^\circ$ localized at the 7 o’clock azimuth. The non-symmetrical distribution of Rodrigues vectors illustrates the strength anisotropy of this material stemming from its crystallographic texture.

Figure 5.7: Side view of a punched zirconium shape at 200 $\mu$s with contours of plastic strain (left) and Rodrigues rotation angle in degrees (right). The constitutive modeling used in this calculation included the MTS flow stress model and a von Mises yield surface. Reproduced from an original color image Tome et al. (2001) with loss of scale information. Important is the difference in the deformed geometry when contrasted to the shape shown in Figure 5.8 obtained using an anisotropic model.

5.5.2 Application to the Limiting Dome Height Test

The limiting dome height test is used to assess the formability properties of metals. The test, shown schematically in Figure 5.9, consists of stretching a rectangular specimen over a hemispherical punch. The draw-in is partially constrained by a circular drawbead. Because the test specimen is rectangular, the extent to which it is constrained depends on its width in comparison to the radius of the drawbead. The constraint acts over the entire periphery for a square specimen, producing a state of stress near to balanced biaxial tension. A narrow specimen is gripped only on its ends and the stress state is closer to uniaxial tension. An intermediate width can be chosen so that straining across the width is nearly zero at the centerline and the deformation mode approximates plane strain.
5.5 Applications

Figure 5.8: Side view of a punched zirconium shape at 200 µs with contours of plastic strain (left) and Rodrigues rotation angle in degrees (right). The constitutive modeling used in this calculation included the MTS flow stress model and a piecewise yield surface. Reproduced from an original color image Tome et al. (2001) with loss of scale information. Important is the difference in the deformed geometry when contrasted to the shape shown in Figure 5.7 obtained using an isotropic model.

The effect of processing practice on formability has been documented for AA-2036-T4 aluminum Bryant et al. (1994). The various processing practices induce different textures and different levels of strength anisotropy. The crystallographic texture affects the performance in the limiting dome height test through the influence of strength anisotropy on the thinning of the sheet. Pole figures for two of the practices reported in Bryant et al. (1994) are shown in Figure 5.10. They differ principally in the intensity of the Goss component, with Texture A having stronger Goss component than Texture B.

Simulation of the limiting dome height test was conducted for three specimen geometries using the direct viscoplastic formulation outlined in Section 5.3.5 Dawson and Beaudoin (1997). The geometries were chosen to give loading states at the pole of balanced biaxial tension, uniaxial tension, and plane strain, as described above. The rectangular specimens were oriented such that the rolling direction of the sheet corresponded to the longer dimension of the specimen. The initial textures prescribed for Textures A and B were represented by sets of weighted orientations Beaudoin et al. (1995b). The orientations and weights of the crystals were determined with the popLA code Kocks et al. (1998). Identical hardening parameters were prescribed for each sheet based on tensile tests performed on samples from the sheet. Each specimen was constrained against movement where it contacted the hold-down tooling; the drawbead was not explicitly modeled. Sticking friction was prescribed for the interface between the punch and the specimen, which corresponded to the unlubricated conditions used in the experiments.

The simulations demonstrate how the analyses may be scaled to deliver differing degrees of resolution, both of the specimen geometry and of the texture representation. For the plane strain case, the mesh had 3600 elements and within each element were 256 orientations to
represent the texture. For balanced biaxial tension case, the simulation employed 3750 elements each having 64 crystal orientations for the texture. Finally, for the uniaxial stress case, the simulation was run with only 208 elements and 64 orientations per element. In all cases, the elements used were 8-node bricks, and the crystal type was face-centered cubic. The plane strain geometry simulations were performed on a parallel computer while the uniaxial stress simulations were performed on a single processor workstation.

Shown in Figure 5.11 are the deformed meshes for each material and specimen geometry. The contours display the specimen thickness. For the plane strain mode of deformation, the simulation corresponding to Texture A shows a greater tendency to thin as the dome height increases, as was observed in experiments. For biaxial tension, the thickness of the specimens for the two cases are roughly the same for equivalent dome heights. However, the location of the thinnest region is different for the two. For the Texture A the most highly deformed region lies along the specimen axis aligned with the sheet transverse direction. For Texture B, the most severe straining is at 45° to the rolling direction. Finally, for the narrow specimen deforming under uniaxial stress the thinning is greatest for Texture B, just the opposite trend to that observed under plane strain deformation. The simulations suggest that the relative formability of the same material in different states is dependent on the loading path. This is borne out by experiments Bryant et al. (1994). The simulations demonstrate that the combined finite element–crystal plasticity formulation scales such that the resolution in either geometry or material state can be tailored to the available computing resources.

![Figure 5.9: Schematic drawing of limiting dome height test geometry. The left figure shows the blank as viewed from above. The ratio of dimensions L and W are chosen to give plane strain deformation along a plane that bisects the sheet in the shorter dimension. The right figure shows the sheet as viewed from the side. The punch is shown along with a schematic depiction of the draw beads. The die and blank holder are not shown. Reproduced from Kocks et al. (1998).](image)
5.5 Applications

Figure 5.10: 111 pole figures for the AA2036 aluminum alloy under two processing conditions. Pole figures were used to evaluate an orientation distribution function and subsequently to select weighted orientations for the simulations. Scale is in multiple of random (uniform) distribution. Taken from Kocks, Tome, and Wenk, 1998.

Figure 5.11: Strain distribution for several strip widths. Strip widths were chosen to give states of plane strain (top), balanced biaxial tension (center), and uniaxial tension (bottom). Various levels of discretization were used to demonstrate scalability of computations. Reproduced from Beaudoin et al. (1996).
5.5.3 Bending of a Curved Component

A U-shaped stainless steel part was loaded in situ in a neutron diffractometer in a manner to produce bending at the root of the U. While in the diffractometer, lattice strains were measured at various points in the loading history to serve as a basis for comparison with simulated strains.

A finite element mesh of half of the part is shown in Figure 5.12. Each element is quite large in comparison with the average grain size. To determine the properties for the computation of the anisotropic stiffness matrix, the responses of crystals in a representative aggregate were averaged using the isostrain assumption outlined in Section 5.2.3.1. The stock material from which the specimen was drawn had been rolled, which gave the material a mild texture. This texture was quantified by diffraction and used to determine an ODF. The lattice orientations of crystals comprising the aggregate were chosen randomly from this distribution. Initially, all elements were assigned the identical aggregates. Over the course of the loading history, the lattice orientations of each aggregate evolved according to the deformation specific to its associated element. The stress distribution over the part is shown on the deformed mesh in Figure 5.12 for the instant prior to the final unloading. Clearly, the distribution is consistent with the root being in a state of bending.

The complete mechanical histories for points in the part coincident with diffraction measurements were recorded in the simulation and used to load sets of individually resolved crystals (also shown Figure 5.12). These sets crystals constitute samples taken at a smaller length scale and correspond to diffraction volumes in the experiment. From the data for the different crystal sets, comparison can be made to the results gathered in the in situ measurements after identifying which crystals in the aggregate satisfy the Bragg condition imposed in the experiment. Such comparisons are shown in Figure 5.13 for two crystals sets: ones with (222) and (004) directions aligned with the tensile axis. Comparison are shown between simulation and experiment for the final loaded and unloaded states. Several points are evident:

1. the emergence of a plastic hinge is clear. The simulations capture this overall, but present a transition from tensile to compressive strains that is too abrupt.

2. the strain levels are substantially higher for the (004) crystals than for the (222) crystals or for others.

3. the quality of the agreement improves as one looks from the central fibers to the outside radius.

A possible explanation for the improving quality nearer the outside edges is that the plastic strain increase as one moves outward. The simplistic model of slip is probably more accurate where strains are larger than it is near the neutral axis of loading where the plastic strains are very small.

5.6 Summary

We conclude this chapter with a short list of important aspects of using crystal plasticity to compute strength and especially its anisotropy.
5.6 Summary

Figure 5.12: Left diagram: finite element mesh of the u-shaped specimen showing the average stress. Right diagram: finite element mesh of polycrystal representing the crystals in a diffraction volume.

Figure 5.13: Lattice strains predicted using the isostrain assumption together with the finite element model.

- Strength anisotropy in polycrystalline materials is inherited from the combination of single crystal anisotropy (that itself stems from the nature of restricted slip) and of nonuniform lattice orientation distributions.

- The basic model for slip presented here is very simplistic, but does capture several critical features of the material behavior, including:

  1. principal features of ODFs induced by deformation, and
  2. directional dependence of the yield or flow stress.

- The basic model for slip does not embody possible influence growing out of a relative size dependence between the dislocations (or dislocation structures) associated with slip and the domains over which slip is assumed to be homogeneous (often referred to as the missing length scale).
References


References


References

35, 1565 – 1588.
Sachs G., 1928. ?? Zeitschrift Verein Deut Ing 72, 734–736.
6 Yield Surface Plasticity and Anisotropy

Frédéric Barlat, Oana Cazacu, Michal Życzkowski, Dorel Banabic, and Jeong Whan Yoon

In this chapter, after a concise review of the classical theory of plasticity, the microstructural origins of plastic anisotropy are discussed. Since crystallographic texture is the main source of anisotropy, a brief account of crystal plasticity is presented. Based on the physics of plastic deformation, the fundamental ingredients necessary to develop sound constitutive equations for anisotropic materials are given. The main anisotropic yield criteria proposed in the literature as well as recent developments are succinctly reviewed. For sheet metals, testing procedures needed to characterize anisotropy are described. The notions of yield and flow surfaces, i.e., the initial yielding limit and the average behavior of the material over a deformation range, respectively, are compared. Finally, examples of formability analysis and sheet forming process simulation are provided as an illustration of the concept of yield surface plasticity.

6.1 Introduction

Beyond the elastic range many materials exhibit plastic, irreversible deformations. The surface in stress space separating elastic and plastic deformations is called the yield surface. The equation of this surface

\[ \phi(\sigma_{ij}) = \phi_0, \]  

(6.1)

where \( \phi_0 \) is a certain constant, is referred to as the yield condition. \( \phi \) itself, which operates on stress tensor components, is called the yield function.

During processing, materials usually undergo very large plastic deformation while, during manufacturing, the product is usually achieved with moderate plastic strains. Even during service, the product may experience significant plastic deformation, for instance, at the tip of a crack during a fatigue phenomenon or while absorbing energy upon impacting with a foreign body. This is the reason why plasticity has been studied quite extensively, both at macroscopic or microscopic scales, over more than a century. Recently, Glazoff et al. (2003) proposed a theoretical framework to bridge the gap between physical metallurgy and plasticity of metals and alloys. The approach described in this chapter is more practical and geared to manufacturing applications, specifically to sheet forming.

The aim of Section 6.2 is to review yield surface formulations as well as relevant flow rules for isotropic and anisotropic materials in classical theory of plasticity. Section 6.3 briefly summarizes the microstructural features that contribute to plastic anisotropy and their influence on the yield surface. Section 6.4 describes a set of representative yield functions for
anisotropic plasticity. Finally, Section 6.5 illustrates the application of the yield surface concept to aluminum alloy sheet anisotropy, including mechanical testing, formability analysis and numerical forming simulation.

6.2 Classical Plasticity Theory

6.2.1 Isotropic Yield Conditions for Perfect Plasticity

The yield function \( \phi(\sigma_{ij}) \) must be invariant with respect to the choice of the reference frame. Hence, for isotropic perfectly plastic bodies, it should depend on stress invariants only. The Cauchy stress tensor, with components \( \sigma_{ij} \), has three independent invariants but various combinations of such invariants are possible. The sets of invariants of the stress tensor, which are mostly used, are compared by Zyczkowski (1981). The principal stresses \( \sigma_1, \sigma_2, \sigma_3 \) are invariants themselves and three other important systems of invariants are discussed below. The coefficients of the cubic equation determining the principal stress \( \sigma_i \) in terms of stress components

\[
P(\sigma_i) = \det(\sigma - \sigma_i I) = -\sigma_i^3 + 3\bar{I}_1 \sigma_i^2 + 3\bar{I}_2 \sigma_i + 2\bar{I}_3 = 0,
\]

(6.2)

where \( I \) is the unit tensor, are called the principal stress invariants:

\[
\bar{I}_1 = 3\bar{I}_1 = \sigma_{11} + \sigma_{22} + \sigma_{33}
\]

\[
\bar{I}_2 = 3\bar{I}_2 = \sigma_{23}^2 + \sigma_{31}^2 + \sigma_{12}^2 - \sigma_{22}\sigma_{33} - \sigma_{33}\sigma_{11} - \sigma_{11}\sigma_{22}
\]

\[
\bar{I}_3 = 2\bar{I}_3 = 2\sigma_{23}\sigma_{31}\sigma_{12} + \sigma_{11}\sigma_{22}\sigma_{33} - \sigma_{11}\sigma_{23}^2 - \sigma_{22}\sigma_{31}^2 - \sigma_{33}\sigma_{12}^2.
\]

(6.3)

The invariant \( \bar{I}_1 \) denotes the mean stress \( \sigma_m \) while the quantity \( p = -\sigma_m \) is called the hydrostatic pressure. It should also be noted that some authors introduce \( \bar{I}_2 \) with opposite sign, Zyczkowski (1981). Moreover, defining the stress deviator components \( s_{ij} \) by the formulae

\[
s_{ij} = \sigma_{ij} - \bar{I}_1 \delta_{ij},
\]

(6.4)

where \( \delta_{ij} \) denotes Kronecker’s symbol, similar principal invariants of the stress deviator \( J_1, J_2, J_3 \) can be introduced. Since \( J_1 = 0, \bar{I}_1 \) is added to this set to form a system of three independent invariants for the stress tensor. It may be shown that \( J_2 \) is always non-negative and hence the signs are defined like in Equations (6.2) and (6.3). Instead of the invariants \( J_1 \) or \( J_3 \), the “angular invariant” may be introduced by the formula

\[
\theta = \arccos \left( \frac{2\bar{I}_3^3 + 3\bar{I}_1 \bar{I}_2 + 2\bar{I}_3}{\bar{I}_1^2 + \bar{I}_2} \right)^{-3/2},
\]

(6.5)

where the ordered principal stresses (\( \sigma_1 \geq \sigma_2 \geq \sigma_3 \)) are (see Barlat et al. 1991 for \( \bar{I}_1 = 0 \))

\[
\sigma_1 = 2\sqrt{\frac{\bar{I}_2^2 + \bar{I}_3}{\bar{I}_1}} \cos \left( \frac{\theta}{3} \right) + \bar{I}_1
\]

\[
\sigma_2 = 2\sqrt{\frac{\bar{I}_1^2 + \bar{I}_2}{\bar{I}_1}} \cos \left( \frac{\theta + 4\pi}{3} \right) + \bar{I}_1
\]

\[
\sigma_3 = 2\sqrt{\frac{\bar{I}_1^2 + \bar{I}_2}{\bar{I}_1}} \cos \left( \frac{\theta + 2\pi}{3} \right) + \bar{I}_1.
\]

(6.6)
In the Haigh-Westergaard space of principal stresses, the invariant \( \varpi = \theta / 3 \) (Lode, 1926) may be interpreted as the angular coordinate in the cylindrical system with the axial direction equally inclined with respect to the axes \( \sigma_1, \sigma_2, \sigma_3 \). Since \( \bar{I}_1 \) is proportional to the axial coordinate and \( \sqrt{J_2} \) to the radial coordinate, the system of invariants \( \bar{I}_1, \sqrt{J_2}, \varpi \) may be interpreted as proportional to the cylindrical coordinates in the principal stress space.

The oldest condition for plastic yielding was proposed by Tresca (1864), expressing that the maximum shear stress reaches a constant critical value

\[
\max \left( \frac{\sigma_1 - \sigma_2}{2}, \frac{\sigma_2 - \sigma_3}{2}, \frac{\sigma_3 - \sigma_1}{2} \right) = \sigma_s,
\]

(6.7)

where \( \sigma_s \) corresponds to the yield-point stress in pure shear. This yield surface is a hexagonal prism in the principal stress space and does not depend on the mean stress. Reuss (1933) expressed the Tresca yield condition in terms of invariants

\[
\phi = 4J_3^3 - 27J_2^2 - 4\sigma_s^4 \left(3J_2 - 4\sigma_s^2\right)^2 = 0
\]

(6.8)

Tresca’s work and the theoretical studies of Levy (1870) led to the foundation of the mathematical theory of plasticity by Saint Venant. In his famous memoir, Saint Venant (1870 and 1871) first proposed a model for rigid ideal plasticity in which elastic strains are neglected and plastic flow does not depend on the mean stress. Next, he discussed the issues that have been and still are of prime importance, namely the choice of the function \( \phi \), which will determine when or if a material enters the plastic stage and what are the corresponding stress-strain relationships under further loading and unloading.

The most popular yield condition (verified for many isotropic metals) may be expressed with the second deviatoric stress invariant

\[
J_2 = \sigma_u^2 / 3
\]

(6.9)

or, in expanded form,

\[
(\sigma_{22} - \sigma_{33})^2 + (\sigma_{33} - \sigma_{11})^2 + (\sigma_{11} - \sigma_{22})^2 + 6\left(\sigma_{23}^2 + \sigma_{31}^2 + \sigma_{12}^2\right) = 2\sigma_u^2.
\]

(6.10)

Here, \( \sigma_u \) denotes the yield-point stress in uniaxial tension. This yield condition is usually ascribed to von Mises (1913), who proposed it as a convenient approximation of Equation (6.7). However, Huber proposed it as early as 1904 as a condition of constant energy of distortion, see Zyczkowski (1981), Engel (1994). The Huber-Mises surface is a cylinder in the Haigh-Westergaard space. The advantage of this criterion is the simplicity of its representation in terms of the stress components, whereas analytical forms for most other yield conditions are simple only when expressed in terms of principal stresses.

Many experimental data showed that the measured yield surface of metals is located between the Tresca and Huber-Mises predictions. Taylor and Quinney (1931) reported extensive test results in which isotropic copper and steel tubes were loaded in combined tension and torsion. These authors found that the deviations from the Huber-Mises criterion were indeed real and that the data were located between the two criteria. These findings led to the conclusion that, in order to reproduce the data, the third invariant \( J_3 \) ought to be included in the expression of the yield criterion. Drucker (1949) proposed the following yield condition

\[
J_2^3 - cJ_2^3 = \sigma_s^6
\]

(6.11)
Assuming $c = 0$, the Huber-Mises yield condition is recovered, whereas direct transition to Tresca is not possible.

Another description of cylindrical yield surfaces including Tresca and Huber-Mises as particular cases, was proposed by Hershey (1954). He applied to plasticity the Norton-Bailey law (Norton 1929, Bailey 1929), well-known in the theory of non-linear creep, to represent yield surfaces calculated with self-consistent polycrystal calculations

\[
\phi = |\sigma_1 - \sigma_2|^a + |\sigma_2 - \sigma_3|^a + |\sigma_3 - \sigma_1|^a
= |s_1 - s_2|^a + |s_2 - s_3|^a + |s_3 - s_1|^a = 2\sigma_u^a
\]  

(6.12)

For $a = 2$ or $a = 4$, Equation (6.12) reduces to the Huber-Mises yield condition, whereas for $a = 1$ and in the limiting case $a \rightarrow \infty$ it leads to the Tresca yield condition. For $2 < a < 4$, the corresponding surface lies outside the Huber-Mises circular cylinder, whereas for $1 < a < 2$ and for $a > 4$, it lies between Huber-Mises and Tresca. According to Hershey, particularly good agreement with many experimental data may be obtained by assuming $a = 6$.

All the yield conditions presented above are not affected by simultaneous change of sign of all the stress components. In other words, these criteria predict equal yield stresses in tension and compression. Such materials, for which the corresponding yield surface exhibits symmetry about the origin, were called “isosensitive” by Życzkowski (1981). Many metals are almost isosensitive but, for instance cast iron and geomaterials, which are sensitive to the mean stress, do not evidently belong to this group. Such materials were called “anisosensitive.” HCP materials, although independent of hydrostatic pressure, are anisosensitive because their deformation mechanisms, slip and twinning, are different in tension and compression. Cazacu and Barlat (2003a) recently proposed a yield criterion for these materials.

A relatively simple but sufficiently general yield condition for anisosensitive materials was proposed by Burzyński (1929)

\[
\phi = bJ_2 + cI_2^2 + dI_1 = \phi_0.
\]  

(6.13)

The relevant yield surface in the space of principal stresses is an ellipsoid (a hyperboloid would be without practical meaning). If $c = 0$, this yield surface turns into a paraboloid, while, if $\phi_0 = -d^2/4c$ and $c < 0$, it turns to a cone (after extracting the root). The conical yield surface was employed by Drucker and Prager (1952) to formulate the theory of plasticity for geomaterials. In general, the constants $b, c, d$ and $\phi_0$ may be determined using the uniaxial tension yield-point stress ($\sigma_u$), the uniaxial compression stress ($\sigma_u - \sigma_m$) and pure shear stress ($\sigma_s$)

\[
\frac{\sigma_u \sigma_u - \sigma_s^2}{3\sigma_s^2} + \left( 9 - \frac{3\sigma_u \sigma_u - \sigma_s^2}{\sigma_s^2} \right) \sigma_m^2 + 3(\sigma_u - \sigma_u) \sigma_m - \sigma_u \sigma_u - \sigma_m = 0.
\]  

(6.14)

In the above expression,

\[
\sigma_s = \sqrt{3J_2} = \sqrt{\frac{3}{2} s_{ij} s_{ij}}
\]  

(6.15)

is associated to the Huber-Mises yield condition and is called the “effective stress” or the “stress intensity.” In Equation (6.15), Einstein’s summation convention was employed. Yield conditions have a very ample literature. Życzkowski (1981) and Yu (2002) reviewed several hundred yield conditions each.
6.2 Classical Plasticity Theory

6.2.2 Flow Rules

Yield conditions must be completed by appropriate constitutive equations relating stress and strain components and their time derivatives. For rigid-plastic bodies, it is usually assumed that plastic strain rate and deviatoric stress components are proportional (Levy, 1870). In order to obtain a tensorial law, a scalar function of stresses called the plastic potential \( \psi = \psi(\sigma_{ij}) \) is usually introduced. The following equation, called the flow rule, results from von Mises maximum plastic dissipation postulate

\[
\dot{\varepsilon}^p_{ij} = \lambda \frac{\partial \psi}{\partial \sigma_{ij}}.
\] (6.16)

\( \dot{\varepsilon}^p_{ij} \) are the plastic strain rate components and \( \lambda \) is a scalar coefficient to be eliminated using the consistency condition \( \dot{\varphi} = 0 \). For many materials, in particular for most metals, the plastic potential \( \psi \) can be identified with the yield function \( \varphi \) leading to

\[
\dot{\varepsilon}^p_{ij} = \lambda \frac{\partial \varphi}{\partial \sigma_{ij}}.
\] (6.17)

This flow rule is said to be associated with the yield condition, whereas if \( \psi \neq \varphi \) it is referred to as non-associated flow rule. Relationship 6.17 is also called “the normality rule” since the plastic strain rate components are normal to the yield surface. If, moreover, the yield surface is convex, it may be shown that the excessive plastic work is positive (for plastic hardening) or non-negative (for perfect plasticity), and the material is defined as stable, or indifferently stable, respectively, Drucker (1949, 1951). In the principal reference frame, a potential \( \psi (\sigma_k) \) is convex with respect to the principal stresses \( \sigma_k \) if its Hessian matrix \( H \), i.e.,

\[
H_{ij} = \frac{\partial^2 \psi}{\partial \sigma_i \partial \sigma_j},
\] (6.18)

is positive semi-definite (Rockafellar, 1972). Lippmann (1970) showed that convexity is preserved from the principal frame to any other reference frames.

The mathematical description of a yield condition is not unique. For many applications it is convenient to present the function \( \varphi (\sigma_{ij}) \) as a homogeneous or semi-homogeneous function of the first degree. A homogeneous function of the first degree satisfies the condition \( \varphi (c\sigma_{ij}) = c\varphi (\sigma_{ij}) \) for any scalar \( c \), whereas a semi-homogeneous function obeys this property for non-negative \( c \) only.

Sometimes the above requirement can be obtained without any difficulty. For example, in the case of the Tresca yield condition, it is readily satisfied. For the Huber-Mises function, it is obtained by replacing the second invariant by the effective stress (Equation (6.15)), and for the Hershey yield condition (Equation (6.12)), by replacing \( \varphi \) by \( \varphi^{1/\alpha} \). However, for other yield conditions, like that by Burzyński (Equation (6.13)), the reduction to a homogeneous or semi-homogeneous function of the first degree is more complicated.

The required property may be achieved by introducing the concept of reduced stress, Życzkowski (1981, p. 71). The reduced stress, \( \sigma_{\text{red}} \), is the stress in uniaxial tension, which from the point of view of plasticity and failure in proportional loading, is equivalent to the given general state of stress. First, all the yield stresses have to be expressed in comparative states...
as dimensionless fractions $\kappa_m$ of the yield stress in uniaxial tension $\sigma_u$. Then, the resulting yield function must be solved with respect to $\sigma_u$ and the result denoted by $\sigma_{\text{red}}$. The equation

$$\sigma_{\text{red}}(\sigma_{ij}, \kappa_m) = \sigma_u$$  \hspace{1cm} (6.19)$$

represents the yield condition in the required form, since dimensional analysis requires $\sigma_{\text{red}}$ to be homogeneous or semi-homogeneous function of the first degree of $\sigma_{ij}$.

The Burzyński yield condition is considered as an example. Its form in Equation (6.14) is evidently non-homogeneous with respect to the stress components. Introducing two dimensionless coefficients $\kappa_c = \sigma_u - /\sigma_u$ and $\kappa_s = \sigma_s /\sigma_u$, it can be rewritten in the form

$$\frac{\kappa_c}{3\kappa_s^2} \sigma_e^2 + \left(9 - \frac{3\kappa_c}{\kappa_s^2}\right) \sigma_m^2 + 3(\kappa_c - 1) \sigma_m \sigma_u - \kappa_c \sigma_u^2 = 0.$$  \hspace{1cm} (6.20)$$

Solving this equation with respect to $\sigma_u$ and denoting the positive root by $\sigma_{\text{red}}$ results in

$$\sigma_{\text{red}} = \frac{1}{2\kappa_c} \left\{3(\kappa_c - 1) \sigma_m + \sqrt{9(\kappa_c + 1)^2 - 12 \frac{\kappa_c^2}{\kappa_s^2} \sigma_m^2 + 4 \kappa_c^2 \sigma_e^2}\right\},$$  \hspace{1cm} (6.21)$$

Indeed, the function in the right-hand side of Equation (6.21) is semi-homogeneous of the first degree and substituted into Equation (6.19) renders the Burzyński yield condition in the required form. If $\kappa_c = 1$, $\kappa_s = 1/\sqrt{3}$, as predicted by the Huber-Mises yield condition, Equation (6.21) leads to $\sigma_{\text{red}} = \sigma_e$ as expected. From the associated flow rule rewritten as

$$\dot{\varepsilon}_{ij}^p = \lambda \frac{\partial \sigma_{\text{red}}}{\partial \sigma_{ij}}.$$  \hspace{1cm} (6.22)$$

$\lambda$ can be eliminated by multiplying both sides with $\sigma_{ij}$ with contraction, by making use of the yield condition (6.19) and by applying Euler’s theorem on homogeneous functions,

$$\lambda = \frac{\sigma_{ij} \dot{\varepsilon}_{ij}^p}{\sigma_u} = \frac{\dot{W}^p}{\sigma_u}.$$  \hspace{1cm} (6.23)$$

Here $\dot{W}^p$ denotes the rate of plastic work per unit volume. Returning to Equation (6.22) and multiplying formally by the time increment $dt$ leads to

$$d\varepsilon_{ij}^p = \frac{\sigma_{kl} \partial \sigma_{\text{red}}}{\sigma_u} \dot{\varepsilon}_{ij}^p,$$  \hspace{1cm} (6.24)$$

which eliminates the time factor. Such materials are called plastic or rate-independent while they are called viscoplastic or rate-dependent if the time factor cannot be eliminated. Rate-independent materials mostly are considered in the present chapter.

### 6.2.3 Subsequent Yield Surfaces during Plastic Hardening

Perfect plasticity is the most simple approximation in the description of plastic deformation. Even if this is a good approximation in certain cases, e.g., for heavily cold rolled materials,
plastic deformation is usually accompanied with plastic hardening, in particular, for annealed metals and alloys. The basic information concerning plastic hardening is given by the tension test. Therefore most theories of plastic hardening attempt to employ that information and to "transfer" this stress-strain curve to the general three-dimensional case. The simplest approach of this type is to formulate a hypothesis of a unique curve for plastic hardening, i.e., that for a certain class of deformation processes there exists a unique dependence between certain invariants of the stress tensor and of the strain tensor, namely

$$\sigma_{\text{red}} = h(\varepsilon_{\text{red}})$$  \hspace{1cm} (6.25)

The reduced stress $\sigma_{\text{red}}$ is determined here as described in Section 6.2.2, whereas the corresponding reduced strain $\varepsilon_{\text{red}}$ may be defined in various ways. Usually, either $\varepsilon_{\text{red}}$ is defined in such a way as to have Equation (6.25) satisfied in the elastic range as well, or the reduced plastic strain is defined by the equation $W_p = \sigma_{ij} \dot{\varepsilon}_{ij} = \sigma_{\text{red}} \dot{\varepsilon}_{\text{red}}$. The first hypothesis of this type was formulated by Ludwik (1909) and was based on the Tresca yield condition while Roš and Eichinger (1926) employed the Huber-Mises yield condition. Hencky (1929) noticed that it could also be extended to finite strains if stresses and strains were respectively taken as the true stresses (related to the deformed cross-sectional area) and logarithmic strains $\epsilon^H$, defined as $\epsilon^H = \ln (1 + \epsilon)$. Experimental verification confirmed the hypothesis described by Equation (6.25) for proportionally increasing stress or strain components (simple loading processes). Nevertheless, in more general cases, this relationship serves as a basis for the description of plastic hardening.

Perfect plasticity is based on the assumption that the yield surface remains constant during the whole plastic deformation process. On the contrary, during plastic hardening, the initial yield surface changes (is subject to a transformation) and subsequent yield surfaces need to be expressed in terms of certain internal state variables, connected with plastic strains.

The classical hardening rules describe the proportional expansion and/or rigid translation of the initial yield surface (isotropic and kinematic hardening, respectively). A more accurate description is obtained if two further elements of transformation of the initial surface are introduced: Affine extension and rigid rotation. The relevant equation starting from the Huber-Mises yield condition may be written in the form proposed by Drucker (1950)

$$N_{ijkl} (s_{ij} - \alpha_{ij}) (s_{kl} - \alpha_{kl}) = 1 = 0,$$  \hspace{1cm} (6.26)

where the second-order tensor $\alpha_{ij}$ describes a rigid translation in stress space, whereas the fourth order tensor $N_{ijkl}$ is responsible for the remaining three elements of the transformation. Equation (6.26) was discussed in detail by Baltov and Sawczuk (1965). If $N_{ijkl} = \delta_{ik} \delta_{jl} / \kappa^2 (\tilde{\kappa})$, where $\kappa$ is a scalar internal state variable, then the affine extension and rotation are eliminated and Equation (6.26) reduces to mixed isotropic-kinematic hardening

$$\left( s_{ij} - \alpha_{ij} \right) \left( s_{kl} - \alpha_{kl} \right) = \kappa^2 (\tilde{\kappa}).$$  \hspace{1cm} (6.27)

Usually $\kappa (\tilde{\kappa})$ is taken from Equation (6.25), but $\varepsilon_{\text{red}}$ is replaced by an integral invariant, for instance, by Odqvist’s parameter (1933)

$$I_{\varepsilon p} = \int_0^t \sqrt{\dot{\varepsilon}_{ij} \dot{\varepsilon}_{ij}} \, dt.$$  \hspace{1cm} (6.28)
For a simple loading process this integral can effectively be evaluated, and turns into the effective plastic strain $\varepsilon_p$ corresponding to the Huber-Mises yield condition. Equation (6.28) may be regarded as the evolution equation for the internal state variable governing isotropic plastic hardening. In any case, the reduction of the initial yield condition to the homogeneous or semi-homogeneous form of the first degree is of essential importance.

The evolution equation for kinematic hardening according to Melan-Ishlinsky and Prager (see Zyczkowski 1981) reads

$$\dot{\alpha}_{ij} = c \dot{\varepsilon}_{ij}$$

or, according to Ziegler (1959),

$$\dot{\alpha}_{ij} = (\sigma_{ij} - \alpha_{ij}) \dot{v}$$

where the proportionality factor $\dot{v}$ should be found from the subsequent yield condition (Equation (6.27)). Isotropic and isosensitive materials subjected to a proportional expansion of the yield surface (i.e., isotropic hardening) remains isotropic and isosensitive. On the other hand, the other elements of transformation of the initial yield surface introduce anisotropy and anisosensitivity. Any of such types of hardening is called “anisotropic plastic hardening.” The Bauschinger effect (1881), i.e., the material softening upon reverse loading, is a form of anisosensitivity introduced by hardening.

Experiments have also shown a fifth element of transformation of the initial yield surface, namely distortion (i.e., non-affine extension). A geometric description of distortional plastic hardening was proposed by Kurtyka and Zyczkowski (1985) and by Boehler (1987). Relevant evolution equations were given in Kurtyka and Zyczkowski (1996).

The flow rules for plastic hardening materials are usually assumed in the non-associated or associated forms (Equations (6.16) and (6.17)), where now $\psi$ and $\phi$ depend on k scalar and m tensorial internal state variables, e.g., $\phi = \phi (\sigma_{ij}, \kappa_k, \alpha_{ijkl(m)})$.

### 6.2.4 Anisotropic Plasticity

Anisotropy means the dependence of material properties on the testing direction. At least three sources of anisotropy can be distinguished: Original or innate anisotropy, e.g. in single crystals; Acquired anisotropy (or deformation induced), due to prior plastic deformations; Finally structural anisotropy, e.g. in fiber-reinforced materials, if their properties are subject to averaging (homogenization).

If the yield condition for perfect plasticity of anisotropic materials is considered, then the stress invariants should be replaced by simultaneous (common) invariants of the stress tensor and of the material (or structural) tensors of plastic anisotropy, for instance

$$\phi (\Pi_{ij}, \Pi_{ijkl} \sigma_{ij} \sigma_{kl}, \Pi_{ijklm} \sigma_{ij} \sigma_{kl} \sigma_{mn}, \ldots) = \phi_0,$$

where $\Pi_{ij}$, $\Pi_{ijkl}$, $\ldots$ denote plastic anisotropy tensors. The invariants shown in the above relationship are the simplest, but by no means general. For example, two tensors of the second order have 10 irreducible polynomial invariants, and the number of irreducible invariants of $\Pi_{ij}$ and $\sigma_{ij}$ amounts to 28 according to the theory of representation of tensor functions (Wang...

\[ \tilde{\sigma}_{ij} = A_{ijkl} \sigma_{kl}, \text{ with } A_{ijkl} = A_{jikl} = A_{ijlk} = A_{klij}, \]

(6.32)

where \( A_{ijkl} \) is a certain anisotropy tensor. The substitution of the stresses by their transformed values into well-known isotropic yield conditions allows for a direct generalization of these conditions for anisotropic materials.

### 6.2.5 Direct Generalizations of Isotropic Yield Conditions

The first anisotropic yield condition was proposed by von Mises (1928)

\[ A_{ijkl} \sigma_{ij} \sigma_{kl} = 1 \text{ with } A_{ijkl} = A_{jikl} = A_{ijlk} = A_{klij}. \]

(6.33)

These symmetry conditions reduce the number of components to 21. Further, the additional requirement of independence of Equation (6.33) on the mean stress \( \sigma_m \) reduces the number of components to 15. In this case, combined with Equation (6.31), this yield condition provides a generalization of the Huber-Mises isotropic criterion (Equations (6.9) or (6.10))

\[ \tilde{s}_{ij} \tilde{s}_{ij} = \frac{2}{3} \sigma_u^2 \]

(6.34)

in which \( \tilde{s}_{ij} \) denote the modified deviatoric stresses. For an orthotropic material, Hill (1948) presented the above relationship in the following form

\[ F (\sigma_{yy} - \sigma_{zz})^2 + G (\sigma_{zz} - \sigma_{xx})^2 + H (\sigma_{xx} - \sigma_{yy})^2 + 2L \sigma_{yz}^2 + 2M \sigma_{zx}^2 + 2N \sigma_{xy}^2 = 1 \]

(6.35)

This yield condition describes anisotropic isosensitive materials.

The generalizations of the Tresca yield condition to anisotropic materials are very simple but only in a particular case: If general anisotropy is restricted to orthotropy and the principal directions of stresses and of orthotropy coincide. Hu (1958) proposed the following system of six equations

\[ \frac{\sigma_1 - \sigma_2}{\sigma_{u1+}} + \frac{\sigma_2 - \sigma_3}{\sigma_{u2-}} = 1 \]

\[ \frac{\sigma_3^2}{\sigma_{u3+}} = 1 \]

(6.36)

with the remaining equations obtained by cyclic permutation of the indices. The symbols \( \sigma_{u1+}, \sigma_{u1-}, \text{ etc.} \), denote yield-point stresses in tension and in compression in the respective principal directions. Berman and Hodge (1959) formulated another generalization

\[ \frac{\sigma_1 - \sigma_3}{b_2} = 1, \quad \frac{\sigma_2 - \sigma_3}{b_1} = 1, \quad \frac{\sigma_2 - \sigma_1}{b_3} = 1 \]

\[ \frac{\sigma_3 - \sigma_1}{c_2} = 1, \quad \frac{\sigma_3 - \sigma_2}{c_1} = 1, \quad \frac{\sigma_1 - \sigma_2}{c_3} = 1 \]

(6.37)
Both these proposals allow for anisosensitivity of the material.

Using the transformed stress tensor (Equation (6.32)), Życzkowski (2001) proposed a direct generalization of the Burzyński yield condition

\[
\frac{3}{2} b \left( A_{ijkl} - \frac{1}{3} A_{ppkl} \delta_{ij} \right) \left( A_{ijqr} - \frac{1}{3} A_{ttqr} \delta_{ij} \right) \sigma_{kl} \sigma_{qr} + \frac{1}{9} c (A_{ppkl} \sigma_{kl})^2 + \frac{1}{3} d A_{ppkl} \sigma_{kl} = \phi_0
\]  

(6.38)

This yield condition obviously describes anisotropic anisosensitive materials.

6.3 Material Structure and Plastic Anisotropy

Plastic anisotropy and anisosensitivity are due to the material microstructural state and result in the distortion of the yield surface compared to an isotropic, isosensitive shape such as Tresca, Huber-Mises or Hershey’s. At moderate temperature, plastic deformation occurs mainly by glide of dislocations and possibly, depending on the material, by twinning. For both slip and twinning, the associated shear strains operate on given crystallographic planes and in certain directions. These mechanisms lead to both anisotropy and anisosensitivity. For HCP materials, and more particularly for magnesium alloys, twinning occurs for compressive stress states, while slip is predominant for tensile states (Kelley and Hosford, 1968). This phenomenon leads to a large anisosensitivity while the strong crystallographic texture, i.e., the distribution of grain orientations usually exhibited by these materials, leads to plastic anisotropy.

Modeling of the microstructural effects will be briefly reviewed in this section but it is worth mentioning that metals and alloys are not necessarily homogeneous and can exhibit microstructure gradients. Thus, the local material behavior might control the performance of a product as a whole. For instance, resistance to fracture in bending of sheet metals exhibiting a gradient through the thickness might be controlled by the microstructure at the surface of the sheet.

6.3.1 Texture

Polycrystal description of plastic anisotropy has been very successful over the last few decades. This approach is based on the physical aspects of plastic deformation, slip and twinning in crystals, and on averaging procedures over a large number of grains. The crystallographic texture is the main input to these models but other parameters, such as the grain shape, can also be included. Because polycrystal models can track the lattice rotation of each individual grain, the material anisotropy is naturally evolutional, which makes this approach very attractive. However, as a drawback, these models are very computationally time-intensive. Many grains have to be considered and the yield surface is not described with an analytical function but with a set of discrete points. Therefore, this approach cannot be used conveniently for industrial applications, particularly for sheet forming analysis and design. The different polycrystal models that were proposed over the years differ in the single crystal constitutive model and the particular procedure used to average the mechanical behavior over a
6.3 Material Structure and Plastic Anisotropy

representative number of grains (Kocks, Tomé and Wenk 1998). Models for rate-independent and visco-plastic single crystals deforming by slip, as well as the self-consistent averaging procedure, are briefly reviewed in this section. More comprehensive descriptions of crystal plasticity are given by Dawson (Chapter 5), Van Houtte (Chapter 22), Tomé (Chapter 23), and Kalidindi (Chapter 26).

6.3.1.1 Single Crystal Behavior

The rate independent formulation, which uses a single crystal yield function $\phi$ (Gambin 1992, 2001), gives an unambiguous relationship between the shear rates and stresses on all the slip systems

$$
\phi = \varphi^{1/a} = \left\{ \sum_s \left| \frac{\tau_s}{\tau_s^c} \right|^a \right\}^{1/a} = 1.
$$

(6.39)

Here, $\tau_s = b_s n_s s_{ij}$ is the resolved shear stress on a system $s$, defined by the slip plane normal $n$ and the slip direction $b$, resulting from the deviatoric stress tensor $s$. $a$ is an exponent which regularizes the singularity of the Schmid description of the single crystal yield surface (Schmid and Boas, 1935). $\tau_s^c$ denotes the critical resolved shear stress for the system $s$, although it is different from a threshold stress. $\tau_s^c$ reduces to a threshold stress in the limit case $a \to \infty$, which corresponds to the classical Schmid crystal. The associated flow rule leads to the strain rate tensor $\dot{\epsilon}^p$ and the plastic spin tensor $\omega^p$:

$$
\dot{\epsilon}^p_{ij} = \frac{\lambda}{2} \left( \frac{\partial \phi}{\partial s_{ij}} + \frac{\partial \phi}{\partial s_{ji}} \right),
$$

(6.40)

$$
\omega^p_{ij} = \frac{\lambda}{2} \left( \frac{\partial \phi}{\partial s_{ij}} - \frac{\partial \phi}{\partial s_{ji}} \right),
$$

where $\lambda$ is a proportionality factor. Practically, when the exponent $a$ is large enough (for instance, 30), the single crystal behavior is very similar to that described by the Schmid law.

In order to relate the shear strain rate $\dot{\gamma}^s$ to the stress state in an unambiguous manner, another approach consists in using a visco-plastic single crystal model (Peirce et al., 1983). The following visco-plastic rule is used with a strain rate sensitivity coefficient $\eta$, typically set to 20 for deformation at room temperature

$$
\dot{\gamma}^s = \dot{\gamma}_0 \left( \frac{\tau^s}{\tau^s_c} \right)^\eta.
$$

(6.41)

$\dot{\gamma}_0$ is a reference shear strain rate and $\tau^s_c$ reduces to the classical critical resolved shear stress when the material becomes rate independent ($\eta \to \infty$).

6.3.1.2 Polycrystal Averaging Procedure

For either rate dependent or rate-independent models, the behavior of a polycrystal can be obtained from the behavior of each single grain using an averaging procedure. In general, a grain is subjected to the constraints of its neighbors in a polycrystal. However, because this grain interaction would require the knowledge of the spatial distribution of all the grain

---

1 Note: Derivatives obtained without assuming the symmetry $s_{ij} = s_{ji}$.
orientations and shapes, the behavior of a grain is usually based on the average constraint of the polycrystal. Hill (1965a) introduced a constraint modulus tensor $\Lambda$ that relates the stress and strain rates in each grain, $\dot{\sigma}$ and $\dot{\varepsilon}$, to their polycrystal average, $\bar{\dot{\sigma}}$ and $\bar{\dot{\varepsilon}}$, respectively:

$$\dot{\sigma} - \bar{\dot{\sigma}} = \Lambda \left( \dot{\varepsilon} - \bar{\dot{\varepsilon}} \right).$$

When a grain is assumed to be free of any constraint ($\Lambda_{ijkl} = 0$), the stress rate is identical in each grain. By contrast, when the grain is fully constrained ($\Lambda_{ijkl} = \infty$, Taylor, 1938), the strain rate in each grain is identical. Many intermediate averaging procedures, such as the visco-plastic self-consistent approach (for instance, Lebensohn and Tomé, 1993), have been proposed in the literature to derive the behavior of a polycrystal from the single crystal properties. Although the most critical feature leading to anisotropy in metal and alloys is texture, other features contribute as well. The shape of the grains affects their mode of interaction, thus promoting activities in certain crystallographic slip planes and directions. This effect is captured by the self-consistent averaging procedure.

### 6.3.1.3 Applications

Bishop and Hill (1951) showed that for a single crystal obeying the Schmid law, i.e., dislocation glide occurs when the resolved shear stress on a slip system reaches a critical value, the resulting yield surface must be convex and the corresponding strain rate increment vector must be normal to it. Furthermore, they extended this result for a polycrystal by averaging the behavior of a representative number of grains in an elementary volume without making any assumption about the interaction mode between grains or the uniformity of the deformation gradient. Thus, the use of a convex yield surface as a stress potential (i.e., associated flow rule) for the description of plastic anisotropy is justified on physical grounds. It is worth noting, though, that non-associated effects in crystals were reported (e.g., Ni$_3$Al) and that appropriate polycrystal formulations were developed for these materials (Qin and Bassani, 1992a; 1992b).

Crystal plasticity can be used to compute yield surfaces, which represent the anisotropy of the flow stress in a multiaxial stress space. Figure 6.1a represents the crystallite orientation distribution function of a typical rolling texture for an aluminum alloy sheet. Figure 6.1b shows the corresponding plane stress yield surface computed with the full constraint and the visco-plastic self-consistent approaches. This yield surface features anisotropy, particularly for plane strain deformation, and exhibits stress state dependent curvatures, departing from a smooth Huber-Mises ellipse.

### 6.3.2 Dislocations

After any given amount of glide, dislocations are trapped into the microstructure leading to strain hardening. Depending on the distribution of these immobile dislocations, anisotropy effects can be observed. Hardening describes the evolution of the critical shear stresses $\tau_c^\varphi$ with plastic deformation. Empirical data and theoretical developments obtained on a large number of metals and alloys (Kocks, 1966, Gil Sevillano et al. 1981) suggest that the shear flow strength on a slip system, $\tau_c^\varphi$, is a linear function of the square root of the forest dislocation density, $\rho_f$. In the Kocks (1976)-Mecking (1977) model the forest dislocation density ($\rho_f$)
6.3 Material Structure and Plastic Anisotropy

Figure 6.1: a) Rolling texture for 5019A-H48 aluminum alloy sheet and; b) Yield loci calculated with full constraint (FC) and visco-plastic self-consistent (VPSC) polycrystal models.

is taken as a state variable whose evolution is controlled by a dislocation storage term and a recovery term. The resulting microscopic shear stress-shear strain flow curve leads to the Voce law, i.e., with a saturation stress. Kubin and Estrin (1990) generalized the Kocks-Mecking model by using both the forest and mobile dislocation densities. Their work can be used to estimate the strain rate sensitivity parameter in alloys \( m = \frac{\partial \ln (\sigma)}{\partial \ln (\dot{\varepsilon})} \), which is often used to characterize viscous effects when materials are slightly rate-dependent. Many other approaches were published over the last decades (e.g., Estrin, 1996; Kocks, 2001), but a review of this topic is out of the scope of this chapter.

Figure 6.2: Dislocation microstructure in simple shear for 1050-O and 6022-T4.

In a grain, the critical resolved shear stresses \( \tau^* \) is often assumed to be the same for all the systems \( \tau^*_c = \tau_c \) and to increase isotropically as a function of the accumulated shear strain...
Γ(Γ = \sum |\dot{\gamma}|) Tomé et al. (1984) proposed an extension of the Voce law that captures the crystal behavior for large strains

\[ \tau_c = \tau_c(\Gamma) = \tau_0 + (\tau_1 + \theta_1 \Gamma) \left[ 1 - \exp \left( -\frac{\theta_0 \Gamma}{\tau_1} \right) \right]. \]  

(6.43)

In this equation, \( \tau_0, \tau_1, \theta_0 \) and \( \theta_1 \) (\( \theta_1 = 0 \) for Voce) are single crystal coefficients. When hardening increases isotropically on the slip systems, the anisotropic hardening behavior of the polycrystal as a whole is due to crystal lattice rotations. However, anisotropic hardening is mostly the results of a different evolution of \( \tau_s^c \) on the different slip systems due to material heterogeneities. The dislocation density is usually not uniformly distributed. Dislocation walls and cells, such as those depicted in Figure 6.2 for a 1050-O aluminum sheet (commercial purity) and a 6022-T4 aluminum alloy sheet (Al-Mg-Si) deformed in simple shear, tend to form during plastic deformation. This figure shows also that the dislocation pattern depends on the chemistry of the alloy. If the material is unloaded and reloaded in a different loading path, anisotropic strain hardening effects appear. For instance, Figure 6.3 shows the stress-strain curve of the 1050-O and 6022-T4 sheets measured in forward and reverse simple shear loading. After reloading, the hardening curves exhibit non-isotropic transient effects that appear to be dependent on the chemistry of the alloy (Rauch et al., 2002).

Different approaches have been suggested to describe this phenomenon (Teodosiu and Hu, 1998; Peeters et al. 2001a&b). Teodosiu and co-workers proposed a model where the yield surface keeps a constant shape and moves in stress space due to a back stress \( \alpha \). Furthermore, the resistance to slip due to dislocation obstacles varies as a function of the orientation of the
dislocation walls with respect to the loading direction
\[ \phi \left( \sigma_{ij} - \alpha_{ij} \right) - \sigma_u - \sigma_R - \sigma_S = 0 \] (6.44)

In this model, \( \sigma_u \) is the tensile yield stress. \( \sigma_R \) and \( \sigma_S \) are, respectively, functions of the randomly distributed dislocations and of the organized dislocation structures, both contributing to isotropic strain hardening. The anisotropy effects are obtained from the evolution of the back stress, which is a function of a 4th order tensor representing the anisotropic strength of the dislocation structure and a 2nd order tensor corresponding to the dislocation structure polarity, i.e., the unbalance of dislocation of different signs on each side of a cell wall. This model has been used successfully to describe the behavior of steel and aluminum alloys (Bouvier et al., 2002).

Dislocations models were also used to explain the pressure effect on plastic deformation. Spitzig and Richmond (1984) showed that high strength steels are anisosensitive because the uniaxial yield stress was larger in compression than in tension. They attributed this strength-differential (SD) effect to the sensitivity of steel to pressure. In fact, these authors conducted experiments under hydrostatic confinement and found a linear dependence between the mean stress (\( \bar{I}_1 \)) and the effective stress (associated to the square root of \( J_2 \)). To a reasonable approximation, Richmond and Spitzig (1980) used the following yield function to describe the experiments
\[ \phi = J_2^{1/2} + \sigma_u \left( \alpha I_1 - 1 \right) = 0, \] (6.45)

where \( \alpha \) is the pressure coefficient. This expression is similar to the yield condition proposed by Drucker and Prager (1952) for soils. Spitzig and Richmond conducted experiments for different steels and obtained approximately the same pressure coefficient, \( \alpha = 20 \text{ TPa}^{-1} \). They also performed experiments on commercial purity aluminum and obtained \( \alpha = 50 \text{ TPa}^{-1} \). The volume changes that they observed experimentally were negligible compared to those calculated by assuming the classical flow rule. They concluded that plastic flow was non-associated, i.e., the strain increment was not normal to the yield surface (Spitzig 1979, Spitzig, Sober and Richmond 1976, Spitzig and Richmond 1984).

Jung (1981) proposed a model based on the additional work needed to induce the motion of a dislocation due to the pressure dependence of the shear modulus. He developed a model that lead to the following relation for \( \alpha \) in Equation (6.45)
\[ \alpha \approx \frac{2}{3} \frac{d\mu}{d\rho} \] (6.46)

In the previous relationship, \( \mu_0 \) is the shear modulus at atmospheric pressure. This model leads to a pressure coefficient \( \alpha \) equal to about 17 and 59 TPa\(^{-1}\) for steel and aluminum, respectively. These values are in good agreement with the experimental values mentioned above. Bulatov et al. (1999) used molecular static calculations with an embedded atom potential to simulate the effect of pressure on dislocation motion. They found that this phenomenon was the result of the interaction of the transient activation dilatancy of moving dislocation with pressure. For aluminum, they computed values for the pressure coefficient \( \alpha \) of 48 and 63 TPa\(^{-1}\) for screw and mixed dislocations, respectively, which is in agreement with the experimental value as well.
Practically, for low to medium strength materials and low confinement pressure, this departure from normality can be neglected.

6.3.3 Porosity and Second Phases

In contrast with the previous approach, pressure has a large influence on volume changes for porous material. Gurson (1977) performed finite element calculations on a periodic array of voids subjected to different loading conditions. He proposed a yield condition, which was able to capture the different results of his computations.

\[
\phi(\sigma_{ij}, f, \sigma_u) = \sigma_2^2 + 2f \cosh \left( \frac{I_1}{2\sigma_u} \right) - (1 + f^2) = 0
\]  (6.47)

where \( f \) is the void volume fraction. Later, Tvergaard (1981) modified this expression in order to give more flexibility to this criterion and describe a wider range of experimental data. Because this criterion includes the \( I_1 \) invariant, it is anisosensitive. Plastic anisotropy can be captured by the formulation if the porosity evolution is a function of specific material directions. To account for porosity, another approach consists in using a damage tensor \( D \), (see Lemaitre and Chaboche 1990) generalizing to plasticity ideas of Kachanov (1958) and Rabotnov (1968).

\[
\tilde{\sigma}_{ij} = \left[ \left( \delta_{ik} \delta_{jl} + \delta_{il} \delta_{jk} \right) / 2 - D_{ijkl} \right]^{-1} \sigma_{kl}.
\]  (6.48)

The effective stress tensor \( \tilde{\sigma} \) is then used in classical continuum mechanics formulations to describe plasticity

\[
\phi(\tilde{\sigma}_{ij}, \sigma_u) = 0.
\]  (6.49)

This formulation is not necessarily anisosensitive or anisotropic, but can be made so if the damage tensor evolution depends both on pressure and on material directions (Chow and Wei 2001, Skrzypek and Ganczarski 2002).

In multi-phase materials, the second phases, whose purpose is to increase strength, also contribute to plastic anisotropy and anisosensitivity. In heat-treatable aluminum alloys, precipitates are intimately linked to texture because they exhibit specific shapes and crystallographic relationships with the grains. They can influence anisotropy in a way that depends on their mode of interaction with dislocations. Wilson (1965) showed that binary Al-4%Cu alloys aged with different thermal treatments, i.e., containing different types of precipitates, exhibit Bauschinger effects of different magnitudes. As noted by Bate et al. (1981), a strong back stress builds up as deformation proceeds in alloys containing non-shearable precipitates. These authors used the results of the elastic inclusion model due to Eshelby (1957) to estimate the values of the back stress. This approach is based on the idea that dislocations accumulate around these particles and produce elastic/plastic strain incompatibilities at the precipitate interface. Barlat and Liu (1998) used this concept to develop a non-linear isotropic/kinematic hardening model that was able to explain the influence of non-shearable precipitates on plastic anisotropy for binary Al-Cu alloys deformed in tension or compression. It is worth noting that the corresponding constitutive equation was similar in form to Equation (6.44). However, in
this case, the back stress was a function of the volume fraction, shape and habit planes of the precipitates, and of the crystallographic texture.

When particles are coarser, such as in dual phase or composite materials, they do not interact with dislocations but produce a back stress that results mainly from the difference in elastic properties of the two phases. Although this discussion is beyond the scope of this chapter, it is worthy to note that, as for polycrystals, Hill (1965b) proposed an interaction relationship between the phases, similar to Equation (6.42), that captures the Bauschinger effect.

6.4 Yield Functions for Metals and Alloys

Phenomenological descriptions of plastic anisotropy are convenient and time-efficient for sheet forming process simulations. Because mechanical data are used as input, these models can be more accurate than polycrystal models when the strains are moderate, which is generally the case for sheet forming. Banabic (2000) provided a review of phenomenological criteria suitable for metals and alloys. For larger strains and for abrupt strain path changes, evolution is an issue and is the subject of much research at the present time (Kurtyka and Zyczkowski, 1996; Teodosiu and Hu, 1998). Nevertheless, when dealing with material containing porosity or second phases, or when assuming isotropic or kinematic hardening, the issue of defining the yield function (stress potential) for the matrix is of prime importance. As mentioned above, because crystallographic texture is the main contributor to anisotropy in metals and alloys, phenomenological yield functions should lead to behaviors consistent with crystal plasticity. In particular, under moderate hydrostatic pressure, the associated flow rule should be used as a reasonable approximation of the material behavior.

In developing yield functions/plastic potentials, it is important that the formulations reduce to isotropy when the material properties used as input are isotropic. As mentioned in Section 6.2, an effective way to achieve this condition is to write isotropic criteria in terms of invariants and to perform mathematical transformations on these invariants. Several sets of invariants have been used to describe isotropic properties, for instance, those given in Section 6.2.1. Some of the criteria reviewed in this section were not written in terms of invariants and can produce spurious anisotropy effects, as indicated by Hosford (1985).

This section describes a selection of anisotropic yield functions, which are relevant for forming processes. A few formulations published recently are reviewed in somewhat more details (BBC2000, Yld2000-2d, CB2001). Finally, the anisotropic strain-rate potential approach, in which the plastic potentials are expressed in strain rate space instead of stress space, is briefly introduced.

6.4.1 Quadratic Yield Functions

Unlike isotropic theories for which the choice of a reference frame is arbitrary, anisotropic formulations must be expressed in a system of axes attached to the material. Due to the symmetry of their thermo-mechanical processing history, sheet metals usually exhibit orthotropic symmetry, i.e., they possess three mutually orthogonal planes of symmetry. In this paper the directions associated to these planes are denoted by $\mathbf{x}$, $\mathbf{y}$ and $\mathbf{z}$, corresponding to the rolling,
transverse and normal directions of the sheet, respectively. Hill (1948) proposed an extension to orthotropy of the isotropic Huber-Mises criterion (see Equation (6.35)). The validity of this yield function has been explored in numerous experiments, the consensus being that it is well suited to specific metals and textures, especially steel (Mellor 1981, Hill 1990). In the case of a sheet, where the only non-zero stress components are the in-plane stresses \((σ_{xx}, σ_{yy}, σ_{xy})\), the criterion may be written as

\[
σ_{xx}^2 - \frac{2r_0}{1 + r_0}σ_{xx}σ_{yy} + \frac{r_0 + r_90}{r_90(1 + r_0)}σ_{yy}^2 + \frac{r_0 + r_90}{r_90(1 + r_0)}(2r_{45} + 1)σ_{xy}^2 = σ_0^2
\] (6.50)

or, alternatively

\[
σ_{xx}^2 + \left(\frac{σ_0^2}{σ_b^2} - 1\right)σ_{xx}σ_{yy} + \frac{σ_0^2}{σ_{45}^2}σ_{yy}^2 + \left(\frac{4σ_0^2}{σ_{45}^2} - \frac{σ_0^2}{σ_b^2}\right)σ_{xy}^2 = σ_0^2
\] (6.51)

\(σ_ξ\) and \(r_ξ\) (\(ξ=0, 45\) and \(90\)) correspond to the yield stress and \(r\) values (i.e., the width to thickness strain ratio, also called Lankford coefficient) in uniaxial tension along an axis at orientation \(ξ\) to the rolling direction \(x\), respectively. \(σ_b\) is the balanced biaxial yield stress. For the same material, these two formulations do not generally lead to the same yield surface. In many applications, the anisotropy in the \(r\) ratio can be disregarded by considering a state of planar isotropy (same properties in any direction of the plane of the sheet) with a uniform mean \(r\) value. The expression of the criterion is

\[
σ_1^2 - \frac{2r}{1 + r}σ_1σ_2 + σ_2^2 = σ_u^2
\] (6.52)

and represents an ellipse whose major and minor axes coincide with those of the Huber-Mises ellipse. For \(r > 1\), the criterion is generally adequate, the effect of anisotropy being to elongate the ellipse along the major axis, and slightly contract it along the minor axis for a given planar yield stress. According to Equation (6.52), the balanced biaxial stress \(σ_b\) is of magnitude \(σ_u\sqrt{(1 + r)}/2\). Woodthorpe and Peace (1969) have found that the yield stress in balanced biaxial tension for aluminum alloy sheet having a \(r\) value lying between 0.5 and 0.6 is significantly higher than the uniaxial yield stress in the plane of the sheet. However, Hill’s criterion cannot describe this behavior, i.e., materials with \(r < 1\) and \(σ_b > σ_u\). To capture this so-called “anomalous” behavior, non-quadratic yield formulations were considered.

### 6.4.2 Non-Quadratic Yield Functions

Hill (1979) proposed a non-quadratic yield criterion and derived four special cases from the general form. Lian, Zhou and Baudelet (1989) proved that these four forms could be expressed as functions of only two coefficients. The general expression of Hill’s (1979) yield criterion accounts for planar anisotropy, provided that the directions of the principal stresses are superimposed with the anisotropy axes. The most widely used expression of this yield criterion is the so-called “Special Case IV,” which applies to materials exhibiting planar isotropy

\[
|σ_1 + σ_2|^{α} + (1 + 2r)|σ_1 - σ_2|^{α} = 2(1 + r)σ_u^{α}
\] (6.53)
and reduces to Equation (6.52) for \( a = 2 \). Although the “anomalous” behavior is captured with this function, the predicted yield surfaces are sometimes different from those either determined experimentally or predicted with polycrystal models. A generalization of this criterion was proposed by Montheillet et al. (1991).

Bassani (1977) developed a family of yield functions, which depend on four parameters, in order to approximate a relatively broad range of transversely (planar) isotropic yield surfaces obtained with the Bishop-Hill (1951) polycrystal model

\[
\phi = \frac{\sigma_1 + \sigma_2}{2\sigma_b} + \frac{\sigma_1 - \sigma_2}{2\sigma_s} = 1.
\]  

(6.54)

In the above equation, \( \sigma_s \) is the yield stress in pure shear (\( \sigma_2 = -\sigma_1 = -\sigma_s \)), while \( a \) and \( a' \) are two coefficients greater than one. Hill’s (1979) criterion, Special Case IV, is recovered for \( a = a' \).

Independently from Hill, Hosford (1972) used Hershey’s isotropic criterion (Equation (6.12)) to describe crystal plasticity results and proposed the following generalization (Hosford 1979)

\[
\phi = F|\sigma_{yy} - \sigma_{zz}|^a + G|\sigma_{zz} - \sigma_{xx}|^a + H|\sigma_{xx} - \sigma_{yy}|^a = \sigma_{red}^a,
\]  

(6.55)

where \( F, G, H \) are material constants. This criterion is a particular expression of Hill’s 1979 yield criterion. Its main advantage is that it leads to a good approximation of yield loci computed using the polycrystal Bishop-Hill model (1951) by setting \( a = 6 \) for BCC materials and \( a = 8 \) for FCC materials (Logan and Hosford 1980). An important drawback of this as well as of Hill’s (1979) criteria is that they do not involve shear stresses. Thus, these criteria cannot account for the continuous variation of plastic properties between the sheet’s \( x \) and \( y \) axes. To overcome this limitation, Hosford (1985) attempted to include a shear stress component into Equation (6.55) but found some inconsistencies in the formulation. Barlat and Lian (1989) successfully extended Hosford’s 1979 criterion to capture the influence of the shear stress

\[
\phi = b|k_1 + k_2|^a + b|k_1 - k_2|^a + c|2k_2|^a = 2\sigma_{red}^a,
\]  

(6.56)

where

\[
k_1 = \frac{\sigma_{xx} + d\sigma_{yy}}{2}; \quad k_2 = \sqrt{\left( \frac{\sigma_{xx} - d\sigma_{yy}}{2} \right)^2 + q^2\sigma_{xy}^2}.
\]  

(6.57)

and \( b, c, d \) and \( q \) are material coefficients. An alternate way to include a shear stress component in the expression of anisotropic yield functions was proposed by Hill (1990)

\[
\phi = |\sigma_{xx} + \sigma_{yy}|^a + (\sigma_b / \sigma_s)^a \left( (\sigma_{xx} - \sigma_{yy})^2 + 4\sigma_{xy}^2 \right)^{a/2} / 2
\]

\[
+ \left\{ b(\sigma_{xx} - \sigma_{yy})^2 - 2c \left( \sigma_{xx}^2 - \sigma_{yy}^2 \right) \right\} \left( \sigma_{xx}^2 + \sigma_{yy}^2 + 2\sigma_{xy}^2 \right)^{(a/2)-1} = (2\sigma_b)^a
\]  

(6.58)

where \( \sigma_s \) is the yield stress in pure shear (\( \sigma_{yy} = -\sigma_{xx} \)) and \( b, c \) are material constants. This function was slightly modified by Lin and Ding (1996).
Barlat et al. (1991) proposed a yield criterion for a full stress state, denoted Yld91, which extends the isotropic Hershey-Hosford criterion to orthotropy. Anisotropy is introduced by replacing the principal values of the stress tensor by those of a stress tensor modified with weighting coefficients. This weighting procedure is equivalent with the application of a fourth order linear transformation operator on the stress tensor (see Section 6.2.4)

\[ \tilde{\sigma} = \mathbf{L} \sigma. \]

Note that the idea of introducing a fourth order linear tensor to describe anisotropy can be found in the early work of Sobodka (1969). A generalization of Hill’s 1979 criterion using a similar methodology was used by Lian and Chen (1991). Karafillis and Boyce (1993) proposed a more general yield criterion

\[ \phi(\tilde{\sigma}) = (1 - c) \phi_1(\tilde{\sigma}) + c \phi_2(\tilde{\sigma}) = 2\sigma_{\text{red}}^a, \]

where \( \phi_1 \) is the Hershey yield function (Equation (6.12)), \( c \) is a material coefficient and

\[ \phi_2(\tilde{\sigma}) = \frac{3^a}{2^{a-1} + 1} (|\tilde{\sigma}_1|^a + |\tilde{\sigma}_2|^a + |\tilde{\sigma}_3|^a). \]

Anisotropy is again introduced using the linear operator \( \mathbf{L} \). This tensor satisfies: (i) the symmetry conditions \( L_{ijkl} = L_{jikl} = L_{klij} \) \( (i,j,k,l = 1,2,3) \), (ii) the requirement of invariance with respect to the symmetry group of the material, and (iii) the three conditions \( L_{1k} + L_{2k} + L_{3k} = 0 \) (for \( k = 1,2,3 \), using \( L_{ik} = L_{ik} \) no summation), which ensures that \( \tilde{\sigma} \) is traceless and, as a result, yielding is independent of the hydrostatic pressure. Consequently, the number of coefficients that specify the anisotropy is the same as in an anisotropic \( J_2 \) type theory such as Hill’s (1948). However, none of the yield functions presented so far can account for the experimentally observed behavior of certain aluminum alloys and brass (see for example, Stout and Hecker, 1983; Muller, 1996) for which \( \sigma_0 = \sigma_{90} \) but \( r_0 \neq r_{90} \) and reciprocally. To account for this so-called “anomalous behavior of second order” Hill (1993) proposed a criterion expressed by a third order polynomial in stresses

\[ \phi = \frac{\sigma_1^2}{\sigma_0^2} - b \frac{\sigma_1 \sigma_2}{\sigma_0 \sigma_{90}} \sigma_1 \sigma_2 + \left\{ (c + d) - \left( \frac{c \sigma_1 + d \sigma_2}{\sigma_b} \right) \right\} \frac{\sigma_1 \sigma_2}{\sigma_0 \sigma_{90}} = \sigma_{\text{red}}^2, \]

where \( b, c, d \) are material parameters.

Barlat and his collaborators (1997a,b) proposed two modified expression of their Yld91 criterion, denoted Yld94 and Yld96 tailored particularly to aluminum alloys. To account for anisotropy, the authors introduced, besides the linear transformation \( \mathbf{L} \) on the Cauchy stress \( \sigma \), functions \( \varsigma_i \) of the angles between the material symmetry axes and the direction associated to the maximum principal values of the linearly transformed stress tensor. The yield function is defined as

\[ \phi = \varsigma_1 |\tilde{\sigma}_2 - \tilde{\sigma}_3|^{\alpha} + \varsigma_2 |\tilde{\sigma}_3 - \tilde{\sigma}_1|^{\alpha} + \varsigma_3 |\tilde{\sigma}_1 - \tilde{\sigma}_2|^{\alpha} = 2\sigma_{\text{red}}^a. \]

Recommended values for the exponent for this criterion are \( \alpha = 8 \) for FCC materials and \( \alpha = 6 \) for BCC materials, respectively, as recommended by Hosford (1979) and Logan and Hosford (1980).

Different non-quadratic yield criteria were developed by Gotoh (1977), Zhou (1990) and others. Additional details can be found in the recent review by Banabic (2000) and Barlat et al. (2002).
6.4 Yield Functions for Metals and Alloys

6.4.3 Yield Functions in Polar Coordinates

For planar isotropy, any two-dimensional yield criterion can be expressed as a function of the stress variables \((\sigma_2 + \sigma_1)\) and \((\sigma_2 - \sigma_1)\). Based on this idea, Budiansky (1984) proposed a general yield criterion written in terms of polar coordinates

\[
x \equiv \frac{\sigma_2 + \sigma_1}{2\sigma_b} = g(\zeta) \cos \zeta
\]

\[
y \equiv \frac{\sigma_2 - \sigma_1}{2\tau_s} = g(\zeta) \sin \zeta,
\]

where \(g(\zeta) \geq 0\) is the radial coordinate of a point located on the yield surface and \(\zeta\) is the associated polar angle. The problem that arises is to establish the function \(g(\zeta)\).

Tourki et al. (1994) proposed a yield criterion using a variable substitution similar to that introduced by Budiansky

\[
x_1 = \frac{\sigma_1 + \sigma_2}{2\sigma_b}, \quad x_2 = \frac{\sigma_1 - \sigma_2}{2\tau_s}.
\]

(6.65)

The plane stress yield criterion is defined as

\[
x_1 = x_1(\zeta, \chi) = g(\zeta, \chi) \cos \zeta
\]

\[
x_2 = x_2(\zeta, \chi) = g(\zeta, \chi) \sin \zeta.
\]

(6.66)

Here, \(\chi\) is the angle characterizing the principal stress directions with respect to the orthotropy axes, \(g(\zeta, \chi)\) is the radial coordinate of a point located on the yield surface and \(\zeta\) is the associated polar angle. This formulation is more general than that proposed by Budiansky (1984) because it allows for planar anisotropy cases due to the dependence of \(g\) on the two parameters \(\zeta\) and \(\chi\). In Tourki et al. (1994), the isotropic Drucker (1949) criterion in Equation (6.11) was first expressed in polar coordinates by deriving the corresponding function \(g(\zeta)\), then extended to anisotropy by generalizing \(g(\zeta)\) to \(g(\zeta, \chi)\).

6.4.4 Other Anisotropic Yield Functions

Other criteria have been proposed on the basis of different principles. Vegter et al. (1995) proposed to use a Bézier interpolation of experimental data points to obtain the first quadrant yield locus (both normal stresses positive). In order to describe planar anisotropy, Vegter’s criterion contains 17 parameters. The most important advantage of the criterion is the flexibility ensured by the large number of parameters. However, this approach requires a large number of experimental data and it is restricted for the description of particular stress states.

6.4.5 BBC2000 Yield Criterion

Recently, Banabic et al. (2000) proposed the following expression for a plane stress yield function

\[
\phi = \left[ d(\beta \Omega + c\Psi)^{2a} + d(\beta \Omega - c\Psi)^{2a} + (1 - d) (2c\Psi)^{2a} \right]^{\frac{1}{2a}},
\]

(6.67)
where $b, c, d$ are material parameters. $\Omega$ and $\Psi$ are functions of the second and third invariants of a transformed stress tensor $\tilde{\sigma} = \mathbf{L}\sigma$, in which $\mathbf{L}$ satisfies the conditions stated above in Section 6.4.2 (Karafillis and Boyce, 1993). Hence, in the reference system associated with the directions of orthotropy, the tensor $\mathbf{L}$ has 6 non-zero components for general stress states and 4 components for plane stress conditions, respectively. For the latter case, the relevant transformed stresses expressed in the reference system $(x, y, z)$ are

\[
\begin{align*}
\tilde{\sigma}_{xx} &= L_{11}\sigma_{xx} + L_{12}\sigma_{yy} \\
\tilde{\sigma}_{yy} &= L_{21}\sigma_{xx} + L_{22}\sigma_{yy} \\
\tilde{\sigma}_{zz} &= -\left( L_{11} + L_{21} \right)\sigma_{xx} - \left( L_{12} + L_{22} \right)\sigma_{yy} \\
\tilde{\sigma}_{xy} &= L_{66}\sigma_{xy}, \\
\tilde{\sigma}_{xz} &= \tilde{\sigma}_{yz} = 0
\end{align*}
\]  

(6.68)

where $L_{11}, L_{22}, L_{12}, L_{66}$ are the four independent components of the tensor $\mathbf{L}$. The expressions of $\Omega$ and $\Psi$ in terms of the stress components are

\[
\begin{align*}
\Omega &= (L_{11} + L_{21})\sigma_{xx} + (L_{12} + L_{22})\sigma_{yy} \\
\Psi &= \sqrt{\left[ \frac{1}{2} (L_{11} - L_{21})\sigma_{xx} + \frac{1}{2} (L_{12} - L_{22})\sigma_{yy} \right]^2 + L_{66}^2\sigma_{xy}^2}
\end{align*}
\]  

(6.69)

(for more details see Banabic et al., 2000). This formulation reduces to the potential proposed by Barlat and Lian (1989) for plane-stress conditions. Two additional parameters, namely $b$ and $c$, were introduced in order to obtain a better representation of the plastic behavior of the sheet metal.

The shape of the yield surface is defined by the material parameters $b, c, d, L_{ij}$ and $a$. Among these parameters, $a$ has a distinct status. More precisely, its value is set in accordance with the crystallographic structure of the material, i.e., $a = 3$ for BCC alloys, and $a = 4$ for FCC alloys. The other seven parameters are determined such that the model reproduces the experimental characteristics of the orthotropic sheet metal as closely as possible, namely, $\sigma_0, \sigma_{45}, \sigma_{90}, \sigma_b, \tau_0, \tau_{45}, \tau_{90}$. It is possible to obtain the values for these parameters by solving a set of seven non-linear equations. However, a more effective strategy of identification is to impose the minimization of the following error function

\[
\Xi(b, c, d, L_{ij}) = \left( \frac{\tau^p_0}{\tau_0} - 1 \right)^2 + \left( \frac{\tau^p_{45}}{\tau_{45}} - 1 \right)^2 + \left( \frac{\tau^p_{90}}{\tau_{90}} - 1 \right)^2 + \left( \frac{\sigma^p_0}{\sigma_0} - 1 \right)^2 + \left( \frac{\sigma^p_{45}}{\sigma_{45}} - 1 \right)^2 + \left( \frac{\sigma^p_{90}}{\sigma_{90}} - 1 \right)^2 + \left( \frac{\sigma^p_b}{\sigma_b} - 1 \right)^2
\]  

(6.70)

where the superscript $(.)^p$ denotes the values predicted by the constitutive equation. For the numerical minimization, the downhill simplex method proposed by Nelder and Mead (see Press et al. 1992) can be adopted because it does not need the evaluation of the gradient.

An improvement of this criterion was proposed recently by Parauan et al. (2003) in order to account for an additional mechanical parameter, namely, the biaxial anisotropy coefficient, introduced independently by Barlat et al. (2003) and Pöhlandt et al. (2002). Aretz (2003) introduced a more flexible yield function in the BBC formulation.
6.4 Yield Functions for Metals and Alloys

6.4.6 Yld2000-2d Yield Criterion

In general, a yield function written in terms of the deviatoric stress tensor fulfills the pressure independence condition. Therefore, another possible linear transformation on stresses is

\[ \tilde{s} = Cs, \]

where \( s \) is the deviatoric stress tensor and \( \tilde{s} \) the linearly transformed deviatoric stress tensor. Moreover, additional coefficients in the context of linear transformations can be obtained by using two transformations in association with two different isotropic yield functions

\[
\begin{align*}
\tilde{s}^I &= C^I s = C^I T \sigma = L^I \sigma, \\
\tilde{s}^{II} &= C^{II} s = C^{II} T \sigma = L^{II} \sigma.
\end{align*}
\]

(6.72)

\( C^I \) and \( C^{II} \) (or \( L^I \) and \( L^{II} \)) represent the linear transformations and \( T \) a matrix relating the deviatoric to the Cauchy stresses

\[
\begin{bmatrix}
\tilde{s}^I_{xx} \\
\tilde{s}^I_{yy} \\
\tilde{s}^I_{xy}
\end{bmatrix} =
\begin{bmatrix}
C^I_{11} & C^I_{12} & 0 \\
C^I_{21} & C^I_{22} & 0 \\
0 & 0 & C^I_{66}
\end{bmatrix}
\begin{bmatrix}
s_{xx} \\
s_{yy} \\
s_{xy}
\end{bmatrix}, \quad T =
\begin{bmatrix}
2/3 & -1/3 & 0 \\
-1/3 & 2/3 & 0 \\
0 & 0 & 1
\end{bmatrix}.
\]

(6.73)

A similar expression with appropriate superscript defines \( C^{II} \). A plane stress state can be described by the two principal values of \( \tilde{s}^I \) and \( \tilde{s}^{II} \)

\[
\tilde{s}_{1,2} = \frac{1}{2} \left( \tilde{s}_{xx} + \tilde{s}_{yy} \pm \sqrt{(\tilde{s}_{xx} - \tilde{s}_{yy})^2 + 4\tilde{\sigma}_{xy}^2} \right)
\]

(6.74)

with the appropriate superscripts (\([.]^I\) and \([.]^{II}\)) for each stress. The expressions of the two isotropic yield functions considered in Barlat et al. (2003) are:

\[
\phi^I (s) = |s_1 - s_2|^a, \quad \phi^{II} (s) = |2s_2 + s_1|^a + |2s_1 + s_2|^a
\]

(6.75)

leading to the resulting anisotropic yield function

\[
\phi = \phi^I (\tilde{s}^I) + \phi^{II} (\tilde{s}^{II}) = 2\sigma_{red}^a.
\]

(6.76)

This expression reduces to Hershey’s isotropy yield condition (Equation (6.12)) when the matrices \( C^I \) and \( C^{II} \) are both taken as the identity matrix so that \( \tilde{s}^I = \tilde{s}^{II} = s \). Out of the ten anisotropy coefficients, two are set to zero \( (C^I_{12} = C^{II}_{21} = 0) \) to reduce the number of parameters in the formulation. The other eight coefficients can be determined using as input the values of the stresses and \( r \) values in tension along three directions, the balanced biaxial flow stress \( \sigma_b \) and the biaxial strain ratio \( r_b \) defined later (Section 6.5.1).

6.4.7 CB2001 Yield Criterion

This criterion is developed in the framework of the theory of representation of tensor functions (e.g., Boehler 1978, Liu 1982). Assuming that yielding is insensitive to hydrostatic pressure, for an isotropic material, the yield function depends on stress through \( J_2 = tr s^2 / 2 \) and \( J_3 = tr s^3 / 3 \), the second and third invariants of the stress deviator \( s \), respectively. To
introduce orthotropy in the expression of an isotropic criterion, Cazacu and Barlat (2001, 2003b) proposed generalizations ($J_2^o$ and $J_3^o$) of the stress deviator invariants. $J_2^o$ and $J_3^o$ were required to be homogeneous functions of degree two and three in stresses, respectively, that reduce to $J_2$ and $J_3$ for isotropic conditions, are insensitive to pressure, and are invariant to any transformation belonging to the symmetry group of the material. Hence, relative to $(\mathbf{x}, \mathbf{y}, \mathbf{z})$, $J_3^o$ must be of the form

$$J_3^o = \frac{1}{27} (b_1 + b_2) \sigma_{xx}^3 + \frac{1}{27} (b_3 + b_4) \sigma_{yy}^3 + \frac{1}{27} \left[ (b_1 + b_4) - b_2 - b_3 \right] \sigma_{zz}^3$$

$$+ 2b_{11} \sigma_{xy} \sigma_{xz} \sigma_{yz} + \frac{2}{9} (b_1 + b_4) \sigma_{xx} \sigma_{zz} \sigma_{yy} - \frac{1}{9} (b_1 \sigma_{yy} + b_2 \sigma_{zz} + b_3 \sigma_{xx}) \sigma_{xy}^2$$

$$- \frac{1}{9} (b_3 \sigma_{zz} + b_4 \sigma_{xx}) \sigma_{yy}^2 - \frac{1}{9} \left[ (b_1 - b_2 + b_4) \sigma_{xx} + (b_1 - b_3 + b_4) \sigma_{yy} \right] \sigma_{zz}^2$$

$$- \frac{\sigma_{yy}^2}{3} \left[ (b_6 + b_7) \sigma_{xx} - b_6 \sigma_{yy} - b_7 \sigma_{zz} \right]$$

$$- \frac{\sigma_{zz}^2}{3} \left[ 2b_9 \sigma_{yy} - b_8 \sigma_{zz} - (2b_9 - b_8) \sigma_{xx} \right]$$

$$- \frac{\sigma_{xx}^2}{3} \left[ 2b_{10} \sigma_{zz} - b_9 \sigma_{yy} - (2b_{10} - b_9) \sigma_{xx} \right]$$

(6.77)

where all the coefficients $b_k$ ($k = 1, \ldots, 11$) reduce to one for isotropic conditions. Similarly, the generalization of $J_2$ can be expressed in $(\mathbf{x}, \mathbf{y}, \mathbf{z})$ as

$$J_2^o = \frac{a_1}{6} (\sigma_{xx} - \sigma_{yy})^2 + \frac{a_2}{6} (\sigma_{yy} - \sigma_{zz})^2 + \frac{a_3}{6} (\sigma_{xx} - \sigma_{zz})^2$$

$$+ a_4 \sigma_{xy}^2 + a_5 \sigma_{xz}^2 + a_6 \sigma_{yz}^2$$

(6.78)

where all the coefficients $a_k$ ($k = 1, \ldots, 6$) reduce to unity in the isotropic case. Note that $J_2^o$ corresponds to Hill’s 1948 quadratic yield function. Using these generalized invariants any isotropic yield criterion can be extended to describe orthotropy. In Cazacu and Barlat (2001), this approach was used to extend Drucker’s 1949 isotropic yield criterion. Hence, the proposed orthotropic criterion is

$$\phi = (J_2^o)^3 - c (J_3^o)^2 = k^2.$$  

(6.79)

For a full stress state, the criterion involves 18 material parameters. In the case of a sheet, where the only non-zero stress components are the in-plane stresses ($\sigma_{xx}$, $\sigma_{yy}$, $\sigma_{xy}$), the criterion reduces to

$$\phi^o = \left[ \frac{1}{6} (a_1 + a_3) \sigma_{xx}^2 - \frac{a_1}{3} \sigma_{xx} \sigma_{yy} + \frac{1}{6} (a_1 + a_2) \sigma_{xy}^2 + a_4 \sigma_{xy}^2 \right]^3$$

$$- c \left\{ \frac{1}{27} (b_1 + b_2) \sigma_{xx}^3 - \frac{1}{9} (b_1 \sigma_{xx} + b_4 \sigma_{yy}) \sigma_{xx} \sigma_{yy} \right\}$$

$$+ \frac{1}{27} (b_3 + b_4) \sigma_{yy}^3 - \frac{1}{3} \sigma_{xy}^2 \left[ (b_5 - 2b_{10}) \sigma_{xx} - b_5 \sigma_{yy} \right] = k^2.$$  

(6.80)
The 10 anisotropy coefficients and the value of $c$ can be determined using different methods, for instance from the measured uniaxial yield stresses $\sigma_\xi$ and strain ratios $r_\xi$ in five different orientations $\xi$ and the value of the balanced biaxial stress, $\sigma_b$ (see more details in Cazacu and Barlat 2001, 2003b).

### 6.4.8 Strain Rate Potentials

Hill (1987) proved the existence of the dual conjugate to a stress potential for rate independent perfect plasticity, i.e., its inverse expressed in terms of the dual strain rate variables. For instance, Fortunier (1989) proposed a dual formulation for the single crystal stress potential obeying the Schmid law and Van Houtte (1994, 2001) for a polycrystal. In general, it is difficult to find the dual conjugate of stress potentials analytically. However, it is possible to define strain rate potentials that describe the material anisotropy as an independent definition of the material behavior. This is formally identical to deriving yield functions. Because a strain rate potential and the associated stress potential are dual of each other, their mathematical properties are similar. Strain rate potentials are useful for rigid plasticity finite elements (FE) analysis and design codes. Stress and strain rate potentials can be used with equal degree of success (e.g., Li et al., 2001) but stress potentials are discussed in more details in this review because they have received more attention in the literature.

The strain rate potential $\tilde{\phi}$ proposed by Barlat et al. (1993) has a structure similar to the stress potential Yld91 but in strain rate space

$$\tilde{\phi} = \frac{2\dot{\varepsilon}_1 - \dot{\varepsilon}_2 - \dot{\varepsilon}_3}{3}^a + \frac{2\dot{\varepsilon}_2 - \dot{\varepsilon}_3 - \dot{\varepsilon}_1}{3}^a + \frac{2\dot{\varepsilon}_3 - \dot{\varepsilon}_1 - \dot{\varepsilon}_2}{3}^a = k\dot{\varepsilon}_{\text{red}}^{a}.$$  \hspace{1cm} (6.81)

This potential operates on a linearly transformed rate of deformation tensor

$$\dot{\varepsilon} = L\dot{\varepsilon}^p$$  \hspace{1cm} (6.82)

which contains six independent parameters. Numerical applications showed that this potential leads to good approximations of polycrystal potentials using $a = 4/3$ and $a = 3/2$ for FCC and BCC materials, respectively. It was also shown that the shape of the corresponding potential was a very close approximation to the dual shape of Yld91. Barlat et al. (1998) introduced a potential, pseudo-dual of the stress potential Yld96 for plane stress cases only

$$\tilde{\phi} = \varsigma_1 |\dot{\varepsilon}_1|^a + \varsigma_2 |\dot{\varepsilon}_2|^a + \varsigma_3 |\dot{\varepsilon}_1 + \dot{\varepsilon}_2|^a = \dot{\varepsilon}_{\text{red}},$$  \hspace{1cm} (6.83)

where $\varsigma_k$ are functions of the principal directions of the linearly transformed strain rate tensor.

### 6.5 Application to Sheet Forming and Formability

#### 6.5.1 Mechanical testing

Many issues have been addressed in determining yield surfaces experimentally. For instance, the definition of yield has been the subject of discussion (Paul, 1968; Hecker, 1976). In the context of this chapter, yield is defined conventionally at a certain offset of plastic strain
(e.g., 0.2%) and cyclic loading is not considered. Multiaxial experiments have been used to characterize a yield surface (see Hecker, 1976, for a review). With respect to the normality and convexity rules, Hecker did not find clear evidence of any violation, which again, supports the associated flow rule. Recently, Banabic et al. (2002a) improved a procedure for biaxial testing of cruciform specimens machined from thin sheets, and for measuring the first quadrant of the yield locus (both stresses positive). In these tests, the onset of plastic deformation is detected from temperature measurements of specimens using an infrared thermo-couple positioned at an optimized distance. This method is based on the fact that the specimen temperature drops first due to thermo-elastic cooling and then rises significantly when the dissipative plastic flow initiates. Figure 6.4 illustrates how the load at yield is determined based on this principle. In spite of this and other types of improvements, multiaxial testing is tedious, difficult to interpret and not suitable for quick characterization of anisotropy. This is more a technique for careful verifications of concepts and theories. Therefore, other methods are necessary to identify material coefficients in constitutive equations, in particular for sheets.

**Figure 6.4:** Load vs. displacement and temperature vs. displacement for a cruciform sheet specimen (3103-O) deformed in biaxial tension.

**Figure 6.5:** Bulge test schematic diagram.
Anisotropic properties can be assessed by performing uniaxial tension tests in the $x$ and $y$ axes (rolling and transverse directions, respectively), and in direction at $\xi$ degrees with respect to $x$. Practically, the anisotropy is characterized by the yield stresses $\sigma_0$, $\sigma_{45}$, $\sigma_{90}$, the $r$ values $r_0$, $r_{45}$, $r_{90}$, their respective average $\bar{q} = (q_0 + 2q_{45} + q_{90})/4$ and variations $\Delta\bar{q} = (q_0 - 2q_{45} + q_{90})/2$. Directional tension of wide specimens can be used to characterize plane strain tension anisotropy (Wagoner 1979; Taha 1995). This test does not produce a uniform state of stress within the specimen and generally leads to more experimental scatter than the uniaxial test.

The balanced biaxial yield stress ($\sigma_b$) is an important parameter to measure. This stress can be obtained by conducting a hydraulic bulge test (Young et al. 1981). Figure 6.5 shows a schematic diagram of this test in which a sheet blank is clamped between a die with a large circular opening and a holder. A pressure $p$ is gradually applied under the blank, which bulges in a quasi-spherical shape. The curvature $C$ and strains at the pole of the specimen are measured independently using mechanical or optical instruments. The stress $\sigma = p/2Ct$ is simply obtained from the membrane theory using the calculated thickness $t$. This test is interesting not only because it gives information on the yield surface but also because it allows measurements of the hardening behavior up to strains of about twice those achieved in uniaxial tension. However, the yield point is not well defined in this test because of the low curvature of the specimen in the initial stage of deformation. As for uniaxial tension, this test can be conducted at different strain rates in order to assess the strain rate sensitivity parameter $m$ (defined in Section 6.3.2).

As noted above, the effective stress $\sigma_{red}$, which is defined by the choice of a plastic potential, is equated to a function of some measure of the accumulated plastic strain and correspond to the flow stress in some particular loading conditions. Frequently, $\sigma_{red}$ is identified to the flow stress in uniaxial or balanced biaxial tension.
Because the biaxial stress state in the bulge test is not exactly balanced, measures of the corresponding strain state might lead to substantial errors. This is because the yield locus curvature is usually high in this stress state (see Figure 6.1). Barlat et al. (2003) proposed the disk compression test, which gives a measure of the flow anisotropy for a balanced biaxial stress state, assuming that hydrostatic pressure has no influence on plastic deformation. In this test, a 12.7 mm disk is compressed through the thickness direction of the sheet. The strains measured in the $x$ and $y$ directions lead to a linear relationship in which the slope ($\frac{d\varepsilon_{p_y}}{d\varepsilon_{p_x}}$) is denoted by $r_b$ by analogy to the $r$ value in uniaxial tension. This parameter is a direct measure of the slope of the yield locus at the balanced biaxial stress state. Figure 6.6 shows deformed specimens for 6111-T4 aluminum alloy sheet samples processed with two different routes and the corresponding strain measures performed during this test. Pöhlandt et al. (2002) have proposed to use the biaxial tensile testing machine to determine the coefficient $r_b$.

Simple shear tests (Rauch and Schmitt, 1989) can be carried out to characterize the anisotropic behavior of the simple shear flow stress. A relatively simple device mounted on a standard tensile machine is needed for this test. A rectangular specimen is clamped with two grips, which move in opposite directions relative to each other. By simply reversing the direction of the grip displacements, forward and reverse loading sequences can be conducted in order to measure the Bauschinger effect. In-plane tension-compression tests of thin sheets have also been used for this purpose (Kuwabara, 1995), but special precautions need to be taken because of the buckling tendency to the specimen during the compressive stage.

Although the above described tests and other experimental procedures are available to test materials in different stress states, it is not always possible to probe all of them. In this case, microstructure modeling can be used to replace the missing experimental data. The resistance to shear through the thickness of a sheet is not readily available from experimental measurements. However, crystal plasticity with a measure of the crystallographic texture of the sheet can be used to compute this data.

### 6.5.2 Analysis and Treatment of the Test Results

Two separate issues have to be considered in developing the material constitutive behavior. First, develop mathematical formulations for yield functions, flow rules and hardening laws, which are appropriate for the class of materials studied and obey physical and mechanical principles, as discussed in the previous sections. Second, analyze the tests and determine the data that best describe the material properties to identify the constitutive parameters. For instance, the yield stresses can be used as input data to calculate the anisotropic yield function coefficients. However, as mentioned above, the yield stress from the bulge test is not very accurate. Moreover, any stress at yield is determined in the region of the stress-strain curve where the slope is the steepest, which might involve additional inaccuracy. Finally, the yield stress is associated with a very small plastic strain and might not reflect the anisotropy of the material over a larger strain range. For these reasons, the flow stresses at equal amount of plastic work along different loading paths could be selected as input data instead of the yield stress. Figure 6.7 shows how the flow stresses in tension and balanced biaxial tension (bulge test) can be defined at equal amount of plastic work ($W_p^u = W_p^b$). For many aluminum alloys, experimental observations show that after a few percent plastic strains, the flow stress anisotropy does not vary significantly.
Similar remarks hold for $r$ values, which can be defined as instantaneous quantities at yield or as the standard slope of the width strain–thickness strain curve over a given deformation range in tension. In the one hand, the yield stresses and instantaneous $r$ values at yield are more appropriate to define the coefficients of the yield function. On the other hand, stresses (called flow stresses) defined at a given amount of plastic work and standard $r$ values can characterize the average behavior of the material over a finite deformation range. These values might be more suitable and more descriptive of the average response of the material for sheet forming simulations. In this case, it would be more appropriate to talk about flow function and flow surface instead of yield function and yield surface although, mathematically, yield or flow functions are identical concepts.

**Figure 6.7:** Flow stress at equivalent amount of plastic work ($W_{p}^{u} = W_{p}^{b}$).

**Figure 6.8:** Tensile anisotropy and yield locus for 3103-O alloy sheet.
6.5.3 Application to 3103-O Aluminum Alloy Sheet Sample

Figure 6.8 shows the experimental normalized yield stress and $r$ value as a function of the tensile direction for a 3103-O aluminum alloy sheet sample and the theoretical curves obtained with the BBC2000, Yld2000-2d and CB2001 yield functions. This Figure also shows the yield locus for this material measured in the first quadrant using biaxial tensile testing of cruciform specimens (Banabic et al. 2002b), and predicted with the three yield functions. The three models are able to describe tensile anisotropy and the yield locus shape of this alloy very well. In general, these models lead to yield surface shapes consistent with those predicted with polycrystal models. Moreover, they can capture the so-called “anomalous behaviors of 1st and 2nd orders.”

6.5.4 Plastic Flow Localization

A heavily cold rolled aluminum alloy is almost perfectly plastic (i.e., does not exhibit strain hardening). As a result, the strains at which plastic flow localization occurs in uniaxial and plane strain tension modes are very small. However, it is well known that the strain at which plastic flow localization occurs in this type of material is significantly higher when the material is deformed in balanced biaxial tension (Hecker 1978). This behavior can be explained with the concept of yield surface shape.

As discussed above, the yield surface shape for an aluminum alloy sheet can be characterized by the existence of regions of high curvatures. These (rounded) vertices are usually associated with plastic flow localization. In fact, based on the interpretation of Sowerby and Duncan (1971), the shape of the yield surface with respect to plane strain states is the important feature because localization is a plane strain mode of deformation. In this mode, compatibilities are preserved when strains are concentrated only in a small zone while the bulk of the material ceases to deform plastically. Equilibrium between the localization zone and the bulk material is governed by

$$\frac{\sigma_1^{(z)}}{\sigma_1^{(z)}} = \frac{A^{(z)}}{A} \leq 1 \tag{6.84}$$
where \( A \) is the load carrying area and the subscript \((.)\) denotes quantities in the zone. Plastic flow localization initiates in a zone for which \( A^{(z)} \) is smaller than \( A \) because of material heterogeneities or deformation gradients. Consequently, the ratio \( \sigma_1/\sigma_1^{(z)} \) must be less than 1. As an illustration, Figure 6.9 shows schematically the yield surfaces of two perfectly plastic materials deforming in balanced biaxial stretching. In Case (a), premature failure is expected because the major stress in the localization zone (plane strain) cannot increase significantly with respect to the major stress in the bulk of the sheet, i.e., \( \sigma_1/\sigma_1^{(z)} = 1 \). Equilibrium within the materials cannot be fulfilled. However, in Case (b), equilibrium between the plane strain localization zone and the bulk of the sheet can be achieved because \( \sigma_1/\sigma_1^{(z)} \) can be less than 1. In this case, the limit strain is expected to be significantly higher than the limit strain in tension. Thus, this simple example underlines the importance of the yield surface shape with respect to plastic flow localization.

### 6.5.5 Cup Drawing Simulation

The numerical simulation of the reverse deep drawing of a cup illustrates the application of yield surface plasticity in the sheet forming area. This test was a benchmark problem proposed by the committee of the Numisheet'99 conference (Gelin and Picart, 1999). The specific dimensions of the tools and the process conditions are given in the Numisheet'99 proceedings. The blank was machined in a 6016-T4 aluminum alloy sheet. Anisotropy was characterized with the results of uniaxial tension tests (yield stresses and \( r \) values) conducted in seven directions in the plane of the sheet.

The yield function Yld2000-2d was implemented in a finite element (FE) code (Yoon et al., 2000a). Isotropic hardening was assumed with \( \sigma_{red} \) obtained from Equation (6.76) and the reduced strain defined as \( \dot{\varepsilon}_{red} = \sigma_{ij}\dot{\varepsilon}^{ij}_{red}/\sigma_{red} \). The hardening curve

\[
\sigma_{red} = \sigma_u = h (\varepsilon_{red}) \dot{\varepsilon}^m_{red} \tag{6.85}
\]

was identified with the uniaxial stress-strain curve in the rolling direction although other hardening curves such as balanced biaxial (bulge test) could be used. The viscosity effects were neglected since \( m \approx 0 \) for aluminum alloys at room temperature. Because \( \sigma_{red} \) is a homogeneous function of 1\textsuperscript{st} degree, the flow rule (Equation (6.22)) was used with \( \lambda = \dot{\varepsilon}_{red} \). Additional information about the implementation of anisotropic yield functions into finite element codes is available elsewhere (Yoon et al. 1995, 1999, 2000b). All the details of the draw and redraw simulation conditions were given by Yoon et al. (2004). Three simulations were conducted assuming either isotropic or anisotropic constitutive coefficients but with an identical stress-strain curve. For isotropy, the exponent \( a \) was equal to 2 (Huber-Mises) and to 8 (Hershey 1954, Hosford 1972), and the matrices \( C^I \) and \( C^{II} \) (Equation (6.75)) were both taken as the identity matrix to reduce \( \tilde{s}^I \) and \( \tilde{s}^{II} \) to the stress deviator \( s \). For anisotropy, \( a \) was equal to 8 and the coefficients of \( C^I \) and \( C^{II} \) were calculated from the mechanical properties of the 6016-T4 sheet. The corresponding tensile anisotropy (normalized yield stress, \( r \) value) and tricomponent yield surface are shown in Figure 6.10.

Figure 6.11 shows the finite element mesh of the deformed cup at the intermediate and final stages of the drawing process for the anisotropic case. The predicted load profiles for forward and reverse drawings (not shown here) of the 6016-T4 sheet agreed qualitatively well with
the experimental profiles although the simulated results were underestimated (Yoon 2000a, 2004). This was mainly attributed to the lack of transverse shear stress in the general thin shell formulation.

Figure 6.12 shows the predicted and experimental thickness strain distributions in the rolling (0° or RD) and diagonal (45° from the RD) directions, respectively, after the drawing and redrawing operations. This figure shows that the thickness profiles in the two directions (0° and 45°) are different and that the Yld2000-2d model with anisotropic coefficients can capture this effect very well. Although the anisotropic simulation leads to a slight overestimation of the thickness profile in the diagonal (45°) direction, the agreement between predicted and experimental thickness distributions is generally excellent. The thin shell formulation without half thickness consideration of the sheet material presumably contributed to this overestimation. The isotropic response with $a = 8$ is characterized by an underestimation of the thickness in the two orientations (0° and 45°) and, of course, cannot lead to an anisotropic thickness profile. With $a = 2$ (Huber-Mises), the redraw simulation could not be carried out completely because plastic flow localization occurred before the end of the punch displacement. This cup drawing simulation shows the importance of a good description of the yield surface and plastic anisotropy. It would be interesting to perform computations and assess the improvements made with an anisotropic hardening law such as the combined
Conclusions

The succinct review presented in this chapter was an attempt to assess the progress made in the field of anisotropic plasticity over the last decades, and to give perspectives for future improvements. Anisotropy in sheet metals was discussed in terms of microscopic origins and macroscopic aspects. Much progress was made in the last decade because of the developments in the area of microstructure modeling, in particular crystal plasticity. This understanding has led to a refinement of anisotropic plasticity theories based on the classical concept of yield surfaces/plastic potentials. These improved formulations, combined with progress in other aspects of modeling such as contact mechanics, numerical methods, etc., have contributed to the success of numerical sheet forming process simulations. Yield surface plasticity offers a good compromise between realistic material modeling and computation speed, which is suitable for industrial applications. However, phenomena such as springback, wrinkling, flow localization and fracture have not been predicted with sufficient robustness and reliability for practical applications yet. Improvements are expected to result from more advanced phenomenological material descriptions, in particular, deformation-induced anisotropy in which, the concept of evolving yield surface will play an important role. These improvements will occur from further understanding and modeling of plasticity at and above the atomic scale.

Acknowledgements

The authors are indebted to Prof. J.J. Grácio and Dr. A.B. Lopes, University of Aveiro, Portugal, for providing the TEM micrographs, Dr. E.F. Rauch, Institut National Polytechnique de Grenoble, France, for providing the simple shear stress-strain curves, Mr. E. Llewellyn, Alcoa Technical Center, USA, for providing the crystallographic texture information, and Mr.
J.C. Brem, Alcoa Technical Center, for providing the disk compression test results. FB and J.W.Y are grateful to Alcoa Inc. for granting the permission to publish this work.

References


References


Burzyński, W., 1929. Über die Anstrengungshypothesen. Schweiz. Bauz. 95, 259-263.


References


References


References

Chapter 7: Artificial Neural Networks

Einar Broese and Hans Ulrich Löffler

Artificial Neural Networks (ANN’s) are a comparatively new way of information processing that was inspired by biological nervous systems and became very popular in the mid of 1980’s. They are composed of a large number of simple processing units which are tied together by weighted connections. A key feature of neural networks is their adaptability, i.e. their ability to improve themselves by learning from external data. This can be exploited in technical applications where a certain adaptation to real-world conditions is required. Neural Networks are often good at solving problems that are too complex for conventional technologies. Fields of application are pattern recognition, empirical modeling, diagnosis systems, robotics and adaptive control. Artificial Neural Networks can be implemented either in software or specialized hardware. Today there are multitudes of different neural network types including the most popular types Multilayer Perceptron (MLP) and Radial Basis Function Networks (RBFN). Typical tasks of any Neural Network application are network type selection, configuration, input preprocessing, prevention of overfitting and diagnostics of the networks internal state. A short outlook on future developments will also be given.

7.1 Introduction

Scientists and Engineers have been fascinated by the amazing capabilities of biological nervous systems, especially the human brain, for a long time. First ideas about artificial intelligence and artificial neural networks arose in the mid of the last century together with first computers (McCulloch 1941, Hebb 1949, Rosenblatt 1958). But it took 30 more years for the computers to become fast enough for substantial applications of artificial intelligence techniques. The big boom of neural network research was in between 1985 and 1995, when both fundamental and applied research was boosted in almost any area – scientific, technological, commercial and others. The expectations planted on neural networks were very high at that time, assuming a breakthrough in artificial intelligence for the next few years. But soon it became evident, that there is still a big gap between artificial and biological intelligence, which is obviously not only a question of computation power, but is also caused by insufficient concepts. Though the public interest in neural networks has calmed down in the meantime, scientific and engineering work in this area continues.
7.2 Basic Terms

An artificial neural network (ANN) is commonly accepted as a system of adaptive processing units, so called neurons, which are operating in parallel and interacting with each other. Neural networks have inputs and outputs, but no predefined input-output behavior. They get their functionality by training on external data.

A set of input values is called input pattern. The corresponding set of desired output values is called target pattern. Making a forward calculation means that the Neural network calculates the outputs for a given input pattern. The remaining error is the difference between output and target values for a given input. Network training minimizes the remaining error for each pair of input and target patterns.

The generalization performance of a trained neural network can be checked with a separate set of input and target patterns, so called validation data, which have not been used in the training before.

7.3 Fields of Application

7.3.1 Pattern Recognition/Classification

Pattern recognition means that noisy input signals (pictures, sounds) are compared with and assigned to certain objects. The neural network makes decisions and gives back the category number (digital output) for that object. Important applications are:

- speech recognition
- picture processing and diagnosis in medicine (X-ray, Ultrasound)
- interpreting of handwriting (mail delivery, electronic diaries)
- object identification in military applications (reconnaissance, target finding)
- quality control (e.g. sound control of tiles, porcelain,...)
- waste separation
- recognition of abnormal noise from engines, gearboxes, etc.
- classification of customers (e.g. credit worthiness)
- personal identification (face, finger prints)

(Bishop 1995)

7.3.2 Empirical Modeling

The task of empirical modeling is to find an approximate functional relationship for a given set of input and output pairs. Both input and output variables may be vector-valued, with each vector entry typically being continuous-valued. Mathematically spoken we look for a
function \( f : \mathbb{R}^n \rightarrow \mathbb{R}^m \), with \( n \geq 1 \) and \( m \geq 1 \). To each input pattern (a vector \( i \in \mathbb{R}^n \)) belongs an output target vector (a vector \( t \in \mathbb{R}^m \)), which allows the training of the neural network, i.e. the procedure to find the function \( f(i) = o \), with \( |o - t| = \text{minimum} \). If it is trained well the neural network will calculate the correct output value for any input. The advantage of neural networks as approximators of functional relationships is that nothing has to be known a priori about the functional relationship and the function may also be highly non-linear. In addition the neural network approach enables step by step learning in case where training patterns become available only successively. In control applications (neural control) the neural network is mostly incorporated into the control loop and works as an adaptive empirical model. Application areas:

- finance data prognosis (e.g. exchange rates)
- process control (chemical industry, basic industry,...)
- materials modeling
- robot control
- engine management of combustion engines
- driving of vehicles
- data filtering

A special case of empirical modeling is materials modeling with neural networks (see e.g. Chapter 46 of this book). The field of materials modeling is characterized by a high degree of complexity and poor possibilities of physical description from first principles. Therefore there is a high need for empirical modeling, e.g. for the calculation of steel hardness from alloy composition. A recent overview of the application of neural computing in steel industry can be found in Schlang et al., 2000.

### 7.4 Implementation

Artificial Neural Networks can be implemented either in software or in specialized hardware.

#### 7.4.1 Software

Today there is a huge variety of neural network simulators and software packages, commercial software and freeware, written in different programming languages for all kinds of computers. The following list gives a few examples, but is not intended to be complete:

- NeuroSolutions (NeuroDimension, http://www.nd.com)
- NeuralWorks (NeuralWare, http://www.neuralware.com/products.jsp)
7.5 Types of Artificial Neural Networks

7.5.1 Multilayer Perceptron

The Multilayer Perceptron (MLP) is the most common neural network type. It is often called “Backpropagation” or “Backprop” Network because of its learning rule (Rumelhart et al. 1986), or also “Sigmoid Network” hinting at the activation functions used.

The Multilayer Perceptron consists of an input layer, one or more hidden layers and an output layer. Neurons of neighboring layers are usually fully interconnected. It can be shown that it makes no sense to take more than one hidden layer, because any continuous function can be constructed by one layer, i.e. by superposition of its sigmoid base functions (Cybenko 1989).

The network’s output is calculated as follows:

\[
OH_j = \tanh \left( B1_j + \sum_i W1_{ij} \cdot I_i \right)
\]

\[
O_k = B2_k + \sum_j \left( W2_{jk} \cdot OH_j \right)
\]  

(7.1)
A basic feature of the Multilayer Perceptron is the sigmoid transfer function (\( \tanh \)) of the hidden units, also referred to as activation function. This function has a number of useful properties: It is linear in the origin and nonlinear at the ends, allowing the Multilayer Perceptron to represent linear and nonlinear functions, it is continuously differentiable, which is needed for gradient descend adaptation methods, and it exhibits a conservative extrapolative behavior, as it tends smoothly towards \( \pm 1 \) as the input tends to infinity.

Output neurons may also contain sigmoid functions. But typically it is better to take linear outputs, especially if the range of output values is not clearly defined in advance.

The weights of the network \( W_{1ij}, W_{2jk}, B_{1j}, \) and \( B_{2k} \) are adjusted such that the network produces for each input an output close to the target output. A certain amount of training patterns, consisting of inputs and related targets, will be used for the network training. The training patterns will be presented to the network in a random order followed by an adaptation of the network’s weights. The weights will be changed according to the so-called backpropagation rule, where the network error will be “propagated back” from the output layer to the
hidden layer:
\[
\Delta_k = (T_k - O_k) \\
\Delta W_{2jk} = \eta \cdot O H_j \cdot \Delta_k \\
\Delta B_{2k} = \eta \cdot \Delta_k \\
\Delta_j = (1 - OH_j^2) \cdot \sum_k (W_{2jk} \cdot \Delta_k) \\
\Delta W_{1ij} = \eta \cdot I_i \cdot \Delta_j \\
\Delta B_{1j} = \eta \cdot \Delta_j
\]  
(7.2)

where
- $\Delta_k$ error of the output layer neurons
- $\Delta_j$ error of the hidden layer neurons
- $T_k$ target vector
- $\eta$ learning rate

In a strict mathematical sense the backpropagation learning rule is a gradient descent optimization step on a quadratic error function $E$:
\[
\Delta W_i = -\eta \cdot \frac{\partial E}{\partial W_i} ; \quad E = \frac{1}{2} \cdot \sum_k (O_k - T_k)^2
\]  
(7.3)

where
- $W_i$ a certain neural network weight ($W_{1ij}, W_{2jk}, B_{1j}, B_{2k}$)
- $O_k$ the output vector of the network

Setting the formulas (7.1) into (7.3) one can derive the backpropagation rules (7.2).

In the literature there exist a lot of variants of the Multilayer Perceptron network type with varying architecture (number of layers, number of connections, activation functions) and learning rules (Riedmiller 1993, Haykin 1999).

Multilayer Perceptrons can be utilized for both pattern recognition and function approximation. But they are unsuitable in applications, which have a focus on fast online training with single patterns, because they cannot reduce the error for a single pattern without affecting the other patterns. The training of Multilayer Perceptrons has to be carried out in batch mode with a complete set of all possible training patterns.
7.5.2 Radial Basis Function Networks

A radial basis function (RBF) is a radially symmetric function around an associated center $C_n$ (or “kernel”) in the n-dimensional input space. It is usually a function of the Euclidean distance of the input vector $I$ to the center $C$:

$$r = \| I - C \| = \sqrt{\sum_i (I_i - C_i)^2} \quad (7.4)$$

Examples of radial basis functions are:

- inverse multiquadrics: $\varphi(r) = \frac{1}{\sqrt{r^2 + d^2}}$ with $d > 0 \quad (7.5)$
- Gaussian functions: $\varphi(r) = \exp\left(-\frac{r^2}{2\sigma^2}\right)$ with $\sigma > 0 \quad (7.6)$

Gaussian radial basis functions are probably the most frequently used ones.

![Figure 7.3: Superposition of 3 Gaussian RBF functions with $\sigma = 0.1$ on a $1 \times 1$ plane.](image)

The interest of neural network researchers in radial basis functions was initially caused by some analogies with the neural activation in the visual cortex (which is the part of the brain responsible for visual input processing). Different combinations of analogue input signals (patterns) will lead to the activation of different neurons. This feature can be immediately used for pattern recognition and other classifying problems.
On the other hand, radial basis functions can be used to construct any non-linear function and therefore they are a suitable choice for function modeling. The simplest type of a RBF network is a weighted sum of Gaussian functions.

The network’s output is the sum of all the activation factors $\Phi_j$, each multiplied with the corresponding weighting factor $W_{jk}$:

$$O_k = \sum_j (\Phi_j \cdot W_{jk})$$  \hspace{1cm} (7.7)

$\Phi_j$ can be the radial basis function itself, or better, the normalized to 1 radial basis function:

$$\Phi_j = \frac{\phi_j}{\sum_m \phi_m}$$  \hspace{1cm} (7.8)

In the latter case the network is called “normalized Radial Basis Function Network”.

Note that the individual “neuron” (center) of the Radial Basis Function Network is not as similar to a biological neuron, as in a Multilayer Perceptron. The neural process which leads from an input pattern to the activation of a single neuron, should be more complex and probably involves more than one neuron.

The learning of Radial Basis Function Networks is usually a two stage learning, consisting of the determination of the center positions (clustering) and the optimization of the weights. A simultaneous optimization of both types of internal parameters toward a common error function is not recommended, since this would be a poorly conditioned optimization problem with lots of similar solutions.

There is no single way to determine the position of local centers, but a variety of different methods, also called cluster methods. The task of clustering in general is to find a set of vectors...
Artificial Neural Networks

which represent a given multitude of data points (patterns) as good as possible. A standard cluster method is k-Means, a least-squares partitioning method, which minimizes the sum of the quadratic distances of all input vectors to their closest centers (MacQueen 1967). Another way is to start with a set of randomly distributed centers and move them pattern by pattern according to the following rule:

$$\Delta C_{ij} = \lambda (j) \cdot (I_i - C_{ij})$$  \hspace{1cm} (7.9)

where learning rate $\lambda$ is a number between 0 and 1 and $C_{ij}$ are the coordinates of the center. If j is the index of the centers ordered by distance to the given input, $I_i$, $\lambda(j)$ should be maximum for the closest center (j=0) and decrease with rising j. $\lambda$ should also decrease with time in order to “freeze” the optimum distribution.

The weight optimization is usually a gradient descent method (Press et al. 1989). In case of the upper figured simple Radial Basis Function Network, the weight changes will be calculated as

$$\Delta W_{jk} = \eta \cdot \Phi_j \cdot (T_k - O_k).$$  \hspace{1cm} (7.10)

RBF models are actually a superposition of local models, each representing a certain section of the multidimensional input space. In the above mentioned simple case these local models are adaptive constants. But they can also be a function of the input values. The next figure shows the output of a Radial Basis Function Network, which consists of 3 adalines (adaptive linear models). In this case the local models are linear. Their contribution to the output is determined by the radial basis functions. In the surrounding of each center the global model is similar to the respective local model. The transition between the local models can be sharp or smooth, depending on the Gaussian parameter $\sigma$.

For an overview about Radial Basis Function Networks see Ghosh et al. 2000.

7.5.3 More Network Types

- **Recurrent Networks**: Recurrent networks differ from feed forward networks by the feature that the output of at least one neuron is fed back to its input, directly or via other neurons. Thus the state and the outputs of a recurrent network will depend not only on recent input values but also on the history.

- **Hopfield Networks**: Hopfield networks, or associative networks, consists of a layer of fully interconnected neurons, which excite or damp each other. They are typically used for classification. Given a distorted input pattern, the Hopfield network associates it with an undistorted pattern stored in the network (Hertz 1991).

- **Kohonen Self Organizing Maps**: Kohonen maps are 1- or 2-dimensional lattices arranging their nodes by a special clustering algorithm, which moves not only the winner (closest center) but also its neighborhood (Kohonen 1995). Kohonen Self Organizing Map’s can be used for classification tasks. The resulting feature maps illustrate both data distribution and similarity relationships between the classes.
7.6 Kinds of Learning

7.6.1 Unsupervised Learning

Learning algorithms will be called unsupervised, if they use input values only. They aim to optimize the input data preprocessing. They are always just a part of the neural network training, i.e., they have to be combined with supervised learning methods. Examples for unsupervised learning are input clustering, self organizing maps (SOM) and Principal Component Analysis (PCA).

7.6.2 Supervised Learning

Neural network learning will be called supervised, if there is some kind of external evaluation of the network’s performance. It is usually a parameter optimization toward a minimum error function, calculated from the difference between output and target values, like gradient descent or backprop learning for Multilayer Perceptrons.

7.6.3 Reinforcement Learning

Reinforcement learning is a special case of supervised learning, which can be used in case of a non-differentiable error function, e.g., if the result is either good or bad. Parameter changes will be kept only if they have led to an improved result. Otherwise they will be reversed. The
Artificial Neural Networks

7.6.4 Bayesian Learning

The result of Bayesian training is not only a single set of network weights, but a probability distribution for each of the weights. This method costs a lot of computation time but has several advantages:

- reduced danger of overfitting due to averaging weights.
- the output uncertainty can be estimated.
- prior knowledge can be integrated by combining the new distribution with a prior distribution

(Bishop 1995)

7.7 Application Details

The success of any neural network application depends on a number of implementational issues, that will be discussed next.

7.7.1 Network Type Selection and Configuration

In principle any neural network application can be done with both Multilayer Perceptron (MLP) and Radial Basis Function Network (RBFN). But there are some differences in detail. The Radial Basis Function Network can be adapted with a single pattern in a local limited area without affecting the rest of the input space, what is impossible with Multilayer Perceptrons. If considering empirical modeling, a Multilayer Perceptron will need much more training steps to learn a certain function, but in exchange the model becomes smoother. So if an application needs a fast single pattern learning and if smoothness and derivative are less important compared to the function itself - a Radial Basis Function Network would be the better choice. But if smoothness and first derivative of the empirical model are important, and the network can be trained offline with a pattern collection, a Multilayer Perceptron should be preferred.

As already mentioned one hidden layer is sufficient for a Multilayer Perceptron in most cases. It has been shown that a Multilayer Perceptron with one hidden layer can model any continuous function (Cybenko 1989). More layers will not decrease the remaining error, but decrease the convergence rate.

The size of the network, i.e. number of inputs and hidden units should be as small as possible, to prevent overfitting and to save computation time. The number of training patterns in a batch training should exceed the number of free internal parameters (weights) by a factor of 3 to 10. This is especially important for noisy data.

The safest way is to start with small networks and to increase the network size step by step, if necessary.
7.7 Application Details

7.7.2 Input Selection

For the choice of input variables the following rules are applicable:

- The input has to be available at the time when the Neural Network is applied.
- The input must have an effect on the required output or decision.

The following inputs do not enhance the performance of a neural network and should therefore be omitted:

- constant values,
- strongly correlated inputs,
- very noisy inputs.

With an intelligent data pre-processing, the influence of the aforementioned inputs can be reduced. However, using these inputs occupies memory space and uses processing resources.

Categories, e.g. material types or colors, are best supplied as binary inputs (0, 1) where each category is defined as one input. If there are only two categories, one input is sufficient.

There is no principal restriction on the number of inputs. The only limitations are memory space and processing power.

7.7.3 Data Preprocessing and Input Scaling

The role of data preprocessing and input scaling is quite important, though often underestimated. Especially neural networks with different kinds of inputs need them to be scaled to a unique data range, e.g. \([-1, +1]\). If \(I\) is within the interval \([\text{min}, \text{max}]\) the linear scaled (or normed) input \(I' \in [-1, +1]\) is calculated as:

\[
I' = a \cdot I + b = \frac{2}{\text{max} - \text{min}} \cdot I - \frac{\text{max} + \text{min}}{\text{max} - \text{min}}
\] (7.11)

The factor “\(a\)” increases or decreases leverage and selectivity of the corresponding input.

Another well approved scaling type is the sigmoid input scaling, which lowers the danger of extrapolation errors.

\[
I' = c \cdot \tanh (a \cdot I + b)
\] (7.12)

If a representative set of input vectors is already given, the redundancy can be lowered using Principal Component Analysis (PCA). This leads to a new input vector of a lower than original dimension and to a better neural network performance.

7.7.4 Prevention of Overfitting

Overfitting is probably the most serious problem in function approximation with neural networks. It always occurs if big networks with lots of inputs are used to approximate a small number of noisy training patterns. Overtrained networks show good results on the training patterns themselves but poor results in case of interpolation or extrapolation.
Several strategies have been developed in the past to fight the overfitting problem:

1. **separation of training and validation data**: Existing training patterns are split into two parts: a training and a validation pattern set. Training patterns are used to train the network and validation patterns to test the network. The training consists of small learning steps with all patterns (epochs). At the beginning both training and validation error usually decrease. In case of overfitting the validation error stops to decrease and comes up again, while the training error continues to decrease. The training will be stopped after reaching the minimum validation error.

2. **weight decay**: Another indication for overfitting are high absolute values of the internal weights of the network. So the overfitting can be limited by keeping internal weights small. In order to lower weights they can be multiplied after each learning step with a factor \((1 - WD)\), with weight decay \(WD \ll 1\) (Werbos 1988).

3. **network pruning**: The risk of overfitting can be lowered by reducing the number of internal weights (free parameters) of the network. After reducing the number of inputs and hidden units to the necessary minimum, individual weights (connections) might be removed too. The importance of a connection can be estimated by tracing the weight value over training time. If the weight keeps changing more than its absolute value, the connection is likely to be not necessary and can be removed (Russel 1993).

4. **mixture of experts**: Instead of training just one neural network a whole set of networks can be trained. These networks might be of different types and configurations. The output is the average value of all networks (experts). If the networks are of the same type and size, they might be replaced by one network with average weights. This approach is close to Bayesian learning algorithms.

### 7.7.5 Optimization of Training Parameters

The method of division in training and validation patterns, outlined in the previous chapter, can also be used to optimize the network configuration (scaling type, input selection, number of hidden units) and training parameters (learning rates, weight decay, …). The optimization target would be the minimum error on validation patterns. If the neural network is aimed to be used as a dynamic model with continuous online training (adaptation), the validation test must be executed in a similar way, i.e. it should be a simulated online training on a time ordered pattern series.

### 7.7.6 Diagnostics of the Internal State

Neural network applications always face the problem, that people want to know, whether the networks are well trained or not. Similar to school, where the teacher cannot see the knowledge opening the student’s brain, it does not make much sense to look at the network’s weights. Students will be evaluated by giving them questions and comparing their answers to the correct ones. Similarly the networks have to be tested by checking their response to certain inputs. It is helpful to develop diagnostic tools for this purpose which
7.8 Future Prospects

- calculate the average error for a given set of training patterns.
- display the input-output-behavior of the neural network
- display the empirical function, learned by the network, together with the original training patterns.

7.8 Future Prospects

Research in neural networks is currently going in two directions – firstly the neurobiological modeling or computational neuroscience with the goal to understand brain processes, and secondly the improvement of existing applications.

A new development in neurobiological modeling are the so called spiking neurons or pulsed neural networks, which take into account the pulsing behavior of individual neurons (Maass 1999).

In current applications of neural networks like voice recognition, hand writing recognition, vision control, stock prediction, process optimization, robot control etc., we will see a continuous progress toward better performance and higher robustness.

Technical applications of neural networks, or adaptive systems in general, face the problem of safety. It is difficult to guarantee that a system capable of learning will never fail. Therefore neural network research will be partially focused on robust learning algorithms and overfitting prevention. In addition, if operational safety is strongly required like in nuclear power plants or auto pilots, neural network control has always to be accomplished by safety arrangements, which cover all possible outputs of the neural network.

Research in neurotechnology aims to replace lost functionality of the peripheral and central human nervous system by means of technical information processing technology. One example is the “Retina Implant Project” (http://www.nero.uni-bonn.de/projekte/ri/ri-index-en.htm), where a visual prosthesis for subjects blind from retinal degeneration is currently being developed.

Much has been written about the prospects of neural networks and artificial intelligence in general. Will computers ever reach human thinking? Will computers be as complex and as functional? Will machines ever get consciousness? The author’s point of view is that there seems to be no principal limitations on what machines can do. It is probably just a question of time. It’s not foolhardy to expect some big inventions in information processing in the next 10–20 years, which will clear up how the brain works and which will result in a rapid development of intelligent programs and machines.

References

8 Multiscale Discrete Dislocation Dynamics Plasticity

Hussein M. Zbib, Masato Hiratani, and Mutasem Shehadeh

8.1 Introduction

Deformation and strength of crystalline materials are determined to a large extent by underlying mechanisms involving various crystal defects, such as vacancies, interstitials and impurity atoms (point defects), dislocations (line defects), grain boundaries, heterogeneous interfaces and microcracks (planar defects), chemically heterogeneous precipitates, twins and other strain-inducing phase transformations (volume defects). Most often, dislocations define plastic yield and flow behavior, either as the dominant plasticity carriers or through their interactions with the other strain-producing defects.

Dislocation as a line defect in a continuum space was first introduced as a mathematical concept in the early 20th century by Volterra (1907) and Somigliana (1914). They considered the elastic properties of a cut in a continuum, corresponding to slip, disclinations, and/or dislocations. But associating these geometric cuts to dislocations in crystalline materials was not made until the year 1934. In order to explain the less than ideal strength of crystalline materials, Orowan (1934), Polanyi (1934) and Taylor (1934) simultaneously hypothesized the existence of dislocation as a crystal defect. Later in the late 50’s, the existence of dislocations was experimentally confirmed by Hirsch, et al. (1956) and Dash (1957). Presently these crystal defects are routinely observed by various means of electron microscopy.

A dislocation can be easily understood by considering that a crystal can deform irreversibly by slip, i.e. shifting or sliding along one of its atomic planes. If the slip displacement is equal to a lattice vector, the material across the slip plane will preserve its lattice structure and the change of shape will become permanent. However, rather than simultaneous sliding of two half-crystals, slip displacement proceeds sequentially, starting from one crystal surface and propagating along the slip plane until it reaches the other surface. The boundary between the slipped and still unslipped crystal is a dislocation and its motion is equivalent to slip propagation. In this picture, crystal plasticity by slip is a net result of the motion of a large number of dislocation lines, in response to applied stress. It is interesting to note that this picture of deformation by slip in crystalline materials was first observed in the nineteenth century by Mügge (1883) and Ewing and Rosenhain (1899). They observed that deformation of metals proceeded by the formation of slip bands on the surface of the specimen. Their interpretation of these results was obscure since metals were not viewed as crystalline at that time.

Over the past seven decades, experimental and theoretical developments have firmly established the principal role of dislocation mechanisms in defining material strength. It is now understood that macroscopic properties of crystalline materials are derivable, at least in principle, from the behavior of their constituent defects. However, this fundamental understanding
has not been translated into a continuum theory of crystal plasticity based on dislocation mechanisms. The major difficulty in developing such a theory is the multiplicity and complexity of the mechanisms of dislocation motion and interactions that make it impossible to develop a quantitative analytical approach. The problem is further complicated by the need to trace the spatiotemporal evolution of a very large number of interacting dislocations over very long periods of time, as required for the calculation of plastic response in a representative volume element. Such practical intractability of the dislocation-based approaches, on one hand, and the developing needs of material engineering at the nano and micro length scales on the other, have created the current situation when equations of crystal plasticity used for continuum modeling are phenomenological and largely disconnected from the physics of the underlying dislocation behavior.

Bridging the gap between dislocation physics and continuum crystal plasticity has become possible with the advancement in computational technology with bigger and faster computers. To this end, over the past two decades various discrete dislocation dynamics models have been developed. The early discrete dislocation models were two-dimensional (2D) and consisted of periodic cells containing multiple dislocations whose behavior was governed by a set of simplified rules (Lepinoux and Kubin 1987; Ghoniem and Amodeo 1988; Groma and Pawley 1993; Van der Giessen and Needleman 1995; Wang and LeSar 1995; Le and Stumpf 1996). These simulations, although served as a useful conceptual framework, were limited to 2D and, consequently, could not directly account for such important features in dislocation dynamics as slip geometry, line tension effects, multiplication, dislocation intersection and cross-slip, all of which are crucial for the formation of dislocation patterns. In the 90’s, development of new computational approaches of dislocation dynamics (DD) in three-dimensional space (3D) generated hope for a principal breakthrough in our current understanding of dislocation mechanisms and their connection to crystal plasticity (Kubin and Canova 1992; Canova, Brechet et al. 1993; Hirth, Rhee et al. 1996; Raabe et al. 1996; Zbib, Rhee et al. 1996). In these new models, dislocation motion and interactions with other defects, particles and surfaces are explicitly considered. However, complications with respect to dislocation multiplications, self-interactions and interactions with other defects, and keeping track of complex mechanisms and reactions have provided a new set of challenges for developing efficient computational algorithms.

The DD analysis and its computer simulation modeling devised by many researchers (Ghoniem and Amodeo 1988; Canova, Brechet et al. 1992; Kubin 1993; Schwarz and Tersoff 1996; Zbib, Rhee et al. 1996; Zbib, Rhee et al. 1998) have been advanced significantly over the past decade. This progress has been further magnified by the idea to couple DD with continuum mechanics analysis in association with computational algorithms such as finite elements. This coupling may pave the way to better understanding of the local response of materials at the nano and micro scales and globally at the macroscale (Zbib and Diaz de la Rubia 2002), increasing the potential for future applications of this method in material, mechanical, structural and process engineering analyses. In the following, the principles of DD analysis will be presented followed by the procedure for the measurement of local quantities such as plastic distortion and internal stresses. The incorporation of DD technique into the three-dimensional plastic continuum mechanics-based finite elements modeling will then be described. Finally, examples are provided to illustrate the applicability of this powerful technique in material engineering analysis.
8.2 Theoretical Fundamentals of the Method

In order to better describe the mathematical and numerical aspects of the DD methodology, first we will identify the basic geometric conditions and kinetics that control the dynamics of dislocations. This will be followed by discussion of the dislocation equation of motion, elastic interaction equations, and discretization of these equations for numerical implementation.

8.2.1 Kinematics and Geometric Aspects

A dislocation is a line defect in an otherwise perfect crystal described by its line sense vector $\xi$ and Burgers vector $b$. The Burgers vector has two distinct components: edge, perpendicular to its line sense vector, and screw, parallel to its line sense vector. Under loading, dislocations glide and propagate on slip planes causing deformation and change of shape. When the local line direction becomes parallel to the Burgers vector, the dislocation may propagate into other slip planes. This switching of the slip plane, which makes the motion of dislocations three-dimensional, and is better known as cross slip is an important recovery phenomena to be dealt with in dislocation dynamics. In addition to glide and cross slip, dislocations can also climb in a non-conservative three-dimensional motion by absorbing and/or emitting intrinsic point defects, vacancies, and interstitials. Some of these phenomena become important at high loads level or temperatures when point defects become more mobile. In summary, the 3D dislocation dynamics accounts for the following geometric aspects:

- Dislocation topology; 3D geometry, Burgers vector and line sense.
- Identification of all possible slip planes for each dislocation.
- Changes in the dislocation topology when part of it cross-slips and or climbs to another plane.
- Multiplication and annihilation of dislocation segments.
- Formation of complex connections and intersections such as junctions, jogs and branching of the dislocation in multiple directions.

8.2.2 Kinetics and Interaction Forces

The dynamics of the dislocation is governed by a “Newtonian” equation of motion, consisting of an inertia term, damping term, and driving force arising from short-range and long-range interactions. Since the strain field of the dislocation varies as the inverse of distance from the dislocation core, dislocations interact among themselves over long distances. As the dislocation moves, it has to overcome internal drag, and local barriers such as the Peierls stresses. The dislocation may encounter local obstacles such as stacking fault tetrahedra, defect clusters and vacancies that interact with the dislocation at short ranges and affect its local dynamics. Furthermore, the internal strain field of randomly distributed local obstacles gives rise to stochastic perturbations to the encountered dislocations, as compared with deterministic forces such as the applied load. This stochastic stress field also contributes to the spatial dislocation patterning in the later deformation stages. Therefore, the strain field of local obstacles adds
spatially irregular uncorrelated noise to the equation of motion. In addition to the random strain fields of dislocations or local obstacles, thermal fluctuations also provide a stochastic source in dislocation dynamics. Dislocations also interact with free surfaces, cracks, and interfaces, giving rise to what is termed as image forces. In summary, the dislocation may encounter the following set of forces:

- Drag force, \( Bv \) where \( B \) is the drag coefficient and \( v \) is the dislocation velocity.
- Peierls stress \( F_{\text{Peierls}} \).
- Force due to externally applied loads, \( F_{\text{external}} \).
- Dislocation-dislocation interaction force \( F_{D} \).
- Dislocation self-force \( F_{\text{self}} \).
- Dislocation-obstacle interaction force \( F_{\text{obstacle}} \).
- Image force \( F_{\text{image}} \).
- Osmotic force \( F_{\text{osmotic}} \) resulting from non-conservative motion of dislocation (climb) and results in the production of intrinsic point defects.
- Thermal force \( F_{\text{thermal}} \) arising from thermal fluctuations.

The DD approach attempts to incorporate all of the aforementioned kinematics and kinetics aspects into a computational traceable framework. In the numerical implementation, three-dimensional curved dislocations are treated as a set of connected segments as illustrated in Figure 8.1. It is possible to represent smooth dislocations with any desired degree of realism, provided that the discretization resolution is taken high enough for accuracy (limited by the size of the dislocation core radius \( r_0 \), typically the size of one Burgers vector). In such a representation, the dynamics of dislocation lines is reduced to the dynamics of discrete degrees of freedom of the dislocation nodes connecting the dislocation segments.

### 8.2.3 Dislocation Equation of Motion

The velocity \( v \) of a dislocation segment \( s \) is governed by a first order differential equation consisting of an inertia term, a drag term and a driving force vector (Hirth 1992; Hirth, Zbib et al. 1998; Huang, Ghoniem et al. 1999), such that

\[
m_s \ddot{v} + \frac{1}{M_s(T, p)} v = F_s \quad \text{with} \quad m_s = \frac{1}{v} \left( \frac{dW}{dv} \right)
\]  

\[
F_s = F_{\text{Peierls}} + F_{D} + F_{\text{self}} + F_{\text{external}} + F_{\text{obstacle}} + F_{\text{image}} + F_{\text{osmotic}} + F_{\text{thermal}}
\]  

In the above equation the subscript \( s \) stands for the segment, \( m_s \) is defined as the effective dislocation segment mass density, \( M_s \) is the dislocation mobility which could depend both on the temperature \( T \) and the pressure \( p \), and \( W \) is the total energy per unit length of a moving dislocation (elastic energy plus kinetic energy). As implied by (8.1b), the glide force vector...
8.2 Theoretical Fundamentals of the Method

\( F_s \) per unit length arises from a variety of sources described in the previous section. The following relations for the mass per unit dislocation length have been suggested (Hirth, Zbib et al. 1998) for screw \((m_s)_{\text{screw}}\) and edge \((m_s)_{\text{edge}}\) dislocations when moving at a high speed.

\[
(m_s)_{\text{screw}} = \frac{W_0}{v^2} (-\gamma^{-1} + \gamma^{-3}) \tag{8.2a} \\
(m_s)_{\text{edge}} = \frac{W_0 C^2}{v^4} \left( -16\gamma_l - 40\gamma_l^{-1} + 8\gamma_l^{-3} \\
+ 14\gamma + 50\gamma^{-1} - 22\gamma^{-3} + 6\gamma^{-5} \right) \tag{8.2b}
\]

where \( \gamma_l = (1 - v^2/C_l^2)^{1/2}, \quad \gamma = (1 - v^2/C^2)^{1/2} \), \( C_l \) is the longitudinal sound velocity, \( C \) is the transverse sound velocity, \( \nu \) is Poisson’s ratio, \( W_0 = \frac{Gb^2}{4\pi} \ln(R/r_0) \) is the rest energy for the screw per unit length, \( G \) is the shear modulus. The value of \( R \) is typically equal to the size of the dislocation cell (about 1000b, with b being the magnitude of the Burgers vector), or in the case of one dislocation is the shortest distance from the dislocation to the free surface (Hirth and Lothe 1982). In the non-relativistic regime when the dislocation velocity is small compared to the speed of sound, the above equations reduce to the familiar expression \( m = \beta \rho b^2 \ln(R/r_0) \), where \( \beta \) is a constant dependent on the type of the dislocation, and \( \rho \) is the mass density.

Figure 8.1: Discretization of dislocations loops and curves into nodes, segments and collocation points.
8.2.3.1 Dislocation Mobility Function

The reliability of the numerical simulation depends critically on the accuracy of the dislocation drag coefficient $B (= 1/M)$ that is material dependent. There are a number of phenomenological relations for the dislocation glide velocity $v_g$ (Kocks, Argon et al. 1975; Sandstrom 1977), including relations of power law forms and forms with an activation term in an exponential or as the argument of a sinh form. Often, however (Johnston and Gilman 1959; Sandstrom 1977) the simple power law form is adopted for expedience, e.g. $v_g \equiv v_s \left( \frac{\tau_e}{\tau_s} \right)^m$, resulting in nonlinear dependence of $M$ on the stress. In a number of cases of pure phonon/electron damping control or of glide over the Peierls barrier a constant mobility (with $m = 1$), predicts the results very well. This linear form has been theoretically predicted for a number of cases as discussed by Hirth and Lothe (1982).

Mechanisms to explain dislocation drag have been studied for long time and the drag coefficients have been estimated in numerous experimental and theoretical works by atomistic simulations or quantum mechanical calculations [see, for example, the review by Al’shitz (1992)]. The determination of each of the two components (phonon and electron drag) that constitute the drag coefficient for a specific material is not trivial, and various simplifications have been made, e.g. the Debye model neglects Van Hove singularities in phonon spectrum (Ashcroft and Mermin 1976), isotropic approximation of deformation potentials, and so on. Also the values are sensitive to various parameters such as the mean free path or core radius. Nevertheless, in typical metals, the phonon drag $B_{ph}$ range is $30 \sim 80 \mu Pa \cdot s$ at room temperature and less than $0.1 \mu Pa \cdot s$ at very low temperatures around $10 \ K$, while for the electron drag $B_e$ the range is a few $\mu Pa \cdot s$ and expected to be temperature independent. Under strong magnetic fields at low temperature, macroscopic dislocation behavior can be highly sensitive to orientation relative to the field within accuracy of $1\%$ (McKrell and Galligan 2000). Except for special cases such as deformation under high strain rate, weak dependences of drag on dislocation velocity are usually neglected.

Examples of temperature dependence of each component of the drag coefficient can be found for the case of edge dislocation in copper (Hiratani and Nadgorny 2001), or in Molybdenum (Jinpeng, Bulatov et al. 1999). Generally, however, the dislocation mobility could be, among other things, a function of the angle between the Burgers vector and the dislocation line sense, especially at low temperatures. For example, Wasserbäch (1986) observed that at low deformation temperatures ($77$ to $195 \ K$) the dislocation structure in Ta single crystals consisted of primary and secondary screw dislocations and of tangles of dislocations of mixed characters, while at high temperatures ($295$ to $470 \ K$) the behavior was similar to that of fcc crystals. In the work of Mason and MacDonald (1971) they measured the mobility of dislocation of an unidentified type in Nb as $4.2 \times 10^4 \ (Pa \cdot s)^{-1}$ near room temperature. A smaller value of $3.3 \times 10^3 \ (Pa \cdot s)^{-1}$ was obtained by Uvabe and Weertman (1975) for the mobility of edge dislocation in Fe. The mobility for screw dislocations in Fe was found to be about a factor of two smaller than that of edge dislocations near room temperature. A theoretical model to explain this large difference in behavior is given in Hirth and Lothe (1982) and is based on the observation that in bcc materials the screw dislocation has a rather complex three-dimensional core structure, resulting in a high Peierls stress, leading to a relatively low mobility for screw dislocations while the mobility of mixed dislocations is very high.
8.2 Theoretical Fundamentals of the Method

8.2.3.2 Dislocation Collisions

When two dislocations collide, their response is dominated by their mutual interactions and becomes much less sensitive to the long-range elastic stress associated with external loads, boundary conditions, and all other dislocations present in the system. Depending on the shapes of the colliding dislocations, their approach trajectories and their Burgers vectors, two dislocations may form a dipole, or react to annihilate, or to combine to form a junction, or to intersect and form a jog. In the DD analysis, the dynamics of two colliding dislocations is determined by the mutual interaction force acting between them. In the case that the two dislocation segments are parallel (on the same plane and or intersecting planes) and have the same Burgers vector with opposite sign they would annihilate if the distance between them is equal to the core size. Otherwise, the colliding dislocations would align themselves to form a dipole, a jog or a junction depending on their relative position. A comprehensive review of short-range interaction rules can be found in Rhee, Zbib et al. (1998).

8.2.3.3 Discretization of Dislocation Equation of Motion

Equation (8.1) applies to every infinitesimal length along the dislocation line. In order to solve this equation for any arbitrary shape, the dislocation curve may be discretized into a set of dislocation segments as illustrated in Figure 8.1. Then the velocity vector field over each segment may be assumed to be linear and, therefore, the problem is reduced to finding the velocity of the nodes connecting these segments. There are many numerical techniques to solve such a problem. Consider, for example, a straight dislocation segment $s$ bounded by two nodes $j$ and $j + 1$ as depicted in Figure 8.1. Then within the finite element formulation (Bathe 1982), the velocity vector field is assumed to be linear over the dislocation segment length. This linear vector field $v$ can be expressed in terms of the velocities of the nodes such that $v = [N^D]^T v^D$ where $v^D$ is the nodal velocity vector and $[N^D]$ is the linear shape function vector (Bathe 1982). Upon using the Galerkin method, Equation (8.1) for each segment can be reduced to a set of six equations for the two discrete nodes (each node has three degrees of freedom). The result can be written in the following matrix-vector form.

$$[M^D] \ddot{V}^D + [C^D] V^D = F^D$$  (8.3)

where

$$[M^D] = m_s \int [N^D][N^D]^T \, dl$$ is the dislocation segment $6 \times 6$ mass matrix,

$$[C^D] = (1/M_s) \int [N^D][N^D]^T \, dl$$ is the dislocation segment $6 \times 6$-damping matrix, and

$$F^D = \int [N^D] F_s \, dl$$ is the $6 \times 1$ nodal force vector.

The integration is performed over the dislocation segment length $l$. Then, following the standard element assemble procedure, one obtains a set of discrete system of equations, which can be cast in terms of a global dislocation mass matrix, global dislocation damping matrix, and global dislocation force vector. In the case of one dislocation loop and with ordered numbering of the nodes around the loop, it can be easily shown that the global matrices are banded with half-bandwidth equal to one. However, when the system contains many loops
that interact among themselves and new nodes are generated and/or annihilated continuously, the numbering of the nodes becomes random and the matrices become un-banded. To simplify the computational effort, one can employ the lumped matrix method. In this method, the mass matrix \([M^D]\) and damping matrix \([C^D]\) become diagonal matrices (half-bandwidth equal to zero), and therefore the only coupling between the equations is through the nodal force vector \(F^D\). The computation of each component of the force vector is described below.

### 8.2.4 The Dislocation Stress and Force Fields

The stress induced by any arbitrary dislocation loop at an arbitrary field point \(p\) can be computed by the Peach-Koehler integral equation given in Hirth and Lothe (1982). This integral equation, in turn, can be evaluated numerically over many loops of dislocations by discretizing each loop into a series of line segments. If we denote,

\[
\begin{align*}
N_l &= \text{total number of dislocation loops} \\
n_s^{(l)} &= \text{number of segments of loop } l \\
n_n^{(l)} &= \text{number of nodes associated with the segments of loop } n_n^{(l)} = n_s^{(l)} + 1 \\
N_s &= \text{total number of segments } = n_s^{(l)} \times N_l \\
N_n &= \text{total number of nodes } = n_n^{(l)} \times N_l \\
l_s &= \text{length of segment } s \\
r &= \text{distance from point } p \text{ to the segment } s
\end{align*}
\]

Then the discretized form of the Peach-Koehler integral equation for the dislocation stress \(\sigma_{ij}^d\) at any arbitrary field point \(p\) becomes

\[
\begin{align*}
\sigma_{ij}^d(p) &= \sum_{l=1}^{N_l} \sum_{s=1}^{n_s^{(l)}} \left\{ -\frac{G}{8\pi} \int_{l_s} b_p \epsilon_{mpq} \frac{\partial}{\partial x_m'} \nabla^2 R \, dx_j' \\
 & \quad \quad - \frac{G}{8\pi} \int_{l_s} b_p \epsilon_{mpq} \frac{\partial}{\partial x_m'} \nabla^2 R \, dx_j' \\
 & \quad \quad - \frac{G}{4\pi(1-\nu)} \int_{l_s} b_p \epsilon_{mpk} \left( \frac{\partial^3 R}{\partial x_m' \partial x_n' \partial x_j'} - \delta_{ij} \frac{\partial}{\partial x_m'} \nabla^2 R \right) \, dx_k' \right\}
\end{align*}
\] (8.4)

where \(\epsilon_{ijk}\) is the permutation symbol, \(R\) is the distance from a point on the dislocation line to the field point \(p\) as depicted in Figure 8.1. The first summation in (8.4) is carried over total number of dislocation loops \(N_l\) and the second summation is over the total number of segments \(n_s^{(l)}\) in each loop. The integral over each segment can be explicitly carried out using the linear element approximation. Exact solution of Equation (8.4) for a straight dislocation segment can be found in DeWit (1960) and Hirth and Lothe (1982). However, evaluation of the above integral requires careful consideration as the integrand becomes singular in cases where point \(p\) coincides with one of the nodes of the segment that integration is taken over, i.e., self-segment integration. Thus,
8.2 Theoretical Fundamentals of the Method

- if $p$ is not part of the segment $s$, there is no singularity since $r \neq 0$ and the ordinary integration procedure may be performed, or
- if $p$ coincides with a node of the segment $s$ where the integration should be carried out, special treatment is required due to the singular nature of the stress field as $r \to 0$. Here, the regularization scheme developed by Zbib and co-workers have been employed. In this scheme, the dislocation stress is decomposed into the following form.

$$d \sigma(p) = \sum_{s=1}^{N_s-2} d \sigma^{(s)} + d \sigma^{(p+)} + d \sigma^{(p-)} \tag{8.5}$$

where $d \sigma^{(s)}$ is the contribution to the stress at point $p$ from any segment $s$ except the two that are adjacent to point $p$. The contribution of these two segments is denoted by $d \sigma^{(p+)}$ and $d \sigma^{(p-)}$. (Here $p^+$ and $p^-$ indicate the segment located to the left and right of the field point $p$ when $p$ is situated at the dislocation line).

Once the dislocation stress field is computed the forces on each dislocation segment can be calculated by summing the stresses along the length of the segment. The stresses are categorized into those coming from the dislocations as formulated above and also from any other external applied stresses plus the internal friction (if any) and the stresses induced by any other defects. A model for the osmotic force $F_{\text{oosomal}}$ is given in Raabe (1998) and its inclusion in the total force is straightforward since it is a deterministic force. However, the treatment of the thermal force $F_{\text{thermal}}$ is not trivial since this force is stochastic in nature, requiring a special consideration and algorithm leading to what is called Stochastic Dislocation Dynamics (SDD) as developed by Hiratani and Zbib (2002). Therefore, the force acting on each segment can be written as:

$$F_s = \left( \sum_{m=1}^{N_s} d \sigma^{(m)} + \sigma^{(s)} + \tau_s \right) \cdot b_s \times \xi_s = d \sigma^{(s)} + \tau_s + F_{\text{thermal}} \tag{8.6}$$

where $d \sigma^{(m)}$ is the contribution to the stresses along segment $m$ from all the dislocations (dislocation-dislocation interaction), $d \sigma^{(s)}$ is the sum of all externally applied stresses, internal friction (if any) and the stresses induced by any other defects, and $\tau_s$ is the thermal stress; $d \sigma^{(s)}$, $\tau_s$ and $F_{\text{thermal}}$ are the corresponding total Peach-Koehler (PK) forces.

Using Equations (8.5), the force $d \sigma^{(s)}$ can also be decomposed into two parts one arising from all dislocation segments and one from the self-segment, which is better known as the self-force, that is,

$$d \sigma^{(s)} = \sum_{m=1}^{N_s-2} d \sigma^{(m)} + d \sigma^{(s)\text{(self)}} \tag{8.7}$$

where $d \sigma^{(m)}$ and $d \sigma^{(s)\text{(self)}}$ are respectively, the contribution to the force on segment $s$ from segment $m$ and the self-force. In order to evaluate the self-force, a special numerical treatment as given by Zbib, Rhee et al. (1996) and Zbib and Diaz de la Rubia (2002) should be used.
in which exact expressions for the self-force are given. This approximation works well in terms of accuracy and numerical convergence for segment lengths as small as 20 \( b \). For finer segments, however, one can use a more accurate approximation as suggested by Scattergood and Bacon (1975). Another treatment has been given by Gavazza and Barnett (1976) and used in the recent work of Ghoniem and Sun (1999).

The direct computation of the dislocation forces discussed above requires the use of a very fine mesh, especially when dealing with problems involving dislocation-defect interaction. As a rule to capture the effect of the very small defects, the dislocation segment size must be comparable to the size of the defect. Alternatively, one can use large dislocation segments compared to the smallest defect size, provided that the force interaction is computed over a many points (Gauss points) over the segment length. In this case, the self-force of segment \( s \) would be evaluated first. Then the force contribution from other dislocations and defects is calculated by computing the stresses at several Gauss points along the length of the segments. The summation as in (8.6) would then follow according to:

\[
F_s = F_{self}^s + \sum_{m=1}^{N_s-2} \frac{1}{n_g} \sum_{g=1}^{n_g} \left( d^{(m)}(p_g) + \cdots \right) \cdot b_s \times \xi_s
\]

where \( p_g \) is the position of the Gauss point \( g \) and \( n_g \) is the number of Gauss points along segment \( s \). The number of Gauss points depends on the length of the segment. As a rule, the shortest distance between two Gauss points should be larger or equal to \( 2r_o \), i.e. twice the core size.

### 8.2.5 The Stochastic Force and Cross-slip

Thermal fluctuations arise from dissipation mechanism due to collision of dislocations with surrounding particles, such as phonons or electrons. Rapid collisions and momentum transfers result in random forces on dislocations. These stochastic collisions, in turn, can be regarded as time-independent noises of thermal forces acting on the dislocations. Suppose the exertion of thermal forces follows a Gaussian distribution. Then, thermal fluctuations most likely result in very small net forces due to mutual cancellations. However, they sometimes become large and may cause diffusive dislocation motion or thermal activation events such as overcoming obstacle barriers. Therefore, the DD simulation model should also account not only for deterministic effects but also for stochastic forces; leading to a model called “stochastic discrete dislocation dynamics” (SDD) (Hiratani and Zbib 2002). The procedure is to include the stochastic force \( F_{\text{thermal}} \) in the DD model by computing the magnitude of the stress pulse \( (\tau_s) \) using a Monte Carlo type analysis.

Based on the assumption of the Gaussian process, the thermal stress pulse has zero mean and no correlation (Röpppel, Streit et al. 1993; Raabe 1998) between any two different times. This leads to the average peak height given as (Koppenaal and Kuhlmann-Wilsdorf 1964; Hiratani and Zbib 2002)

\[
\sigma_s = \sqrt{2kT/M_s b^2 \Delta l \Delta t}
\]

where \( k \) denotes Boltzmann constant, \( T \) absolute temperature of the system, \( b \) the magnitude of Burgers vector, \( \Delta t \) time step, and \( \Delta l \) is the dislocation segment length. Some values of
the peak height are shown in Table 8.1 for typical combinations of parameters. Here, $\Delta t$ is chosen to be 50 fs, roughly the inverse of the Debye frequency. Although $\Delta l$ or $\Delta t$ does not have a fixed value, such a restriction is imposed so that the system at the energetic global minima should reach thermal equilibrium. The validity of these parameters are checked by measuring the assigned system temperature ($T$) with the kinetic temperature of dislocation of which both ends are fixed, which should coincide on average according to the equipartition law. Although there is no unique way for choosing $\Delta t$ or $\Delta l$ in Equation (8.9), the justification can be checked if the segment velocity distribution is correct Maxwellian and if the average kinetic energy of the segment is equivalent to $kT/2$ for each degree of freedom at thermal equilibrium (Hiratani and Zbib 2002). The size of $\Delta l$ may also depend on the microstructure being considered, for example when the size of the local obstacle is in the order of nm, $\Delta l$ is also chosen to be in the same order.

Table 8.1: The stress pulse peak height for various combinations of parameters, $\Delta t = 50$ fs.

<table>
<thead>
<tr>
<th>$T$(K)</th>
<th>$1/M$ (µ Pa · s)</th>
<th>$\sigma_s$ (MPa)</th>
<th>$\sigma_s$ (MPa)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>($\Delta l = 5b$)</td>
<td>($\Delta l = 10b$)</td>
<td>($\Delta l = 10b$)</td>
</tr>
<tr>
<td>0</td>
<td>2</td>
<td>11.5</td>
<td>8.11</td>
</tr>
<tr>
<td>50</td>
<td>5</td>
<td>40.6</td>
<td>28.7</td>
</tr>
<tr>
<td>100</td>
<td>10</td>
<td>81.1</td>
<td>57.4</td>
</tr>
<tr>
<td>300</td>
<td>30</td>
<td>256</td>
<td>181</td>
</tr>
</tbody>
</table>

Numerical implementation includes an algorithm where stochastic components are evaluated at each time step of which strengths are correlated and sampled from a bivariate Gaussian distribution (Allen and Tildesley 1987)$^1$. In the stationary case, distribution of the dislocation segment velocity becomes also a Gaussian with variance of $kT/\mu^* \Delta l$. Under a constant stress without obstacles, Maxwell distributions around a mean velocity, which is equal to final viscous velocity $b \sigma_M$, will be realized as a steady state.

With the inclusion of stochastic forces in DD analysis, one can treat cross-slip (a thermally activated process) in a direct manner, since the duration of waiting time and thermal agitations are naturally included in the stochastic process. For example, for the cross-slip in fcc model one can develop a model based on the Escaig-Friedel (EF) mechanism where cross-slip of a screw dislocation segment may be initiated by an immediate dissociation and expansion of Shockley partials. This EF mechanism has been observed to have lower activation energy than Shockey-Seeger mechanism where the double super kinks are formed on the cross slip plane (this model is used for cross-slip in bcc (Rhee, Zbib et al. 1998)). In the EF mechanism, the activation enthalpy $\Delta G$ depends on the interval of the Shockley partials ($d$) and the resolved shear stress on the initial glide plane ($\sigma$). (See, for example, the MD simulation of Rasmussen and Jacobs (1997) and Rao, Parthasarathy et al. (1999)). The constriction interval $L$ is also dependent on $\sigma$. For example, for the case of copper, the activation energy for cross-slip can be computed using an empirical formula fitted to the MD results of Rao, Parthasarathy et al. (1999).

$^1$ Here we generate stress pulses as $\tau_s = \sigma_s \sqrt{-2 \ln r_1 \cos (2\pi r_2)}$ where $r_1$ and $r_2$ are uniform random numbers between zero and unity (Allen, M. P. and Tildesley, D. J., 1987).
Figure 8.2 depicts the $\Delta G (\sigma)$ for the case of copper where the value of the activation free energy is 1.2 eV, and for stacking fault energy is equal to 0.045 J/m². This activation energy for stress assisted cross-slip is entered as an input data into the DD code. Usually, within the DD code, dislocations are represented as perfect dislocations while a pair of parallel Shockley partials are introduced in the case of screw dislocation segments only for stress calculation. Then a Monte Carlo type procedure is used to select either the initial plane or the cross slip plane according to the activation enthalpy (Rhee, Zbib et al. 1998). For simplicity, one can set the regime of the barrier with area of $L \times d$ and strength of $\Delta G/Ld$. The virtual Shockley partials move according to the Langevin forces in addition to the systematic forces until the partials overcome the barrier and the interval decreases to the core distance. The implementation of this model captures the anisotropic response of cross-slip activation process to the loading direction, and consideration of the time duration (waiting time) during the cross-slip event, which have been missing in the former DD simulations.

### 8.2.6 Modifications for Long-Range Interactions: The Super-Dislocation Principle

Inclusion of the interaction among all the dislocation loops present in a large body is computationally expensive since the number of computations per step would be proportional to $N_s^2$ where $N_s$ is the number of dislocation segments. A numerical compromise technique termed the super-dislocation method, which is based on the multipolar expansion method (Wang and LeSar 1995; Hirth, Rhee et al. 1996; Zbib, Rhee et al. 1996), reduces the order of computation to $N_s \log N_s$ with a high accuracy. In this approach, the dislocations far away from the point of interest are grouped together into a set of equivalent monopoles and dipoles. In the numerical implementation of the DD model, one would divide the 3D computational domain into sub-domains, and the dislocations in each sub-domain (if there are any) are grouped together in terms of monopoles, dipoles, etc. (depending on the desired accuracy) and the their far stress field is then computed.
8.2 Theoretical Fundamentals of the Method

8.2.7 Evaluation of Plastic Strains

The motion of each dislocation segment gives rise to plastic distortion, which is related to the macroscopic plastic strain rate tensor $\dot{\varepsilon}^p$ and the plastic spin tensor $W^p$ via the relations

\[
\dot{\varepsilon}^p = \sum_{s=1}^{N_s} \frac{l_s v_{gs}}{2V} (n_s \otimes b_s + b_s \otimes n_s) \tag{8.10a}
\]

\[
W^p = \sum_{s=1}^{N_s} \frac{l_s v_{gs}}{2V} (n_s \otimes b_s - b_s \otimes n_s) \tag{8.10b}
\]

where $n_s$ is a unit normal to the slip plane, the symbol “$\otimes$” is the dyadic operator, $v_{gs}, b_s$ and $l_s$ are the magnitude of the glide velocity, Burgers vector and length of the dislocation segment $s$, respectively, $V$ is the volume of the representative volume element and $N_s = N_l \times n_s^{(l)}$ is the total number of dislocation segments. The above relations provide the most rigorous connection between the dislocation motion (the fundamental mechanism of plastic deformation in crystalline materials) and the macroscopic plastic strain, with its dependence on strength and applied stress being explicitly embedded in the calculation of the velocity of each dislocation. Nonlocal effects are explicitly included into the calculation through long-range interactions. Another microstructure quantity, the dislocation density tensor $\alpha$, can also be calculated according to

\[
\alpha = \sum_{s=1}^{N_s} \frac{l_s}{V} b_s \otimes \xi_s \tag{8.11}
\]

This quantity provides a direct measure for the net Burgers vector that gives rise to strain gradient relief (bending of crystal) (Shizawa and Zbib 1999).

8.2.8 The DD Numerical Solution: An Implicit-Explicit Integration Scheme

An implicit algorithm to solve the equation of motion (8.3) with a backward integration scheme may be used, yielding the recurrence equation

\[
v^{t+\delta t} = v^t + \Delta t \left( \frac{1}{m_s M_s} \right)^{t+\delta t} n_s b_s + \frac{\Delta t}{m_s} \varepsilon^{t+\delta t} . \tag{8.12}
\]

This integration scheme is unconditionally stable for any time step size $\delta t$. However, the DD time step is determined by two factors: i) the shortest flight distance for short-range interactions, and ii) the time step used in the dynamic finite element modeling to be described later. This scheme is adopted since the time step in the DD analysis (for high strain rates) is of the same order of magnitude of the time required for a stable explicit finite element (FE) dynamic analysis. Thus, in order to ensure convergence and stable solution, the critical time $t_c$ and the time step for both the DD and the FE ought to be $t_c = l_c/C_l$, and $\Delta t = t_c/20$, respectively, where $l_c$ is the characteristic length scale which is the shortest dimension in the finite element mesh.
In summary, the system of equations given in Section 8.2 summarizes the basic ingredients that a dislocation dynamics simulation model should include. There are a number of variations in the manner in which the dislocation curves may be discretized, for example zero order element (pure screw and pure edge), first order element (or piecewise linear segment with mixed character), or higher order nonlinear elements but this is purely a numerical issue. Nonetheless, the DD model should have the minimum number of parameters and, hopefully, all of them should contain basic physical and material parameters and not phenomenological ones for the DD result to be predictive. The DD model described above has the following set of physical and material parameters:

- Burgers vectors,
- elastic properties,
- core size (equal to one Burgers vector),
- thermal conductivity and specific heat,
- mass density,
- stacking fault energy, and
- dislocation mobility.

Also there are two numerical parameters: the segment length (minimum segment length can’t be less that three times the core size) and the time step (as discussed in Section 8.2.8), but both are fixed to ensure convergence of the result. In the above list, it is emphasized that in general the dislocation mobility is an intrinsic material property that reflects the local drag mechanisms as discussed above. One can use an “effective” mobility that accounts for additional drag from dislocation-point defect interaction, and thermal activation processes if the defects/obstacles are not explicitly impeded in the DD simulations. However, there is no reason not to include these effects explicitly in the DD simulations (as done in the model described above), i.e. dislocation defect interaction, stochastic processes and inertia effects, which actually permits the prediction of the “effective” mobility from the DD analysis (Hiratani and Zbib 2002; Hiratani, Zbib et al. 2002).

### 8.3 Integration of DD and Continuum Plasticity

#### 8.3.1 Continuum Elasto-Viscoplasticity

The discrete dislocation model can be coupled with continuum elasto-visoplasticity models, making it possible to correct for dislocation image stress and to address a wide range of complex boundary value problems at the microscopic level. In the following, a brief description of this coupling is provided. The coupling is based on a framework in which the material obeys the basic laws of continuum mechanics, i.e. the linear momentum balance.

\[
\text{div } \sigma = \rho \dot{\mathbf{v}} \tag{8.13a}
\]
and the energy equation

\[ \varrho c_v \dot{\theta} = k \nabla^2 T + \boldsymbol{\sigma} \cdot \dot{\varepsilon}^p \]  

(8.13b)

where \( v = \dot{u} \) is the particle velocity, \( u \), \( \varrho \), \( c_v \) and \( k \) are the displacement vector field, mass density, specific heat and thermal conductivity respectively. In DD the representative volume cell analyzed can be further discretized into sub-cells or finite elements, each representing a representative volume element (RVE). Then the internal stresses field induced by the dislocations (and other defects) \( \sigma^D \) and the plastic strain field within each RVE can be calculated at any point within each element. However, and in order to be consistent with the definition of a RVE, the heterogeneous internal stress field can be homogenized over each RVE, resulting into an equivalent internal stress \( S^D \) which is homogenous within the RVE, i.e.

\[ S^D = \langle \sigma^D \rangle = \frac{1}{V_{\text{element}}} \int_{\text{element}} \sigma^D(x) \, dv \]  

(8.14)

Furthermore, the plastic strain increment results from only the mobile dislocations that exit the RVE (or sub-cell) and is computed as in Equation (8.10) but with \( V \) being the volume of the element (or sub-cell). The dislocations that are immobile (zero velocity) do not contribute to the plastic strain increment. Therefore, the dislocations within an element induce an internal stress due to their elastic distortion. When some of these dislocations (or all) move and exit the element they leave behind plastic distortion in the element, and the internal stress field should be recomputed by summing the stress from the remaining dislocations in the element. With this homogenization procedure Hooke’s law for the RVE becomes

\[ \sigma + S^D = [C^e] [\varepsilon - \varepsilon^p] \]  

(8.15)

where \( C^e \) is the fourth-order elastic tensor. In the continuum plasticity theory one would need to develop a phenomenological constitutive law for plastic stress-strain. Here, this ambiguity is resolved by using the explicit expressions given by Equation (8.10a) for the plastic strain tensor as computed in the dislocation dynamics.

### 8.3.2 Modifications for Finite Domains

The solution for the stress field of a dislocation segment (Hirth and Lothe 1982) is true for a dislocation in an infinite domain and for homogeneous materials. In order to account for finite domain boundary conditions, Van der Giessen and Needleman (1995) developed a 2D model based on the principle of superposition. The method has been extended by Yasin, Zbib et al. (2001) and Zbib and Diaz de la Rubia (2002) to three-dimensional problems involving free surfaces and interfaces as summarized below.

#### 8.3.2.1 Interactions with External Free Surfaces

In the superposition principle, the two solutions from the infinite domain and finite domain are superimposed. Assuming that the dislocation loops and any other internal defects with self induced stress are situated in the finite domain \( V \) bounded by the surface \( S \) and subjected
to arbitrary external tractions and constraints. Then the stress, displacement, and strain fields are given by the superposition of the solutions for the infinite domain and the actual domain subjected to

\[ \sigma = \sigma^\infty + \sigma^*, \quad u = u^\infty + u^*, \quad \varepsilon = \varepsilon^\infty + \varepsilon^* \]  

(8.16)

where \( \sigma^\infty \), \( \varepsilon^\infty \) and \( u^\infty \) are the fields caused by the internal defects as if they were in an infinite domain, whereas \( \sigma^* \), \( \varepsilon^* \) and \( u^* \) are the field solutions corresponding to the auxiliary problem satisfying the following boundary conditions

\[ t = t^a - t^\infty \quad \text{on } S \\
 u = u^a \quad \text{on part of the boundary } S \]  

(8.17)

where \( t^a \) is the externally applied traction, and \( t^\infty \) is the traction induced on \( S \) by the defects (dislocations) in the infinite domain problem. The traction \( -t^\infty = \sigma \cdot n \) on the surface boundary \( S \) results into an image stress field which is superimposed onto the dislocations segments and, thus, accounting for surface-dislocation interaction.

The treatment discussed above considers interaction between dislocations and external free surfaces, as well as internal free surfaces such as voids. Internal surfaces such as micro-cracks and rigid surfaces around fibers are treated within the dislocation theory framework, whereby each surface is modeled as a pile-up of infinitesimal dislocation loops (Demir, et al. 1993; Demir and Zbib 2001). Hence, defects of these types may be represented as dislocation segments and loops, and their interaction with external free surfaces follows the method discussed above. This subject has been addressed by Khraishi, et al. (2001).

### 8.3.2.2 Interactions with Interfaces

The framework described above for dislocations in homogenous materials can be implemented into a finite element code. The model can also be extended to the case of dislocations in heterogeneous materials using the concept of superposition as outlined by Zbib and Diaz de la Rubia (2002). For bi-materials, suppose that domain \( V \) is divided into two sub-domains \( V_1 \) and \( V_2 \) with domain \( V_1 \) containing a set of dislocations. The stress field induced by the dislocations and any externally applied stresses in both domains can be constructed in terms of two solutions such that

\[ \sigma = \sigma^{\infty_1} + \sigma^*, \quad \varepsilon = \varepsilon^{\infty_1} + \varepsilon^* \]  

(8.18)

where \( \sigma^{\infty_1} \) and \( \varepsilon^{\infty_1} \) are the stress and strain fields, respectively, induced by the dislocations (the infinite solution) with the entire domain \( V \) having the same material properties of domain \( V_1 \) (*homogenous solution*). Applying Hooke’s law for each of the sub-domains, and using (8.18), one obtains the elastic constitutive equations for each of the materials in each of the sub-domains as:

\[ \sigma^* = [C_{11}^e] \varepsilon^*, \quad \sigma^* = [C_{21}^e] \varepsilon^* + \sigma^{\infty_2}, \quad \sigma^{\infty_2} = [C_{22}^e - C_{11}^e] \varepsilon^{\infty_1}, \quad \text{in } V_1 \]  

\[ \sigma^* = [C_{12}^e] \varepsilon^* + \sigma^{\infty_1}, \quad \text{in } V_2 \]  

(8.19)
8.4 Typical Fields of Applications and Examples

where $C^e_1$ and $C^e_2$ are the elastic stiffness tensors in $V_1$ and $V_2$, respectively. The boundary conditions are:

$$t = t^a - t^\infty_1 \text{ on } S; \quad u = u^a \text{ on part of } S \quad (8.20)$$

where $t^a$ is the externally applied traction and $t^\infty_1$ is the traction induced on all of $S$ by the dislocations in $V_1$ in the infinite-homogenous domain problem. The “eigenstress” $\sigma^{\infty 21}$ is due to the difference in material properties. The method described above can be extended to the case of heterogeneous materials with $N$ sub-domains (Zbib and Diaz de la Rubia 2002).

The above system of Equations (8.13)–(8.15) can be combined with the boundary corrections given by (8.16)–(8.20). The resulting set of field equations can be solved numerically using the finite element method as described by Zbib and Diaz de la Rubia (2002). The end result is a model, which they call a multiscale discrete dislocation dynamics plasticity (MDDP), coupling continuum elasto-visoplasticity with discrete dislocation dynamics. The MDDP consists of two main modules, the DD module and the continuum finite element module. The DD module computes the dynamics of the dislocations, the plastic strain field they produce, and the corresponding internal stresses field. These field values are passed to the continuum finite element module, in which the stress-displacement-temperature field is computed based on the boundary value problem at hand. The resulting stress field, in turn, is passed to the DD module and the cycle is repeated.

8.4 Typical Fields of Applications and Examples

Over the past decade, the discrete dislocation dynamics has been utilized by a number of researchers to investigate many complicated small-scale crystal plasticity phenomena that occur under a wide range of loading and boundary conditions (see for example a recent review by Bulatov et al. (2001)), and covering a wide spectrum of strain rates. Some of the major phenomena that have been addressed include:

- Dislocation pattern formation during monotonic and cyclic loading.
- Dislocation-defect interaction problems, including dislocation-void interaction (Ghoniem and Sun 1999), dislocation-SFT/void-clusters interaction in irradiated materials and the role of dislocation mechanisms on the formation of localized shear bands (Diaz de la Rubia, Zbib et al. 2000; Khraishi, Zbib et al. 2002).
- Effect of particle size on hardening in metal-matrix composites (Khraishi and Zbib 2002).
- Crack tip plasticity and dislocation-crack interaction (Van der Giessen and Needleman 2002).
- The role of various dislocation patterns (Raabe et al. 1996) and geometrically necessary boundaries (GNB’s) in hardening phenomena (Khan, Zbib et al. 2001).
• Plastic zone and hardening in Nano-indentation tests (Fivel, Roberston et al. 1998).

• The role of dislocation mechanisms in increased strength in nano-layered structures (Zbib and Diaz de la Rubia 2002; Schwartz 2003).

• High Strain Rate Phenomena and shock wave interaction with dislocations (Shehadeh et al. 2003; Zbib et al. 2003).

In what follows, and in order to illustrate the utility of this approach in investigating a wide range of small-scale plasticity phenomena, representative results for a set of case studies are presented. This includes, dislocation behavior during monotonic loading, the evolution of deformation and dislocation structure during loading of a cracked specimen, and dislocation interaction with shockwaves during impact loading conditions2. The results given below are for both copper and molybdenum single crystals whose materials properties are as follows. For copper, mass density \( \rho = 8900 \) kg/m\(^3\), \( G = 54.6 \) GPa, \( \nu = 0.324 \), \( b = 0.256 \) nm, \( M = 10^4 \) Pa·s. For molybdenum, mass density \( \rho = 10200 \) kg/m\(^3\), \( G = 123 \) GPa, \( \nu = 0.305 \), \( b = 0.2725 \) nm, \( M_{\text{mixed}} = 10^3 \) Pa·s, \( M_{\text{screw}} = 0.1 \) Pa·s.

### 8.4.1 Evolution of Dislocation Structure during Monotonic Loading

A classical problem in crystal plasticity is the nature of strain hardening behavior and the underlying dislocation microstructure evolution during both loading and stress relaxation. Both patterning and the concomitant strain hardening are thought to result from attractive non-planar dislocation interactions that lead to formation of sessile junctions and intersections. The first simulation result shown is that of deformation of a single crystal with periodic boundary conditions under monotonic loading with low strain rate. For this case, one may constructs a simulation cell with either reflected boundary conditions (Hirth, Rhee et al. 1996; Zbib, Rhee et al. 1996) or periodic boundary conditions (Bulatov, Rhee et al. 2000), maintaining dislocation continuity and flux across the cell boundaries. In either case, the long-range stress field is computed using the super-dislocation method described in Section 8.2.6. Figure 8.3a shows a cube unit cell whose size is 10 \( \mu \)m with initial screw dislocations (with jogs and kinks) that are distributed randomly in the crystal. The crystal considered is molybdenum with randomly distributed initial dislocation density on the \(<111\>{\{011}\} \) systems. The load is applied in the \([2\,9\,20]\) direction for reasons described in (Lassila, LeBlanc et al. 2002). A constant strain rate of \( 1/s \) is imposed in that direction. At this strain rate the DD simulation is performed with the time step varied between \( 10^{-8} \) s to \( 10^{-6} \) s. The dislocation velocity and the shortest distance between two dislocations control the time step. The stress field is assumed to be uniform throughout the cell, and therefore, the finite element part of the analysis is suppressed. The number of steps in this simulation was over one million (one processor on Dec Alpha workstation) to reach a strain of about 0.3 %. Typical results can be seen in Figure 8.3b showing the dislocation morphology at 0.3 % strain. From the data collected, one can extract various interesting information that could be useful in many ways. For example, by analyzing the spatial distribution of all dislocation segments, one can construct pair-distribution functions

---

2 The dynamic evolution of the dislocation structures presented in this paper can be best visualized by viewing the video clips available on the website http://www.cmm.wsu.edu/
as shown in Figure 8.4, for projections in various crystallographic directions, from which one can extract a wavelength and indication of a dislocation pattern. Further analysis also reveals that not only the $<111>|{011}$ systems are active but also some of the $<111>|{112}$ systems also become activated, resulting mainly from multiple cross-slip.
8.4.2 Dislocation Crack Interaction: Heterogeneous Deformation

Figures 8.5 and 8.6 show results of a mode I crack in a single crystal copper. The crack is located at the right side of the bottom of the specimen as can be deduced from the figures. The crack length is one third of specimen width. The bottom surface of the specimen (the un-cracked portion) and the left and right sides of the specimen are assumed to be symmetric boundaries. The crystal orientation is depicted in the figure with the x, y and z-axes being in the (110), (110) and (001) directions respectively. Initial dislocation loops (Frank-Read sources) are distributed randomly in the crystal on two slip planes: the (111) and (111) planes. The initial dislocation density is $10^{12}/m^2$. The upper surface of the specimen is displaced a constant distance such that the overall macroscopic strain is 1.67% (stress relaxation condition). This orientation induces initially double slip deformation, but as the simulation proceeds, some of the dislocations segment cross-slip and multi-slip deformation prevails. The DD simulations is performed in parallel with the finite element analysis which corrects for boundary tractions, image stresses from the surfaces of the crack, and computes the heterogeneous stress field.

Figures 8.5a depicts the distribution of shear stresses in the cracked specimen and after the dislocations attained the distribution shown in Figure 8.6a with the dislocation density having increased by about two orders of magnitudes to $7 \times 10^{13}/m^2$ (see Figure 8.6b). Figure 8.5b shows the corresponding plastic strain distribution. The stress field here does not have the exact same characters of the elastic stress field of a dislocation free cracked-specimen. This
is due to two factors, one is the internal stress field of the dislocations, and two is the plastic strain field induced by the mobile dislocations, both of these fields cause stress relaxation and reduce the order of the stress singularity at the crack tip. This can be deduced from Figure 8.5b where the plastic strain around the crack tip is about 1.5%. Close examination of Figure 8.6a
Figure 8.6: a) The dislocation structure around the crack tip. b) The evolution of the dislocation density.

reveals the formation of pile-ups of dislocation near a boundary separating the region with high stresses and a region with very low stress (almost insignificant). The dislocation piles up consist of [101] and [011] type dislocations extended in the (110) direction. The density of these pile ups is high enough that could lead to the nucleation of a microcrack that eventually
would link with the main crack. The coupled DD-FE analysis also predicts the evolution of the temperature field. In this simulation the rise in temperature at these strain levels was about 20 K in a region where the plastic strain is about 3.5% around the crack tip (Figure 8.5b).

8.4.3 Dislocations Interaction with Shock Waves

![Figure 8.7: Dislocation-Shockwave interaction. a) Computational setup geometry with initial dislocation distribution (dislocation loops, Frank-Read sources). b) The dislocation morphology in crystal shocked at strain rate $1 \times 10^6$ for 3.6 nanoseconds.]

In this final example, we give representative results pertaining to the deformation of a pure copper single crystal under extreme pressures and high strain rates (Shehadeh and Zbib 2002). The main objective of these simulations is to investigate the interaction of dislocations with stress waves under extreme strain rates up to $10^8/s$ and pressures ranging from a few GPa to tens GPa, under which the elastic properties are pressure-dependent. The simulation consists of a domain shown in Figure 8.7a. The cross-section is a square whose size 2.5 µm. The length of the cell in the z-direction is 25 µm. The cell is considered to be in an infinite domain with the four side surfaces having confined boundary conditions, i.e. the FE nodes can slide in the plane of their respective surface but not normal to it. This condition is assumed in order to keep the specimen under confinement to better represent shock wave experiments.
Figure 8.8: Dislocation-shockwave interaction. a) Effect of dislocations on pressure profile.
b) Effect of pulse period on the dislocation density.

(Meyers 1994; Kalantar et al. 2001). The upper surface is displaced (compressively) with a constant velocity over a small period of time (e.g. 3 ns) to correspond to an average strain rate of $10^6$/s, the surface is then released and the simulation continues long enough for the elastic waves to pass through the entire domain and dissipate at the lower boundary. During this process the dislocation sources, which are situated in the simulation cell as can be deduced from Figure 8.7a, interact with the stress wave and an avalanche of dislocations takes place the instant the wave impacts the sources. As the wave passes through the sources, dislocations continue to emit from the sources but eventually the rate of production diminishes and a dislocation structure with high density emerges as can be deduced from Figure 8.7b. Figure 8.8a
depicts the pressure profile calculated by the finite element analysis along the length of the specimen after the wave has traveled past the dislocations sources. The figure also shows the wave profile when there are no dislocations present in the crystal. It can be deduced from the figure that the dislocations cause reduction in the pressure peak resulting from plastic strain and internal stress relaxation. The effect of the length duration of the imposed pulse on the dislocation density is shown in Figure 8.8b. Finally, the FE analysis also provides information about the temperature field as can be seen in Figure 8.9a. The increase in temperature is due to dissipation of energy from dislocation drag. Figure 8.9b shows the history profile of the temperature in the region with the highest rate of energy dissipation.

Figure 8.9: Contour plot of temperature distribution in the computational cell, the local temperature increase reached value as high as 105°C, and a plot of the temperature history in the element of maximum temperature increase.

8.5 Summary and Concluding Remarks

The discrete computational approach of dislocation dynamics is a bridging model that is based on the fundamental carriers of plastic deformation – dislocations. The DD approach developed over the past decade has overcome many hurdles and established a useful framework for predictive simulations. In order to expand the range of engineering and scientific issues that can be addressed, the DD models should become more realistic and computationally efficient. A much-needed extension of the DD methodology is to anisotropic elasticity and a consistent treatment of local lattice rotations that become increasingly important with increasing strain. A completely new set of challenging problems arises when dealing with high strain rates phenomena. When it comes to putting dynamics into dislocation dynamics, relevant extensions should incorporate relativistic distortions of the stress field of the fast moving dislocations and various peculiar behaviors associated with the short-range interactions.
Despite the impressive recent developments in the DD methodology, it is important to avoid unrealistic expectations. DD models were intended to address crystal plasticity at the microscopic length scale. Stretching their computational limits is important but may not be the most constructive way to building a predictive framework for macroscale plasticity. This is because deformation processes at still larger (meso) scales are important, involving collective behavior of huge dislocation populations and making direct DD simulations computationally prohibitive. Further coarse-graining and homogenization are required so that the results of DD simulations at the microscale are used to inform a less detailed meso-scale model. Although no consistent framework of this kind currently exists, several recent developments in this area look promising. In addition to the reaction-diffusion approaches (Holt 1970; Walgraef and Aifantis 1985), variant forms of strain gradient plasticity theory have been advanced attempting to capture length scale effects associated with small-scale phenomena (Mindlin 1964; Dillon and Kratechvil 1970; Aifantis 1984; Zbib and Aifantis 1989; Fleck, Muller et al. 1994; Arsenlis and Parks 1999). Closely related are several recent attempts (Anthony and Azirhi 1995; Shizawa and Zbib 1999) that follow the original ideas of Kröner (1958) and describe the populations of crystal defects as continuous fields generating geometric distortions in the crystalline lattice. One way or another, computational prediction of crystal strength presents the over-arching goal for micro- and meso-scale material simulations.

Acknowledgement

The support of Lawrence Livermore National Laboratory to WSU is gratefully acknowledged. This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory (contract W-7405-Eng-48). Many thanks to Professor John P. Hirth for his advise and encouragement.

References


Kalantar, D. H, Allen, A.M., Gregori, F. et al., 2001. Laser driven high pressure, high strain-rate materials experi-
ments. 12th biennial International Conference of the APS Topical Group on Shock Compression of Condensed Matter, Atlanta, Georgia.


References


9 Physically Based Models for Industrial Materials: What For?

Yves Brechet

An overview of some recent development in modelling of microstructures and mechanical properties is provided, both from an academic point of view and from an industrial application perspective. Examples of different type of modelling, both analytical and computer simulations, are proposed, in relation with their possible applications in materials design. Fields of promising development are outlined in conclusion.

9.1 Introduction

It is generally thought that the sign of scientific maturity is the ability to model and predict, as well as to observe and classify. In that sense, materials science is definitely becoming a mature field. But materials science has a special status due to the close relation it bears with engineering applications. In that sense, maturity also presupposes the ability to provide guidelines for materials design. The purpose of this chapter is to illustrate from examples of modelling of materials, recent developments and trends in the field which are expected to lead to a better understanding of microstructure genesis and materials properties. Another claim of this chapter is that physically based models can provide guidance for the optimization of materials processing route and implementation in structures. In the present chapter, we are concerned with industrial materials, meaning materials which are designed to be applied, not for the investigation of a specific scientific problem. These materials are intrinsically complex, and the existing modelling in the literature have to be developed accordingly. In addition, the questions of interest in industrial materials also suggest new development from a fundamental viewpoint. The examples of this chapter mean to illustrate this constant interplay between application and fundamental research.

In §2, we will outline recent trends in the development of modelling in materials science in the last 10 years, insisting on the emergence of intensive computer simulations. In §3, we will illustrate these trends and their consequence on materials design procedure. In §4, we will state some personal view on directions in which more effort is needed in modelling for the coming years.

9.2 Recent Trends in Modelling Materials Behavior

In the present contribution, Materials are to be understood as structural materials, and the properties under consideration are mechanical properties. Although metals, polymers, ceram-
ics and composite materials are all present in this class of applications, the strongest effort in modelling the relation process/microstructures/properties, which is the core of materials science, has been developed in physical metallurgy. The reason for that situation is simply one of chronology, and it is expected in the coming years that a comparative approach between the different classes of materials will lead to a similar development in modelling for the other classes (the trend is already visible as far as mechanics is concerned).

The general strategy for modelling materials science problems has been outlined with great clarity in a paper by M. Ashby (1992): it relies on a constant interplay between order of magnitude calculations, simple analytical modelling, feedback from experiments, and, when required, the use of computer simulations at various levels of sophistication. The examples presented in §3 will illustrate this strategy.

9.2.1 Analytical Models

The basic paradigms for modelling in physical metallurgy are well developed:

- the theory of phase transformations allows us to describe, depending on the level of sophistication required, the volume fraction, or the shape and scales of the different phases resulting from simple heat treatments.
- the theory of dislocations allows us to understand the interaction of these defects with refined second phases and their mutual interaction.
- the theory of inclusions allows us to treat the problems of incompatibilities and load transfer present in coarse heterogeneous microstructures.
- the linear theory of fracture, and the theory of cavity growth by plasticity are the two entrance doors of failure mechanics for brittle and ductile materials.

These basic paradigms, applied in now classical works, have allowed the development of a “tool box” for modelling. An exhaustive list would be out of the scope of this paper, but we can list some examples of problems for which some relatively consensual models are available.

- nucleation, growth and coarsening of second phase precipitates: this aims at predicting the fine microstructure resulting from a precipitation heat treatment.
- theory of interface driven transformations: this aims at predicting the invasion rate of a grain boundary nucleated reaction such as pearlite, bainite, or discontinuous precipitation.
- theory of structural hardening: this aims at predicting the yield stress of a material knowing its state of precipitation and some characteristics of the precipitates.
- theory of work hardening: this aims at predicting the subsequent hardening post yielding associated either with dislocation accumulation, or with Polaris ed internal stresses development.
- theory of ductile fracture: this aims at relating fracture characteristics (toughness, ductility) with both microstructural features and loading conditions (with a special emphasis on triaxiality of the stress state).
This doesn’t mean that everything is solved in these problems: raging controversies on work hardening, unsettled questions on bainite reaction, shadows in the quantitative understanding of nucleation are in everybody’s mind to remind us of the existing pitfalls in our current understanding. However, some models exist and can be used as starting points to understand experimental facts. All these classical models have in common to simplify (and sometimes oversimplify) the assumptions: the inclusions are assumed to be ellipsoids, precipitates are considered as spherical, the interfaces are planar, dislocations are described by their density only. Moreover, the thermomechanical treatments and the loadings are simple (isothermal, uniaxial, etc.). One can always derive more sophisticated models, replace analytical solutions by FEM calculations, but the question of the “ratio quality to price” of a model is always a relevant one: in modelling, economy and efficiency are the rule.

Still, some annoying questions are shadowing the landscape, otherwise rather reassuring, that was outlined above. Some of them are of practical relevance, others are intellectual puzzles which might become of practical relevance, but have their interest for the simple sake of understanding.

Lets start by the questions of practical relevance. The quantitative understanding of nucleation kinetics is still under question: it is not clear that modelling precipitation can always be dealt with homogeneous nucleation theory, it is not known how to include into nucleation kinetics the sequence of poorly known metastable phases. The interface conditions at a transforming interface (local equilibrium, paraequilibrium, . . . ) are still selected a posteriori to describe the observed kinetics, and the coupling between solute drag, interface conditions, and the diffusion field ahead of the interface is not totally clear. Nucleation of recrystallization has long been known to be impossible to describe by classical nucleation theory, but in spite of this long history of non-understanding, we are still lacking a proper theory for the critical strain for recrystallization. Nucleation seems to be a fateful word in physical metallurgy: the nucleation step for deformation induced cavities is still a puzzle.

On the side of the puzzles “teasing the mind”, some of them are long standing ones: What are the fundamental reasons for dislocation patterning, and the relative importance of short range interactions, long range interaction and self screening associated with patterning? What are the mechanisms for pattern selection in phase transformations: are they stability conditions, noise induced transitions, optimal dissipation principles? Some of them are more recent: What are the length scales in mechanics and what could be the physical origin, if any, of the corrective terms in second gradient theories? What are the laws of plasticity in confined structures and what are the mechanisms of deformation for nanostructured materials?

All these questions are still unsolved, and however modelling has proceeded . . . are we facing a giant with clay feet, or simply applying the statement due to Jean Rostand “Waiting to have understood everything before acting amounts to condemn oneself to inaction”?

### 9.2.2 Computer Simulations

The examples listed above can be coined “classical models” in physical metallurgy. They can be found in good textbooks as Raabe (1998) and Phillips (2000) and an extensive review of the literature on these questions can be found in the series directed by Cahn et al. (1991), in particular the volumes Mughrabi (1991) and Kostorz (2001).
In recent years, computer simulations have entered the field with all the more strength that physical metallurgy was, in this respect, somewhat lagging behind physics, chemistry and mechanics. The reason for this relative delay in using computer simulations in physical metallurgy may be cultural, but there is a deeper reason for it: by nature, materials science is an exercise of changing scales, and, with the exception of irradiation damage and diffusion mechanisms (a field in which computer simulation as a tool emerged much before . . .), the relation between the atomistic level and the macroscopic behavior involves mechanisms at many different time and space scales. As a result, computer simulations emerged as an efficient tool in physical metallurgy when it became able to deal, at least partially, with this exercise of changing scales. Recent reviews can be found in Kirchner et al. (1996) and Finel et al. (2003). The developments in computer simulations applied to physical metallurgy problems can be grouped in three classes:

- At the level of atomistic simulations, after the pioneering works on irradiation damage and diffusion mechanisms, the modelling of order/disorder transitions and of diffusion controlled phase transitions has now reached a level of maturity which allows to make predictions which can be tested experimentally, such as the recent prediction of an intermediate phase in NbC precipitation from ferritic steels (Gendt, 2000). Structural defects, both in their structure, their dynamics, and their interactions, have also been successfully investigated by molecular dynamics (Rodary et al., 2003). Crack under loading have also been investigated (Gumbsch, 1999), bringing both some new insight into the ductile/brittle transition, but also pointing at the limits of atomistic simulations when complex far field boundary conditions are applied. The semicontinuum approach seemed to be a possible cure for these problems, but in spite of promising features, it remains today still unpractical (Rodney, 2003).

Another class of simulation, developed originally by Khachaturyan (for a review, see Wang et al. (2003)), takes advantage of using a lattice, but not necessarily of atomic scale, and allows for the description of microstructures involving long range terms such as elastic effects: the question of the noise introduced by brute force in the simulation is somewhat worrisome, but the simulated microstructures are strikingly similar to the ones observed experimentally.

- The basic understanding lies at the atomistic level (which relies on a proper derivation of interatomic potentials) but the core problems of materials science are at the mesoscopic scale, at the level of the collective behavior of structural defects. The number of entities to be considered here excludes for the moment the direct use of atomistic simulations. At this mesoscopic scale, simulations build on the known dynamics interactions between elementary entities. Some striking examples of this strategy in recent years are the 3D collective behavior of dislocations (Kubin, 1991), and the simulation of grain growth and recrystallization using either a Potts model (Anderson et al., 1984), or a vertex dynamics simulation (Weygand et al., 1999a). Since computer simulation can be relatively heavy, it is always worth asking what is the understanding they bring that an analytical model couldn’t provide. Lets give a few examples. The 3D dislocations simulation has allowed to identify the role of local cross slip on work hardening (Kubin, 1991), to simulate slip patterning associated with local softening (Brechet et al., 1996), to investigate the consequence of various anomaly mechanism in intermetallics (Devincre
Physically Based Models for Industrial Materials: What For?

et al., 1997), to simulate nanoindentation (Fivel et al., 1998), and is probably the tool needed to understand dislocation patterning (Kubin, 1991). Another version of discrete dislocation simulation, taking advantage of both a 2D situation computationally efficient, and of incorporating local rules mimicking the 3D behavior, has been recently developed and allows to deal with larger strains and complex boundary problems (Benzergera et al., 2004). The Vertex simulation of grain growth has allowed to investigate the possible role of energy and mobility anisotropy on grain growth (Weygand et al., 1998a), to quantify the influence of triple junctions restricted mobility (Weygand et al., 1998b), to test the various statistical summations in Zener Pinning (Weygand et al., 1998c), to investigate the difference between subgrain growth and grain growth (Weygand et al., 1999b), and is probably the tool needed to better understand the mechanisms of nucleation of recrystallization (Weygand et al., 2000). A third example of “mesoscale simulation” has been developed by the Lausanne team for solidification structures, the so called CAFE model (Rappaz and Gandin, 1993) relies on a cellular automata simulation of the solidification mechanism coupled with a FEM calculation of the thermal and solute field. This type of model have proven to be very useful in understanding the relation between a component geometry and its solidification microstructure, and is a true example of “scale transition”. In all these cases, intensive computer simulations were necessary to perform “thought experiments” necessary to identify the relevant approximations for an analytical treatment.

• When the collective nature of the phenomenon is not crucial, one may use “phenomenological simulations” which amount to solve numerically evolution equations on a limited number of “internal variables” describing the microstructure or its behavior. Phase field models in phase transformations (Karma, 2003), constitutive equations for plasticity (Es-trin, 1996), self consistent models for polycrystals plasticity (Kocks et al., 2000), FEM calculations on a representative cell, and the whole group of models corresponding to the so called “local approach to fracture”, fit in this category. This last class of simulation is less closely related to the elementary mechanisms familiar to the physicist, but they are closer to the answers required by an engineering approach. It is likely that the coming years will see a closer connection between the atomistic simulations providing the input to the mesoscale simulation, and the mesoscale simulations proving the local laws required to run the phenomenological simulations.

9.2.3 Materials Modelling and Materials Design: Some Examples

Before turning to the applications of materials modelling to materials design, it is worth to keep in mind the two driving forces for research in materials science. Pure science is certainly one. But since materials is matter with a function, the interaction between industrial questions and academic research is bound to be naturally closer in materials science than in many other sciences. The industrial push is both for increased properties (the search for ultrafine grain ferritic steels is one example), and for a decreased dispersion of properties. The first of these request speaks for a better understanding of the relation between process parameters, microstructures and properties. It drives the need for application of the physically based models on complex alloys and thermomechanical treatments. The second of these request speaks for
9.2 Recent Trends in Modelling Materials Behavior

...a better understanding of the role of the process parameters. It drives the need for a sensitivity analysis of any model developed on physical understanding of the basic phenomena.

The examples given above of recent development in modelling in physical metallurgy have their own value per se. But in order to be used as guide lines for materials design, they need to be adapted to industrial materials. This adaptation is not an easy task, so industrial problems may be dealt with in a number of ways. Empirical approach and statistical analysis of factory data is a possible route. Physically based models is another one, and can be integrated to deal with complex thermomechanical processes. A recent trend is now to have design driven models, in order to develop materials for a given set of requirements. Let's examine in sequence these four types of modelling.

9.2.4 Sophisticated Statistical Analysis

The use of statistics of properties from plants has always been a way to derive empirical relations between process parameters and materials properties. It is not intellectually very satisfactory since it provides no understanding of the underlying physical phenomena, but it is certainly very useful to capture existing correlations. A refined version of statistical analysis using neural networks have been developed in recent years, mainly in Cambridge (Gavard et al., 1996) and Delft (Van Der Volk et al., 1998). The principle is relatively simple. A number of input parameters \( x_i \) are supposed to influence a property \( y \). These parameters may be composition, heat treatments characteristics, rolling conditions. The property (for instance creep resistance) is correlated to the inputs via the function:

\[
y = \sum_i w_i^{(2)} \cdot \tanh \left( \sum_j w_{ij}^{(1)} x_j + \theta_i^{(1)} \right) + \theta^{(2)}
\]

The unknowns of the problem are the values of the constants \( w \) and \( \theta \) which capture the correlations. These values are adjusted in order to describe accurately (in the sense of a least square minimization) experimental results on a “training sample”. Such methods are now current in chemical engineering where process parameters are continuously measured. They have been applied to various problems in materials science such as fatigue strength of welds, recrystallization and grain structures in forged aluminum components, all these problems having in common the difficulty to derive a physically based model. The success of these methods is, at the same time, interesting (since no model physically based could describe so accurately the results), and have to be taken with caution, since it can only integrate correlations which are present in the “training sample”, which have to be sufficiently large in order for the correlation coefficients to be estimated with a sufficient accuracy. Note that this method has no physical content whatsoever. It might be possible to input extra information such as “the behavior is thermally activated and behaves as \( \exp(-Q/RT) \)”, or “the yield stress goes through a maximum as a function of the ageing time”: this would lead to a mixture between pure statistical analysis and physically based models (Ashby (2002) and ?). But the strength and the weaknesses of the methods are identical: very efficient to treat a large number of data, they can be trusted only if trained on a large number of data.
9.3 Some Examples of Physically Based Models for Industrial Materials

The examples selected here have all in common the following features:

- They are motivated by an industrial problem
- Modelling is a key issue in the development of better materials
- The basic concepts for modelling are present in the literature, but they have to be adapted to the industrial problem under consideration.

These examples do not pretend to give a complete overview, neither of the problems, nor of the techniques currently used in the literature, they are selected to illustrate how academic research can be useful on questions of industrial interest, and wish to give some motivation for replacing the “blind phenomenological approaches” by more physically based strategies.

The examples we will present here rely either on analytical approaches, or simple computer simulations. Extensive computer simulations at the atomistic level, as well as simulations on the collective behavior of defects will not be presented in this volume devoted to continuum modelling methods. The reason for this choice is that atomistic simulations are restricted to binary or at most ternary alloys, and the simulations for collective behavior of defects have been validated mainly on pure metals. That doesn’t mean they are useless for industrial materials, but it means that they have to be used “as it” without real adaptation to the industrial situation. Since our aim here is to illustrate the evolution from the academic systems toward industrial ones, we have chosen situations in which the evolution is more transparent.

The three examples described below illustrate three classical situations. In the first case, a phenomenological model is unable to be predictive and has to be refined with better physical basis. In the second case, two phenomena well described in the literature have to be coupled to understand complex thermomechanical treatments. In the last case, an exercise of “scale transitions” is presented which allows to integrate existing models from the solidification stage to the heat treatment stage, in order to optimize castings for their yield stress.

9.3.1 Recovery of Aluminum Alloys

Recovery of aluminum alloys is a key issue for beverage can making industry: the strength of the can body comes from cold work, and the processing of the varnish requires mild heating during which part of the strength is lost via recovery of the dislocation structure generated during the previous processes. A physically based model has to rely on microstructural mechanisms. There may be some “adjustable parameters”, but the value of these constants should have a sensible behavior when the experimental parameters are varied. Recovery of Aluminum Magnesium alloys will provide us with an example of such a “physically based model” (Verdier et al., 1999). Cold rolled Aluminum Magnesium alloys where recovered at various temperature. The observed kinetics where logarithmic in time and the density of dislocations only slightly decreased during recovery. This indicates that the basic idea of Friedel (1964) to interpret recovery as a relaxation of internal stresses associated with the dislocation
structure obtained during cold rolling is essentially correct. The simplest way to write this relaxation kinetics is:

\[
\frac{d\sigma}{dt} = -K \cdot \exp(-U^\circ + V\sigma/kT)
\] (9.2)

where \(\sigma\) is the flow stress, \(U^\circ\) and \(V\) are activation energy and activation volume, \(k\) is Boltzmann constant and \(T\) the temperature and \(K\) is an adjustable constant of the model giving a reference rate for recovery. This equation seems to capture most of the physics of the problem: a strain rate driven by the internal stress and an exponential dependence with this stress. This equations provides a good description of the recovery kinetics for a given temperature after a given rate of cold rolling. However when the model is applied to different initial cold rolling states the constant \(K\) required to provide a good description of the experimental data behaves erratically and changes by two orders of magnitude. The conclusion is that, in spite of a correct fit of the experimental results, the model is missing some physical point. Its not only disturbing from an intellectual point of view, it also prevents the application of such models to interpolated or extrapolated process parameters (temperatures, cold work reduction).

The cure of the problem is simple. The relaxation rate is indeed given by:

\[
\frac{d\sigma}{dt} = -E \dot{\varepsilon}
\] (9.3)

where \(E\) is the young modulus and \(\dot{\varepsilon}\) the plastic strain rate. In order to describe the strain rate, the density of dislocations has to be included, and since it can be experimentally proven that this density scales as the square of the flow stress, the new relaxation equation is:

\[
\frac{d\sigma}{dt} = - \left( \frac{64}{9M^2 \cdot \alpha^2} \right) \cdot \left( \frac{\sigma^2}{E} \right) \nu \cdot \exp \left( \frac{-U^\circ}{kT} \right) \cdot \text{sh} \left( \frac{\sigma V}{kT} \right)
\] (9.4)

\(M\) is a Taylor factor and \(\nu\) is the Debye frequency which sets the time scale. Note that there are still two adjustable parameters \(U^\circ\) and \(V\), that the fit of the experimental curves is as good as previously but that the values of \(V\) and \(U^\circ\) are in the correct range, and vary coherently with the initial rolling ratio. The moral of the story is that accurately describing experimental data is a necessary but not sufficient condition for a model to be reliable!

A recovered structure will work harden in a different manner than a non recovered structure (Verdier et al., 1999). The physical reason for that different behavior is that the subgrain size \(D\) is no longer related to the dislocation density \(\rho\) by the simple scaling law \(D = \rho^{-1/2}\). The rate of accumulation of dislocations has to take into account separately the mechanisms occurring inside the cell walls and inside the cell interior. As a result, one can write:

\[
\frac{d\rho}{d\varepsilon} = M \cdot (k_1 \cdot \sqrt{\rho} - k_2) + M \cdot \left( \frac{k}{D} - \frac{K}{D^\eta} \right)
\] (9.5)

The consequence of this equation is that the strain hardening coefficient \(\vartheta\) can be written:

\[
\vartheta = \vartheta_{II} + \frac{P_1}{\sigma} - P_2 \cdot \sigma
\] (9.6)

with a linear relation between the two constants \(P_1\) and \(P_2\). Moreover, this relation allows to compute a subgrain size which can be compared with experimental observations in TEM.
Both the work hardening behavior and the predicted subgrain size are in agreement with the experimental results. The moral of the story is that internal variable models should be, when it is possible, compared to microstructural observations.

Industrially, several alloys are used in can making. The previous model has been applied to all of them, and the value of the adjustable parameters $U^c$ and $V$ have been shown to vary smoothly with composition (Courbon et al., 2003). There is no model available in the literature to predict a priori the evolution of these parameters with composition, so the industry has to rely on a phenomenological approach. But the empirical relations are now identified on parameters which have a transparent physical meaning, and a smooth variation with both the composition and the initial cold work. Interpolation or even extrapolation to other compositions and other process parameters are now more reliable.

\textbf{Figure 9.1:} Comparison between the yield stress evolution during recovery at different temperature and initial cold work (closed symbols), and the model given by Equation (9.4) (open symbols).
9.3.2 Competition Between Recrystallization and Precipitation

Many industrial alloys, Fe-C-Nb being an example, benefit from the hardening both of precipitation and dislocation storage. Under suitable processing conditions, the hot deformation of a supersaturated solid solution may be followed by simultaneous recrystallization, recovery and (at lower temperatures), strain-induced precipitation processes. A quantitative description of this microstructural evolution can be built upon existing descriptions of the separate processes, but to realistically describe the complex behavior, the various elementary descriptions must be coupled to account for the key interactions between recovery and recrystallization, recrystallization and precipitation and precipitation and recovery.

The interactions of both recrystallization and recovery with strain-induced precipitation are very similar. The presence of precipitates affects the mobility of structural defects through ‘pinning’ and ‘solute drag’ effects and the presence of structural defects may affect precipitation through a modification of the activation barrier for nucleation and the nucleation site density. In the case of recovery and recrystallization, the driving force for both processes is the stored energy of deformation. The progress of recovery will reduce the driving force available for the migration of recrystallization boundaries and should slow down the recrystallization process: recovery and recrystallization are viewed as competing processes.

All the phenomena outlined above are well described in the literature. For an overview on recrystallization and recovery, see Humphreys and Hatherly (1996), for an overview on precipitation Kostorz (2001). Assuming site-saturated nucleation \( N_{\text{rex}} \) sites per unit volume and a linear relation between the recrystallization boundary velocity and the stored energy due to dislocations, \( G(t) \), leads to an Johnson-Mehl-Avrami law of the form (9.7) for the extended volume fraction recrystallized.

\[
X_{\text{ext}} = N_{\text{rex}} \left( \int_0^t M(t) G(t) \, dt \right)^3, \quad (9.7)
\]

where \( G(t) \) is the driving force for recrystallization and \( M(t) \) is the grain boundary mobility. \( F \) and \( R \) are the precipitate size and volume fraction, and \( \rho \) is the dislocation density.

The driving force for recrystallization, \( G(t) \), has been approximated as the stored energy of deformation, modified by a retarding Zener drag term to capture the potential pinning effect of precipitates on the recrystallization boundary. Appreciating that the dislocation density, \( \rho(t) \), is subject to time dependent recovery processes, the time dependent driving force for recrystallization can be expressed as:

\[
G(t) = \frac{\rho(t) \mu b^2}{2} - \frac{3 \gamma_{gb} F_v(t)}{R(t)}. \quad (9.8)
\]

The effect of solute elements on the mobility of the grain boundary has been captured by using a simplified mobility based on the classic treatment of Cahn, where the mobility is constant with velocity and proportional to the solute content. Since the solute content of the matrix is modified by precipitation processes, the mobility is also time dependent.

The recovery kinetics of the dislocation contribution to the hardening is described using the approach already used in Section 9.3.1. The simplest possible approach has been used to
capture the effect of precipitation on the kinetics of recovery. For the sake of simplicity, and in accordance with TEM observations, it is assumed that precipitation first occurs on the nodes of a dislocation network and ‘pins’ those segments between precipitates. These segments are unable to recover until the precipitation coarsening process decreases the number density of precipitates and ‘unpins’ dislocation nodes. When the number density of precipitates, $N(t)$, exceeds the number density of dislocation nodes, $N_c(t) = 0.5\rho(t)^{1.5}$, recovery processes are halted.

The precipitation kinetics of Nb(CN) in austenite have been described using a simple approach focusing on the prediction of the average radius and the volume fraction (Deschamps and Brechet, 1999). Precipitation occurs in two stages, firstly, a nucleation and growth stage followed by a growth and coarsening stage. The mean particle diameter and the number density of precipitates are monitored. The effect of capillarity is included in the analysis. The effect of recovery on precipitation is captured through a time dependent decrease in the number of available nucleation sites for strain induced precipitation. It must be stressed that more refined models exist in the literature, at the same scale: the so called “model in class”. In addition, this new generation of models have been extended to ternary alloys, using the framework of the solubility product. These models are useful to deal with non isothermal heat treatments and reversion kinetics, and to non stoichiometric precipitates, and have been applied to the prediction of precipitation evolution during welding (Nicolas, 2002). In the present situation, the NbC precipitates are stoichiometric, and the heat treatments under consideration are isothermal: the simpler model, relying on a pseudo binary approach and focussing on the average quantities is sufficient. Its especially important in complex industrial materials to select the simplest model required, compatible with the knowledge we have of the system.

The model predictions have been compared with the experimental measurements of Kang et al. (1997). These authors measured the precipitate particle diameter (TEM), the softening fraction (double-hit compression) and the recrystallized fraction (optical metallography) as a function of time in a Nb-microalloyed steel at several temperatures. A comparison of the experimental measurements with model predictions at 850°C are shown in Figures 9.2a and b.

The predicted time evolution of particle diameters and the softening and recrystallization fractions are in excellent agreement with the measured values. It is especially interesting to see that the ‘hump’ in the experimental softening curve between 1 and 100 s is very well reproduced by the model. The softening fraction was evaluated using a rule of mixtures of the recrystallized and unrecrystallized fractions coupled to a strengthening model which included contributions from dislocation hardening, precipitation hardening and solid solution strengthening.

The occurrence of plateau’s in the softening fraction vs. time curve is well-documented in the literature (eg. Medina et al. (1999)). From the present investigation, it has become clear that one must consider carefully the different meanings of a plateau in the recrystallization curve and a plateau in the softening curve. A plateau in the time evolution of the true recrystallized fraction arises when the Zener pinning pressure exceeds the stored energy driving the migration of the recrystallization boundary. In all the cases we have investigated, the stored energy greatly outweighs the Zener pressure for much of the investigated time frame. The Zener pressure only begins to outweigh the stored energy at very long times ($\sim 30,000$ s, corresponding with the onset of the recrystallization plateau seen in Figure 9.2a, and only after the stored energy has been substantially diminished through recovery processes. Re-
9.3 Some Examples of Physically Based Models for Industrial Materials

Figure 9.2: (a) Predicted and experimentally measured Nb(CN) particle size evolution in a Nb microalloyed steel containing 0.03 Nb (wt.%) and 0.076 C (wt.%) treated at 850°C. (b) Predicted and experimentally measured time evolution of the softening and recrystallized fractions at 850°C. (strain = 0.3, strain rate = 10−1).

Recovery appears to be a key ingredient for a quantitative understanding of the plateau in the recrystallization fractions. The morphology of the softening fraction curve is a manifestation of the time evolution of several hardening and softening contributions, including hardening by precipitation and softening by recrystallization, recovery and decreases in solid solution strengthening. The softening curve in Figure 9.2a exhibits a ‘hump’ between 1 and 100 s during which the softening fraction steadily rises to a maximum and then falls. The ‘hump’ is due to the interaction between recovery and precipitation and is not related to the recrystallization process. The experimental recrystallization data clearly show that recrystallization does not occur for 2000 s in this alloy at 850°C. The decrease in stored energy is a consequence of recovery processes and the plateau arises from the inhibition of recovery by precipitate pinning. Coarsening of the precipitate distribution ‘unpins’ the dislocation network and allows
recovery processes to resume after $\sim 200$ s. During the period where recovery is inhibited, precipitation is the dominant microstructural change occurring and this ‘hardens’ the material giving rise to a ‘hump’ in the softening curve rather than a ‘plateau’. Interpretation of the experimentally observed softening curves must be made carefully with a consideration of the relative magnitudes of all the operative contributions to hardening. Recovery processes are an essential ingredient for a quantitative understanding of the morphology of softening fraction curves.

9.3.3 Optimizing Casting Process in Precipitation Hardenable Alloys

Wrought aluminum alloys have been thoroughly investigated as far as the relation between microstructure and mechanical properties is concerned. Comparatively, cast aluminum alloys have been much less investigated. The reason for this apparent lack of interest lies in the industrial situation: for a long period, the limiting factors for cast alloys have been the processability (mould filling, hot tearing, porosities) and much of the emphasis in research have been focussed on these issues. The driving force to use cast alloys (which are in a way a natural Metal Matrix Composite) as structural components has motivated new research on the optimization of microstructure. The example we are presenting here is dealing with a key feature of cast alloys: the heterogeneity of solute distribution and its consequence on the efficiency of further heat treatments. We will focus our attention on binary alloys, with off eutectic compositions, but the procedure we propose is to be seen as a generic strategy to optimize heat treatments in cast alloys for strength design.

Modelling of solidification is based on the resolution of the solute conservation equation in spherical coordinates, coupled with a heat balance. The size of the spherical domain represents the average grain size. Since the dendritic nature of the grains is not accounted for, only the formation of “globulitic” grains is modelled, whose approximation by spheres is well justified. Such grain structure is often observed in aluminum casting alloys, a common industrial practice being to add into the melt very efficient inoculation particles. During cooling of the system from an initially liquid state, each new solid phase starts to form without nucleation undercooling once the uniform temperature domain reaches a new equilibrium temperature (i.e., either the liquidus or the eutectic temperature). The inputs of the model are thus the average grain size after solidification (equivalent to the final grain density), and the cooling rate during solidification imposed through a constant enthalpy change over time. The outputs of the calculations are the evolution with time of the phase fractions as well as the evolution with time of the solute profile in the grain envelope at room temperature. Homogenization is modelled using the same one-dimensional solute diffusion model developed for solidification. It is done by heating the system (through a positive enthalpy change over time) up to the desired homogenization temperature, and by keeping the system at that temperature for a given time. As a result, the Al$_2$Cu phase is progressively dissolved, eventually disappears, and the concentration profile within the grain is flatten. Modelling of precipitation is carried out by using the model first proposed by Langer and Schwartz (1980) and modified by Kampmann and Wagner (1984). The model considers nucleation, growth and coarsening of the precipitates as concomitant processes. Rate equations are thus written and solved using a Runge-Kutta algorithm. As a result of the calculation, the size and volume fraction of the precipitates are determined at each radial position within the spherical grain as a function of the
local composition and for a given artificial ageing temperature. From the output of the phase transformation models, the structural hardening model developed by Deschamps and Brechet (1999) is used. It accounts for the volumetric fraction and for the size of the precipitates, as well as for the remaining solute content of the matrix. The flow stress is thus calculated as a function of the interaction of the dislocations with the precipitates, accounting for the transition occurring between the shearing and the Orowan mechanisms, as well as for solid solution strengthening. The model is applied at different location within the grain to deduce the space dependence of the flow stress, $\sigma_y(r)$. Finally, the global yield stress of the material is deduced from a mechanical model developed by Hervé and Zaoui (1993) for a heterogeneous material. The model is based on a self-consistent Eshelby-type approach and is extended to a material made of spherical embedded layers. Each layer is characterized by individual mechanical properties computed by the structural hardening model.

The main results of the model are summarized in Figure 3. For a given alloy composition, with different combinations of the process parameters, it is found that the variation of the yield stress of the alloy is a direct function of the residual eutectic fraction. This simple final result

Figure 9.3: Variation as a function of the residual eutectic fraction, $g_\text{eut}$, of the normalized conventional yield stress of cast alloys, with respect to a perfectly homogenized and segregation free alloy (no eutectic phase left and no solute gradient). The various curves correspond to different process parameters, and all can be described by a master curve given by the volume fraction of eutectic phase resulting from the thermal history Gandin et al. (2002).
deduced from several calculations is consistent with the usual industrial practices. In order to optimize the yield stress of cast alloys, the process parameters are adjusted so as to maximize the solute content of the Al solid solution. In turn, this practice corresponds to a minimization of the residual eutectic fraction.

All the unit models are grouped into a single integrated model which describes the microstructure change of aluminum-copper alloys during its process route, and its effect on the final yield stress. This “integrated model” establishes direct links between the main process parameters of cast alloys and its final yield stress (Gandin et al., 2002).

9.4 Perspectives

The last ten years have seen a considerable development of modeling in materials science, both from the academic viewpoint, and from the industrial viewpoint. Metallurgy has been at the leading edge of this evolution, leading to so-called “computer assisted metallurgy”. It is likely that this trend will remain for a number of years. It doesn’t mean that modelling is replacing experiments, it means that it has reached a degree of maturity which allows a better interaction with experimentalists.

A number of questions have been listed along this paper, on which further work is required. We would like to conclude this contribution by some personal views on some questions in which modelling might prove to be very useful. These questions fit in three categories: the unsolved questions, the integrated approaches, the fields to explore.

The unsolved questions are many, but it seems to the author that the solution of some of them is necessary for further progress in other unsolved problems. For instance dislocation patterning and its relation to stored energy seems a stumbling stone in the problem of modelling recrystallization. It is not enough to get a cellular structure in a computer simulation, it is necessary to understand why such a structure appears in some systems and later in other. The solution of dislocation patterning is likely to come from the joint effort of atomistic simulation feeding information into mesoscopic dislocation dynamics simulations. Another example is the problem of nucleation of precipitation: Monte Carlo simulation provide new insight in the problem, which should be incorporated into more mesoscopic simulations such as cluster dynamics, to have an input from atomistic information into modeling of precipitation. Again, alloy design is the perspective. Interface driven transformations are at the same time one of the most important issues in the metallurgy of steels, and one on the most forgotten by physicists: a proper understanding of the reaction kinetics requires a better modelling of the conditions at the interface: this requires both studies on model alloys, appropriate thermodynamical and diffusion data, and a theoretical effort to understand solute drag on interphases.

Most of the models developed in academic world are not really building the bridge between different scales. At best, they use an information from a lower scale, and provide an information to the upper scale. They also very often remain at the level of simple loadings or simple thermal history whereas the real challenge is now to build models robust enough to be predictive for complex thermomechanical history. This challenge is what we call “integrated modeling”. In this respect, the modelling of welding can be seen both as a paradigm and as a challenge. Welding brings together all physical metallurgy in 3 cm. Heat affected zone can present precipitation, dissolution, coarsening, recrystallization and grain growth. Understand-
9.4 Perspectives

Modeling the molten region requires to model the fluid flow in the bath, in the mushy zone, the solidification structure, the stress build up and the possible hot tearing. On the larger scale, the quantification of building of internal stresses requires the implementation of appropriate constitutive equations for plasticity with an evolving microstructure. The subsequent loading of weld is also a complex problem from a damage mechanics viewpoint, one where the structure problem and the materials problem become closely interrelated. A proper modeling of weld microstructures and properties, on physical grounds, is still a distant goal, but not so distant that it cannot be given as a reachable challenge. This challenge is of considerable industrial importance, and it also implies further improvements in all the pending questions listed in this paper. Understanding the precipitation kinetics in a non isothermal treatment, understanding the conditions for recrystallization (with its related problem of the dislocation patterning), understanding the competition between recrystallization and precipitation, understanding hot tearing criterion and its occurrence during the solidification procedure, understanding the yielding and damage of a mechanically heterogeneous structure.

A number of fields of considerable interest, both scientifically and industrially, have received comparatively little attention from the community from the viewpoint of modeling. Very often, one is left with a few semi-empirical rules, but not a deep understanding of the underlying mechanisms. These fields are certainly not unexplored, but their importance would require an enhanced activity. Among these questions, one can list as examples: fatigue damage and its relation to slip irreversibility, coupling between corrosion and plasticity, relation between wear and alloys microstructures, hot tearing and processing defects and their evolution during further thermomechanical treatments.

Last, but not least, an open challenge to the modeller has emerged in recent years from a comparative approach between materials. For instance, the mechanics of heterogeneous materials has to face new challenges while dealing with materials when the inclusion/matrix paradigm is no longer operative (such in the case of foams or entangled materials), and these challenges will be better solved if polymers and metals are approached in a comparative manner. The modeling of mechanical properties of bulk amorphous materials will probably benefit a lot from the work done for organic and inorganic glasses.

Finally, we will be facing in the next years more and more design driven modeling. It is the belief of the author that such a procedure, using modeling to design a material for a given application, is a natural and promising development of both modeling in materials science, and of materials and process selection in mechanical design (Ashby and Brechet (2003), Gasser et al. (2004)). Materials by design is the future, and the coupling of modeling and design approaches, especially in hybrid structures, will be a key issue to be tackled in the coming years, both for academic interest, and for industrial applications.

This brief outline of recent developments in modeling microstructures and materials behavior, having in mind industrial applications doesn’t pretend to be exhaustive, neither objective. They reflect even only partially the author’s own center of interest. Many other problems, emerging from industrial demand, would lead to a positive answer to the opening question of the title: physically based models, what for?
Acknowledgements

It is a pleasure to thank colleagues and former students for collaborations and discussions over the years which have shaped the content of the present contribution, among which the ones directly involved in some of the studies presented here. The four cases studies presented here came from the work of Dr. M. Verdier, Dr. H. Zurob, Dr. C. Hutchinson, Dr. C.A. Gandin. Constant collaborations over the years with M. Ashby, A. Deschamps, D. Embury, P. Guyot, G. Purdy, L. Salvo, Y. Estrin, H. Sherciffl, M. Rappaz, T. Pardoen, have greatly contributed to shape the ideas presented here. Last but not least, the long lasting collaborations with Pechiney and Arcelor companies have provided both the industrial origin of the problems analyzed here, and partial funding for their study.

References

A. Benzerga, Y. Brechet, A. Needleman, E. Van Der Giessen, J. Comp. Sim. Modelling, Mat Sc. Eng (2004), 12, 159.
G. Kostorz, (2001), Phase Transformation in Materials, Wiley VCH.
References


Part II

Application to Engineering Microstructures
10 Modeling of Dendritic Grain Formation During Solidification at the Level of Macro- and Microstructures

Michel Rappaz, Alain Jacot, and Charles-André Gandin

Modeling of dendrite formation during solidification has made substantial progress thanks to the development of the phase-field (PF) method. However, this method is still limited when the undercooling is small, as it is difficult to control the attachment kinetics term, unless specific corrections are introduced [Karma 2001]. An alternative to the PF method, the so-called pseudo-front tracking (PFT) method, has been recently developed. It is able to solve the diffusion equation in the liquid and solid phases with appropriate interfacial conditions, using a fixed grid as the PF method, while the curvature of the interface is calculated using a reconstruction algorithm. The method has been coupled to thermodynamic databases using different strategies and can be applied to multi-component systems. The PFT and PF methods have been compared with analytical models of dendrite growth and give similar results. On the other hand, the computation time associated with either the PF or PFT is still very large and calculations are limited to small computational domains. For applications to large scale castings, a Cellular Automaton (CA) method can be used. This physically-based numerical method combines analytical models of nucleation and growth for the simulation of dendritic grain formation with Finite Element (FE) heat flow calculations. The so-called CAFE model predicts many features observed in solidification processes such as grain morphologies, transition from columnar-to-equiaxed regions, evolutions of grain size and grain texture.

10.1 Introduction

Grain structures and microstructures, which form during solidification processes of metallic alloys, are mainly governed by the temperature field and convection effects at the macroscopic scale, and by solute diffusion and interfacial solid-liquid energy, at the microscopic level [Kurz and Fisher 1989, Flemings 1974]. It is only under high solidification speed that solute trapping and/or interfacial kinetics may become important. In analytical models of solidification microstructure formation, it is common to introduce the various causes of equilibrium loss of the interface through several contributions to the total undercooling, $\Delta T$, i.e.: 

$$ \Delta T = T^{SI} - T^{\infty} = \Delta T_i + \Delta T_s + \Delta T_r + \Delta T_k + \Delta T_{st} $$

(10.1)

where $T^{SI}$ would be the normally expected equilibrium temperature (i.e., liquidus temperature associated with the local composition of the alloy) and $T^{\infty}$ the temperature of the bath that could be measured with a thermocouple (Figure 10.1). Please note in Figure 10.1 that the thermal undercooling, $\Delta T_i$, measures the actual temperature difference between the interface...
temperature, $T^*$, and the bath temperature. The other undercooling contributions, which measure the difference ($T^{eq} - T^*$) are associated with: solute diffusion ($\Delta T_s$), curvature effect ($\Delta T_r$), attachment kinetics ($\Delta T_k$) and solute trapping ($\Delta T_{st}$) [Kurz and Fisher 1989]. Recently, a new contribution, $\Delta T_b$, has been introduced by Rappaz et al. in order to describe near the end of solidification, the undercooling necessary to overcome the grain boundary energy of two coalescing (bridging) dendrites [Rappaz et al. 2003].

![Figure 10.1: Schematics of the various undercoolings of a dendrite tip growing in an undercooled melt. The major contribution encountered under normal solidification conditions is the solutal undercooling associated with diffusion of solute in the liquid, $\Delta T_s$.](image)

In principle, the phase-field method, first applied to metallic alloys by Warren and Boettinger, is capable of handling all these contributions [Warren and Boettinger 1995]. It might be therefore tempting to use such a technique to model microstructures and grain structure formation in actual solidified parts. However, such an idea is unrealistic because of the length scales involved (Figure 10.2):

- The length scale associated with thermal diffusion is typically equal to the size of the component (i.e., cm – m) since solidification has to be induced by external cooling of the domain boundaries;
- The length scale associated with solute diffusion and curvature is directly linked with the microstructural features (e.g., dendrite tip radius, 0.1 – 10 $\mu$m);
- The two last contributions, $\Delta T_k$ and $\Delta T_{st}$, (as well as $\Delta T_b$) must consider the atomistic level (i.e., 0.1 – 1 nm).

It is clear that no numerical model can account at the moment for a $10^{13}$ length scale (i.e., $10^{30}$ volume scale) difference, even with adaptive grid methods [Provatas et al. 1999, Jeong
10.1 Introduction

Figure 10.2: The various scales of solidification: turbine blade, grain structure, microstructure and atomic level. Each level requires adapted observation and simulation techniques: eye, optical microscopy (OM), scanning electron microscopy (SEM) and high resolution transmission electron microscopy (HRTEM) for the observation (top); FDM or FEM, cellular automata, phase-field or equivalent method, molecular dynamics for the simulation.

et al. 2001, Burman et al. 2003]. Accordingly, in a way similar to the various observation techniques applied at each specific level (from eyesight- to optical-, scanning- and high resolution electron microscopy observations), various numerical methods have been developed to tackle the different length scales (Figure 10.2). At the atomistic level, molecular dynamics or its variants (e.g., embedded atom method) can predict the interface morphology (facetted or non-facetted) and defects (vacancies, twins). On the scale of the whole solidification process, finite element or finite difference/volume methods based on average conservation equations have been developed in order to model heat and mass transfer, stresses and strains development.
and other macroscopic features (e.g., macrosegregation). The present contribution will present two numerical methods especially developed for the intermediate scales of the microstructure and grain structure. In Section 10.2, the pseudo-front tracking method, which is somewhat equivalent to the phase-field method [Karma, this book], will be presented. Coupled with phase diagram computations (Section 10.3), this technique allows to predict microstructure formation in multi-component alloys (see Section 10.5). In Section 10.4, a physically-based cellular automaton (CA), which integrates analytical laws for nucleation and growth kinetics, will be briefly presented. Coupled with Finite Element heat flow calculations (so-called CAFE method), this allows to predict dendritic grain structures and texture evolutions in various solidification processes. Whereas industrial applications of the CAFE approach are given in the contribution of Gauemann et al. [this book], Section 10.5 will show validations related to the Columnar-to-Equiaxed Transition (CET) in directionally solidified ingots and to the grain selection occurring among columnar grains during directional solidification.

10.2 Pseudo-Front Tracking Model

The pseudo-front tracking method is a 2-dimensional model which allows to describe solidification microstructures according to two calculation steps. The first one describes the formation of the primary phase by solving the diffusion equations in the solid and liquid phases [Jacot and Rappaz 2002]. In a second stage, the model gives a description of the formation of eutectic or peritectic phases from the remaining interdendritic liquid [Qiang and Jacot 2003]. The model is coupled to thermodynamic calculations and can be applied to multi-component alloys.

10.2.1 Primary Phase Formation

The primary phase model is based on the assumption that the temperature is uniform (or given at each location) on the scale of the microstructure and that the phase transformation is governed only by diffusion of the solute element(s) and curvature effects (i.e., $\Delta T_k$, $\Delta T_{st}$ and $\Delta T_t$ in Figure 10.1 are assumed to be negligible).

The growth of the primary phase ($\alpha$) from the liquid ($l$) is described with the following set of equations:

$$\frac{\partial w^\alpha_i}{\partial t} = \text{div} \left[ D^\alpha_i \text{grad} w^\alpha_i \right] \quad \text{with} \quad i = 1, 2, \ldots, n \quad \text{and} \quad \nu = \alpha, l \quad (10.2)$$

where $n$ is the number of alloying elements in the multi-component system, $w^\alpha_i$ is the concentration and $D^\alpha_i$ is the diffusion coefficient. Please note that cross-diffusion has not been considered. The following solute balance must be satisfied at the $\alpha/l$ interface:

$$D^n_i \left[ \frac{\partial w^n_i}{\partial n} \right]^* - D^l_i \left[ \frac{\partial w^l_i}{\partial n} \right]^* = \nu \cdot n \left( w^l_i - w^n_i \right) \quad (10.3)$$

where $\nu$ is the interface velocity and $n$ is the normal vector to the interface pointing towards the liquid. The superscript "*" denotes concentrations (or gradients) taken at the curved interface. The interfacial concentrations $w^l_i$ and $w^n_i$ in each phase are deduced from the phase diagram, taking into account the curvature of the interface.
With the pseudo-front tracking method, the solution to this problem is based on a finite volume method formulated for a regular hexagonal grid as shown schematically in Figure 10.3(a). A volume element or cell of the mesh has three possible states: \( \alpha, l \), or interface \( (\alpha/l) \). A layer of \( \alpha/l \) interface cells always separates the two phases. An explicit formulation of Equation (10.2) provides a variation of solute concentrations in each of the \( \alpha \) or \( l \) cells. For the interfacial cells, these variations correspond to an average over the two phases. Accordingly, the variation of the average concentration in the interfacial cells is converted into a variation of volume fraction of liquid using a local lever rule, the phase diagram information and the local curvature of the interface.

**Figure 10.3:** Hexagonal grid used in the pseudo-front tracking method. Part (a) illustrates the layer of interfacial cells (thick lines) which separates the liquid and solid domains. Part (b) shows the PLIC segments, normal vectors and distance field (grey levels) used to describe the interface.

The curvature of the interface is obtained through the calculation of an additional field: the signed distance to the interface, \( \delta \). This quantity is negative in the solid, positive in the liquid and null on the interface. The calculation of \( \delta \) requires a precise knowledge of the position of the interface. The PLIC algorithm (Piecewise Linear Interface Calculation), which was originally developed for fluid flow problems [Kothe et al. 1996], is first used to reconstruct the interface from the solid fraction field. The PLIC method is based on the assumption that the interface is a straight line inside any given interfacial cell. The position of the linear segment representing the interface in a cell, \( \xi \), is univocally defined from the solid fraction in the cell, \( g_\xi^\alpha \), and the normal vector to the interface, \( n_\xi \).

The interface curvature is obtained through the following algorithm: (1) Calculation of the PLIC segments for each interfacial cell; (2) Calculation of the distance, \( \delta \), of each cell located within a certain bandwidth surrounding the interface. The distance \( \delta \) is defined as the shortest distance to the PLIC segments present in the calculation domain; (3) Calculation of the normal vector to the interface in each interfacial cell using the following expression: \( n = \frac{\text{grad} \delta \cdot \text{grad} \delta^{-1}}{\|\text{grad} \delta\|} \); (4) Calculation of the interface curvature with the relationship \( K = \text{div} n \). Figure 10.3(b) illustrates the calculated distance field and the reconstructed interface obtained with the PLIC algorithm in a small region in the vicinity of the interface. As can be seen, the PLIC segments of adjacent interfacial cells are not necessarily connected,
although they almost match if the mesh size is small with respect to the radius of curvature of the interface.

The growth directions of the dendrites being dictated by the weak 4-fold anisotropy of the interfacial solid-liquid energy, it has to be verified that the grid used in the PFT method does not introduce a larger numerical anisotropy which would then force the dendrites to grow along the 6-fold symmetry axes of the hexagonal grid.

10.2.2 Secondary Phases Formation

As soon as the liquid becomes undercooled for another solid phase, the calculation enters a second stage aimed at predicting the formation of secondary phases in the remaining interdendritic regions [Qiang and Jacot 2003]. In this approach, the interdendritic regions are considered as a mixture of liquid and solid phases. The model is based on the following assumptions: (i) The composition of the interdendritic liquid is uniform when the first secondary phase starts to form; (ii) The interdendritic region is locally in thermodynamic equilibrium and all the phases have uniform concentrations, $w^\nu_i$, including the $\alpha$-phase formed within the mixture. (There is therefore no need to use an index \( \nu^* \) for the concentrations $w^\nu_i$ in the phases $\nu \neq \alpha$). For the $\alpha$-phase, these concentrations are given by the concentrations, $w^\alpha_i^*$, at the boundary between the primary phase and the mixture. If some liquid remains in the mixture, they are given by the solidus of the $\alpha$-phase; (iii) Back-diffusion affects the interdendritic region in a uniform manner.

A solute balance is performed over the interdendritic domain, $\Omega_m$, for all the solute elements. Assuming uniform concentrations in the mixture region, the overall flux, $\phi^\alpha_i$, of solute $(i)$ flowing in the primary phase $\alpha$ is given by:

$$
\phi^\alpha_i = - \int_{\Gamma_{\alpha/m}} J^\alpha_i \cdot n \, d\Gamma = \int_{\Gamma_{\alpha/m}} D^\alpha_i \left[ \frac{\partial w^\alpha_i}{\partial n} \right]^* \, d\Gamma = -V_m \frac{dw^m_i}{dt} + \frac{dV_m}{dt} \left[ w^m_i - w^\alpha_i^* \right] \quad \text{for } i = 1, n \quad \text{(10.4)}
$$

where $w^m_i$ is the average concentration of element $(i)$ in the interdendritic region, $\Gamma_{\alpha/m}$ is the interface between the primary phase and the interdendritic region (Figure 10.4), $V_m$ is the volume of the mixture region. The mixture composition can be expressed as:

$$
w^m_i = \sum_{\nu=1}^{p} g_\nu w^\nu_i \quad \text{for } i = 1, n \quad \text{(10.5)}
$$

where $p$ is the number of phases in the mixture and $g_\nu$ is the volume fraction of phase $\nu$ in the interdendritic region. As for the primary phase solidification, the fluxes $\phi^\nu_i$ can be calculated explicitly at each time step, using the solute profile in the primary phase.

In order to describe the evolution of the mixture (and of the primary phase), one has to describe the evolution of all the concentrations in all the phases present in the mixture, $w^\nu_i(t)$, the volume fraction of all the phases, $g^\nu(t)$, and the volume of the mixture, $V_m(t)$, i.e., $p(n+1)+1$ variables if the temperature history, $T(t)$, is known. For that purpose, one has Equation (10.4) ($n$ equations) and the information given by the phase diagram, i.e.:
10.3 Coupling with Thermodynamic Databases

The thermodynamic data needed for the calculation, i.e., the liquidus/solidus or solvus temperatures and the partition coefficients are obtained with thermodynamic calculation tools based on the CALPHAD approach. Since thermodynamic calculations of multi-component alloys are quite computationally intensive, two different strategies have been implemented for the primary and secondary phases solidification.

10.3.1 Primary Phase Formation

Due to the evolution of the primary phase during solidification, it is necessary to consider the spatial variations of the concentrations along the $\alpha/l$ interface. However, these variations...
are small, since they are essentially associated with the small curvature undercooling and the slightly different solidification paths of each portion of the interface. Instead of calculating phase equilibria for each interfacial cell, Jacot and Rappaz (2002) make at each time step only \((n + 1)\) calls. Indeed, the phase diagram, i.e., the liquidus temperature, \(T^{l/\alpha}\), and the partition coefficients, \(k_{\alpha/l_i}^{\alpha/l}\), can be locally linearized around the average concentrations, \(\bar{w}_i\). In order to deduce the variations of these thermodynamic parameters with the variations \(\delta w_i = (w_i - \bar{w}_i)\), \((n + 1)\) calls to the equilibrium phase diagram calculations have to be performed: one call with the average concentrations and \(n\) calls for a small increment along each of the concentration. A few additional calculations are needed to test if secondary phases can form.

### 10.3.2 Secondary Phases Formation

In this case, assuming uniform concentrations in the mixture, only one phase diagram computation is needed per time-step in order to determine the amount of phases from the temperature and the concentrations, \(w_i^\nu (\nu \neq \alpha)\) and \(w_i^\alpha\).

### 10.4 Cellular Automaton – Finite Element Model

Two- (2D) and then three-dimensional (3D) physically-based Cellular Automaton (CA) methods with a regular array of cubic cells have been developed over the past ten years for the prediction of dendritic grain structures [Gandin et al. 1993-1999]. Contrary to the PF and PFT methods, the CA method briefly presented hereafter does not aim at simulating the complex development of the dendritic patterns. Instead, the development of the external envelop of each dendritic grain is simulated using simplified growth kinetics laws, its inner solid-liquid mixture being characterized by an internal volume fraction of solid. Therefore, dendritic solidification is mimicked at the scale of the grain- or macro-structure, which, for many solidification processes, is as important as the simulation of the microstructure itself. The CA method is thus applied at a scale which is much closer to the scale of the casting, as compared with the PF and PFT methods (Figure 10.2).

![Figure 10.5: 2D schematic representation of the cellular automaton (CA) square grid superimposed onto one triangular element of a FE mesh. The interpolation coefficients, \(\varphi_k(\nu)\), are defined for each CA cell, \(\nu\), with respect to the FE nodes, \(k\).](image)
Figure 10.5 shows a schematic 2D representation of a regular array of square cells which defines the CA grid. It is superimposed onto one triangle of a Finite Element (FE) mesh; this later being used to solve macroscopic average conservation equations (e.g., heat flow, fluid flow, stress-strain) at the scale of the casting. A state index, \( I_\nu \), the value of which corresponding to a physical state, is attributed to each cell \( \nu \): \( I_\nu = 0 \) if the cell is liquid and \( I_\nu \neq 0 \) if the cell is already partially solid (i.e., mushy). Temperature- and orientation-dependent CA rules have been defined in order to modify the cells state indices based on nucleation and growth laws for dendritic grains.

10.4.1 Nucleation Law

Heterogeneous nucleation sites which can give rise to nucleation events are randomly distributed among the cells of the solidifying domain. A distinction is made with another reference index between the cells belonging to the bulk of the domain (heterogeneous nucleation in the volume) and those at the boundaries (heterogeneous nucleation at the surface of the domain). Therefore, several nucleation rules can be defined for the volume and surface boundaries of the domain. The location of each nucleation site, i.e., its cell number \( \nu \), is randomly chosen in each category. Each of these sites is characterized by a critical nucleation undercooling, \( \Delta T_{\text{nucl}} \). These undercoolings are assumed to follow a Gaussian distribution, \( (dn/d(\Delta T)) \), describing the density of nucleation sites, \( dn \), becoming active in an undercooling increment \( d(\Delta T) \), i.e., in a temperature interval \([T_L - \Delta T, T - \Delta T]\). Therefore, for the bulk/surfaces of the solidifying domain, \( n \) is the density of grains that can possibly form up to an undercooling \( \Delta T \). Such nucleation laws require three parameters: the mean undercooling, \( \Delta T_N \), the standard deviation, \( \Delta T_\sigma \), and the integral of the distribution, \( n_{\text{max}} \) [Rappaz 1989].

During a thermal calculation (see hereafter), as soon as the local undercooling, \( \Delta T_\nu \), at a given nucleation site location, \( \nu \), becomes larger than the critical nucleation undercooling associated with its nucleation site, \( \Delta T_{\text{nucl}} \), nucleation takes place and a new grain forms with a random orientation, providing the cell was still liquid. The value of the state index of cell \( \nu \), \( I_\nu \), is then changed to a non-zero integer value which identifies a crystallographic orientation chosen among predefined orientation classes. These orientation classes are initialized at the beginning of the calculation by a random selection of the three Euler angles [Bunge 1982] in order to produce a uniform distribution of the \( \langle 100 \rangle \) directions on a unit sphere.

10.4.2 Growth Law

The growth of the newly nucleated grain is modeled by increasing the size of a regular octahedron whose center is located at the center of cell \( \nu \). The diagonals of the octahedron are aligned with the \( \langle 100 \rangle \) crystallographic directions defined by the Euler angles. They correspond to the directions in which dendrite stems and arms of fcc metals preferentially develop. Their extension is calculated by integrating over time the growth kinetics law of the dendrite tips, \( v(\Delta T) \), computed with the local explicit undercooling of the cell, \( \Delta T = \Delta T_\nu \). This analytical growth kinetics law is based on the assumption that the tip has a parabolic shape with a radius given by the marginal stability limit criterion [Kurz et al. 1986]. Once the size of the octahedron is sufficiently large to capture the cell center of one of its liquid neighbors,
\( \mu \) (with \( I_\mu = 0 \)), the state index of cell \( \mu \) is switched to that of the parent cell \( \nu \) (i.e., \( I_\mu = I_\nu \)). If cell \( \nu \) is fully surrounded by mushy cells (\( I_\mu \neq 0 \) for all the neighboring cells), its growth is no longer considered. Both the center and the size of the octahedron associated with a newly captured cell \( \mu \) are initialized carefully in order to propagate the crystallographic orientation of the grain and to integrate correctly the dendrite tip growth kinetics using the local undercooling, \( \Delta T^*_k \) [Gandin et al. 1997]. The CA growth procedure has to verify that, after several steps and in a uniform temperature field, a simple octahedral shape is obtained with the correct misorientation with respect to the CA grid, in agreement with experimental observations [Ovsienko et al. 1974].

### 10.4.3 Coupling of CA and FE Methods

The local temperature or undercooling of the cells is a key parameter of the CA nucleation-growth algorithm. It has to be deduced from the solution of the average energy conservation equation, i.e., the heat flow equation averaged over the solid and liquid phases. Such a solution is calculated with FEM and a grid which can be much coarser than the CA cell size. Within a time step \( \Delta t \), the variation of the volumetric enthalpy at a FE node \( k \), \( \Delta (\rho H)_k \), is given by:

\[
\Delta (\rho H)_k = (\rho c_p)_k \Delta T_k - (\rho L)_{s/l}^k \Delta g_{s,k},
\]

where \( (\rho c_p)_k \) is the specific heat, \( (\rho L)_{s/l}^k \) is the volumetric latent of heat fusion, \( \Delta T_k \) and \( \Delta g_{s,k} \) are the associated variations of temperature and volume fraction of solid, respectively. (Please note that, unfortunately, "\( \Delta T \)" is the standard notation for both the undercooling and the temperature variation during a time step). The aim of the coupling of the CA and FE methods is to convert the variation of the volumetric enthalpy, \( \Delta (\rho H)_k \), into variations of temperature, \( \Delta T_k \), and volume fraction of solid, \( \Delta g_{s,k} \).

In a “weak” coupling mode, a unique solidification path, \( g_s(T) \) (e.g., the Gulliver-Scheil microsegregation path [Kurz and Fisher 1989]), can be used on the macroscopic scale (i.e., at the scale of the FE nodes). The variation of the volumetric enthalpy becomes a simple function of the temperature variation alone and the temperature field is directly solved on the macroscopic scale. Since the CA nucleation and growth rules defined above are temperature dependent, the temperature of a cell \( \nu \) of a time \( t \), \( T^*_t \nu \), is calculated from the temperature of the FE nodes \( k \), \( T^*_t k \). This is carried out by using the shape functions associated with the FE nodes \( k \), \( \varphi_k(r) \), and calculated at the cell position, \( r_\nu \). The temperature of cell \( \nu \) is then simply given by:

\[
T^*_t \nu = \sum_{k=1}^{N_n} \varphi_k(r_\nu)T^*_t k \tag{10.9}
\]

where \( N_n \) is the number of FE nodes. Since the value of the shape function at the cell location \( r_\nu \) is different from zero only for the nodes defining the element in which it is located, the summation is restricted to these nodes (i.e., \( T^*_t \nu = \varphi_1(r_\nu)T^*_1 + \varphi_2(r_\nu)T^*_2 + \varphi_3(r_\nu)T^*_3 \) in Figure 10.5).

In a “full” coupling mode, the enthalpy variation at a cell location, \( \Delta (\rho H)_\nu \), as well as the explicit temperatures, \( T^*_\nu \), are first interpolated from those calculated at the FE nodes \( k \) (equation similar to, or given by Equation (10.9)). Nucleation and growth can then be calculated at the cell level using the CA algorithm and the explicit temperatures. The variation of the solid fraction within each cell, \( \Delta g_{s,\nu} \), is calculated using a truncated Gulliver-Scheil
model [Hunt 1984, Gandin et al. 1999]. The variation of the volume fraction of solid at the node location, \( \Delta(\rho H)\), is obtained from a restriction:

\[
\Delta g_{s,k} = \frac{\sum_{\nu=1}^{N_v} \varphi_k(r_{\nu}) \Delta g_{s,\nu}}{\sum_{\nu=1}^{N_v} \varphi_k(r_{\nu})}
\]  

(10.10)

Please note again that, since the shape function \( \varphi_k(r_{\nu}) \) differs from zero only when the cell \( \nu \) and node \( k \) both belong to the same element, the summation in Equation (10.10) can be limited to the cells \( \nu \) located in the elements connected to node \( k \). Once \( \Delta g_{s,k} \) is known, the variation of temperature \( \Delta T_k \) can be calculated using the energy balance. This coupling scheme allows to predict for example recalescence at specific locations of the solidifying domain.

It should be emphasized that heavy additional numerical tools had to be developed in order to build a practical 3D CAFE model. Indeed, for a solidifying domain \( 10 \times 10 \times 10 \text{ cm}^3 \), the number of CA cells can be very large, typically \( 10^{19} \), since the cell size must be on the order of the secondary dendrite arm spacing, typically 100 \( \mu \text{m} \), if one wants to correctly predict the grain selection and competition. These tools include a dynamic activation of the cells, the definition and handling of micro-time steps, an optimal storage of the results (e.g., for cells located in predefined cuts or at the surface of the domain), etc. Specific details of the CAFE model can be found in [Gandin et al. 1999].

### 10.5 Results and Discussion

#### 10.5.1 PFT Model

The PFT model was used to predict microstructure formation in an Al-1wt%Mg-1wt%Si alloy (Figure 10.6). Two calculations were performed using a low cooling rate and high grain density (A) and a higher cooling rate and lower grain density (B). Both calculations were started at the liquidus temperature, where solid seeds were nucleated randomly in the 500 \( \mu \text{m} \times 433 \mu \text{m} \) calculation domain according to the selected grain density. Periodic boundary conditions were applied at the limit of the domain and the thermodynamic data were obtained from the database of Ansara et al. [1998]. Some of the parameters are listed in Figure 10.6, whereas more details of the numerical parameters can be found in [Jacot and Rappaz 2002]. Figure 10.6 illustrates the results of calculations A and B at different stages during solidification. The figures at the top show the grain morphology and the concentration field of Mg in the liquid at some early stage of solidification, whereas the two figures at the bottom present the grain structures at a later stage with the distribution of Mg in the solid phase. The latter figures correspond to the end of the primary phase calculation, just before the Mg\(_2\)Si phase could precipitate.

In the case of low cooling rate and high grain density (calculation A), it can be seen that the gradient of Mg concentration in the liquid is rather low due to the overlap of the diffusion layers around the grains at an early stage of solidification. As a result, the grains keep their
globular morphology throughout the entire solidification process. Some coalescence occurs among the grains as no grain boundary energy was introduced [Rappaz et al. 2003]. In the case of high cooling rate and low grain density (calculation B), the interaction between the diffusion layers in the liquid occurs later (except for liquid regions separating two close neighbor grains) and larger variations of Mg concentration are observed in the liquid. This situation leads to a destabilization of the solid-liquid interface and to a more dendritic-type grain morphology.
A few dendrite tips can grow freely in the liquid without much influence from other growing tips. However, the diffusion fields start to overlap before fully developed dendrites can form. Some thermal noise was intentionally added in this calculation to enhance the formation of secondary arms. It was found however that the grain morphology is not much sensitive to noise. It was concluded that the overlap of the diffusion layers is responsible for the scarcity of secondary arms.

In order to assess the ability of the model to correctly predict the grain morphology in two dimensional problems, systematic comparisons between the PFT and the PF methods were performed for binary alloys. The PF model used for these tests is close to the formulation presented in [Tiaden et al. 1998]. The two models were compared for a case of isothermal solidification in an Al-2wt%Cu alloy at a constant undercooling $\Delta T = 1$ K. As an initial condition, a small solid seed with a crystallographic orientation of 11 deg. with respect to the horizontal axis was placed at the center of a $400 \mu m \times 400 \mu m$ domain. The crystal orientation was voluntarily misaligned with respect to the grid in order to reveal potential grid anisotropy effects. Periodic boundary conditions and a grid spacing of 1 $\mu m$ were used in both the PF and PFT simulations. More details about the numerical parameters can be found in [Jacot and Rappaz 2002]. Figure 10.7 compares the shape of the dendrite and the concentration maps obtained at a given time with the PFT (left) and PF (right) methods. Although small differences can be noticed between the concentration fields, the overall agreement is very good. No particular effects of grid anisotropy are observed, neither in PFT nor in PF, since the dendrite arms are correctly oriented at 11 deg. with respect to the horizontal and vertical axes, despite the use of hexagonal or square grids.

![Figure 10.7: Concentration field calculated with the pseudo front tracking (left) and phase-field (right) models for an Al-2wt%Cu dendrite growing at a constant undercooling of 1 K. The domain dimensions are 400 $\mu m \times 400 \mu m$.](image)
The last example obtained with the PFT method and shown in Figure 10.8 is related to the formation of secondary phases during solidification and their partial dissolution during homogenization. The calculation was performed for a simple binary Al-1%Si alloy. As can be seen, the conditions correspond to a globulitic-type primary phase structure (grains represented with an intermediate grey level). The liquid (light grey) progressively disappears during the eutectic reaction and a coupled eutectic (dark grey) forms. By keeping the alloy just below the eutectic temperature, diffusion in the solid grains (back-diffusion) tends to homogenize the concentration and to dissolve some of the eutectic (i.e., the amount of eutectic initially given by a Gulliver-Scheil relationship tends, over extended homogenization times, to a value given by the lever rule).

### 10.5.2 CAFE Model

#### 10.5.2.1 Columnar-to-Equiaxed Transition in Directionally Solidified Ingots

A columnar-to-equiaxed transition (CET) experiment has been used for comparison with the predictions of a 3D CAFE simulation. This experiment was already used to validate the 2D CAFE model [Gandin et al. 1994b]. The mould was made of a highly insulated ceramic
material with an inner cylindrical shape. It lied on a water-cooled copper chill that ensured directional cooling and solidification of an Al-7wt%Si alloy. Seven thermocouples equally spaced from the bottom to the top of the ingot were used to characterize the temperature evolution and to determine the boundary condition to be applied at the chill-metal interface in the simulation. The nucleation parameters have been adjusted in order to reproduce both the position of the CET and the measured cooling curves. Figure 10.9 shows the comparison between the experimental (a) and the simulated (b) grain structures observed in a longitudinal cross section of the ingot. The overall result, in particular the grain appearance and its size evolution, is quite satisfactory, even though a fairly large cell size has been used in this case. It is found that the CET takes place once no superheat remains in the liquid (i.e., the thermal gradient has vanished) and the columnar growth front velocity starts to decrease. Detailed interpretations of such findings have been carried out recently [Gandin 2000, Martorano et al. 2003]. They suggest that the CET in the present directional solidification experiment is the result of a fragmentation of the columnar growth front.

**Figure 10.9:** Transition from a columnar to an equiaxed grain structure: (a) observed in a longitudinal cross section of an Al-7wt%Si directionally solidified alloy (diameter: 70 mm, height: 178 mm); and (b) simulated with the 3D CAFE model (cell size: 300 µm).
10.5.2.2 Grain Selection Occurring among Columnar Grains during Directional Solidification

The 3D CAFE model is a unique tool in predicting the competition that occurs among columnar dendritic grains. Validation of such competition has been carried out, following the analysis described in [Gandin et al. 1995]. The main differences in the results shown here with respect to [Gandin et al. 1995] were in the use of: (i) an instrumented one-dimensional solidification experiment of a cmsx6 nickel-base superalloy; (ii) more intensive Electron Back-Scattered Diffraction (EBSD) measurements for the characterization of the grain structure and texture evolution; and (iii) the use of the 3D CAFE model described above that accounts for non-uniform temperature situations. The temperature history during solidification was recorded using six thermocouples equally distant and localized at the center line of the 100 mm-height cylindrical ingot. These measurements gave the possibility to characterize precisely the cooling conditions. They were used in an inverse method for the determination of the heat flux as a function of time at the chill-metal interface. This heat flux was then applied as a boundary condition for the 3D CAFE calculation.

The nucleation boundary condition imposed at the bottom of the simulation domain was the grain density (set equal to the measured grain density), no nucleation undercooling and a random distribution of the crystallographic orientation of the grains. Figure 10.10 presents the result of the evolution of the \(\langle 100 \rangle\) pole figures recorded or calculated at 500 \(\mu\)m, 2 mm and 10 mm from the copper chill. Differences in the intensity are only because the size of the surface used to draw the simulated pole figures are larger than that used for the EBSD measurements. From an initially random distribution of the \(\langle 100 \rangle\) directions close to the chill, it can be seen that the \(\langle 100 \rangle\) fiber texture that forms after only a few millimeters is well reproduced by the CAFE model. Similar agreement is obtained when considering the evolution of the grain density as a function of the distance from the chill.

Applications of the 3D CAFE model to the prediction of grain structures in various solidification processes can be found in this book [Gauemann et al. 2003].

10.6 Conclusion

The phase-field or alternative methods, such as the pseudo-front tracking technique, are very powerful in predicting detailed solidification microstructures and several applications are given throughout this volume. Conceptually, a coupling of such methods with phase diagram computations for multi-component alloys would provide an ideal “virtual metallography” tool. However, this is not yet the case for several reasons: i) the computation power is still limited considering the ratio of the grain size over the interface thickness (typically 10^6). This limits the applications to two dimensions mostly; ii) the appearance of several phases requires to handle correctly triple points (e.g., using Lagrangian multiplier method [Nestler and Wheeler 2000]); iii) the competition between various phases depends upon nucleation as much as upon growth, and nucleation parameters are still empirically determined. This is why more pragmatic approaches, such as the mixture model presented here, are still interesting for the prediction of microstructures during solidification and homogenization.
Along the same line, the prediction of grain structures in solidified parts has to rely on approaches such as the CAFE model. While a lot of efforts have been made in the past to couple the heat flow equation (FE) with a robust CA algorithm for the nucleation-growth of dendritic envelopes, clear experimental evidences of the relationships between grain structures, convection and macrosegregation have defined new objectives for the CAFE model, namely the accounting of solute- and fluid-flow. Such objectives are nothing but the goals already achieved in the mid-nineties by Wang and Beckermann [Wang 1996] using a fully deterministic approach and a fully Eulerian multiphase representation. Preliminary steps to reach these goals have recently been published [Gandin et al. 1998, 2003]. These efforts should lead to a fully coupled CAFE model, with the advantages offered by the stochastic approach and the Lagrangian representation of the grain structures.

Acknowledgements

The development of the PFT and 3D CAFE models was founded by the Commission pour la Technologie et l’Innovation (Bern, CH - Grants #2861.1, #3666.1), by the Office Fédéral de l’Education et de la Science (Bern, CH - Grant #99.0644-1 / EU project Vircast #G5RD-CT-1999-00153) and by a consortium of several industries. The close collaboration and financial support of Calcom S.A. (Lausanne, CH) is also gratefully acknowledged.
References


11 Phase-Field Method Applied to Strain-dominated Microstructure Evolution during Solid-State Phase Transformations

Long-Qing Chen and Shenyang Hu

11.1 Introduction

Many of the important properties of a material can be engineered by controlling its solid state phase transformations and the accompanying microstructure evolution. Examples include the improvement of mechanical properties through solid state precipitation reactions in alloys such as Ni-based superalloys and age-hardened Al-alloys, the useful dielectric properties and electro-mechanical coupling effects by manipulating the phase transitions in ferroelectric crystals, the memory effect of shape-memory alloys by utilizing martensitic transformations. One of the key features common to solid state phase transformations is the fact that almost all microstructures produced at the early stages of phase transformations are coherent. In a coherent microstructure, the lattice directions and planes are continuous across the interfaces separating the parent and product phases or separating different orientation domains of the product phase. In order to maintain this lattice continuity, the lattice mismatch between the product and parent phases and among the orientation domains of the product phase must be accommodated by elastic displacements of atoms from their equilibrium lattice positions. Therefore, formation of coherent microstructures generates coherency elastic strain energy whose magnitude depends on the degree of lattice mismatch, the elastic properties of both the parent and product phases, and the shape and spatial distributions of coherent particles or domains (Mura 1982; Khachaturyan 1983). The effect of elastic strain energy on coherent precipitate morphology and its temporal evolution has been a subject of intensive theoretical study, and various simulation methods for modeling coherent precipitate microstructures were recently reviewed by Fratzl (Fratzl, Penrose et al. 1999).

This chapter reviews the applications of the phase-field approach to predicting the microstructure evolution during various solid state phase transformations including phase separation, order-disorder transformations, precipitation reactions, and ferroelastic and ferroelectric transformations. The emphasis is on the coupling between the phase-field and the elasticity equations as well as the role of strain in determining the microstructure stability and evolution. Different approaches for treating elastic inhomogeneity in the diffuse-interface description will be discussed.
11.2 Phenomenological Description of Solid State Phase Transformations

Before discussing the phase-field approach to microstructure evolution, let us briefly review the phenomenological description of solid state phase transformations. In principle, any solid state phase transformation can be characterized by physically well-defined order parameters that distinguish the parent and product phases. For example, the magnitude of a long-range order parameter in an order-disorder transformation is proportional to the intensity of a superlattice reflection corresponding to the ordered superstructure, and the order parameter for a ferroelectric phase transition is the local polarization vector which is related to the relative displacement of opposite-charged ions. Since solid state phase transformations always involve structural changes, strain, which is a measure of the strain state of a crystal during a phase transformation with respect to the parent phase, is also an order parameter. For most of the examples that will be discussed in this chapter, strain is a secondary order parameter in a sense that there is another primary physical order parameter such as a long-range order parameter or polarization which acts as the primary order parameter and characterizes the nature of a phase transformation.

In a phenomenological description, the local free energy function is expressed as a polynomial of order parameters using a conventional Landau-type of expansion. All the terms in the expansion are required to be invariant with respect to the symmetry operations of the high-temperature phase. The dependence of a phase transformation on strain is primarily determined by the coupling between the primary order parameter and strain, which has been discussed in great detail in an excellent textbook by Salje (Salje 1990). 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, \( \eta \) (Figure 11.1). 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 (\( \varepsilon_{ij} \)) and the order parameter (\( \eta \)) is linear-quadratic, the thermodynamics of the system can be described by

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

(11.1)

where \( \alpha_o, \beta, \) and \( \gamma \) are phenomenological coefficients which 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 \( \delta_{ij} \) the Kronecker-delta function defined as

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

Equation (11.1) 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. Throughout the chapter, the Einstein summation convention is employed i.e. repeated indices imply summation.
The critical temperature, i.e., the temperature below which the high temperature phase becomes unstable, is obtained from

\[
\frac{\partial^2 f}{\partial \eta^2} \bigg|_{\eta=0} = \alpha_o (T - T_c) - 2\eta \delta_{ij} \varepsilon_{ij} = 0 \tag{11.2}
\]

Therefore, \(T_c\) is the critical temperature at \(\varepsilon_{ij} = 0\). The equilibrium value for the order parameter as a function of temperature at \(\varepsilon_{ij} = 0\) is given by

\[
\eta_e = \pm \sqrt{\alpha_o (T - T_c) / \beta} \tag{11.3}
\]

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

\[
\varepsilon_{ij}^0 = s_{ijkl} \delta_{kl} \gamma \eta^2 \tag{11.4}
\]

which can be obtained by minimizing the free energy with respect to the strain, \(\varepsilon_{ij}\). In Equation (11.4), \(s_{ijkl}\) 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. Substituting the strain, \(\varepsilon_{ij}\), into expression (11.1), the free energy as a function of order parameter under the stress-free boundary condition can be expressed as

\[
f = \frac{1}{2} \alpha_o (T - T_c) \eta^2 + \frac{1}{4} \beta' \eta^4 \tag{11.5}
\]

where \(\beta' = \beta - 2\gamma^2 \delta_{ij} s_{ijkl} \delta_{kl}\). \(\alpha_o\) and \(\beta'\) should be measured under the stress-free boundary condition.

With Equation (11.5), the free energy at a given strain state can be rewritten as

\[
f = \frac{1}{2} \alpha_o (T - T_c) \eta^2 + \frac{1}{4} \beta' \eta^4 + \frac{1}{2} \lambda_{ijkl} (\varepsilon_{ij} - \varepsilon_{ij}^{0}) (\varepsilon_{kl} - \varepsilon_{kl}^{0}) \tag{11.6}
\]

Therefore, Equation (11.1) and (11.6) are equivalent. One can also formulate the local free energy as a function of order parameter at a given stress state,

\[
g = \frac{1}{2} \alpha_o (T - T_c) \eta^2 + \frac{1}{4} \beta' \eta^4 + \frac{1}{2} \lambda_{ijkl} (\varepsilon_{ij} - \varepsilon_{ij}^{0}) (\varepsilon_{kl} - \varepsilon_{kl}^{0}) - \sigma_{ij} \varepsilon_{ij} \tag{11.7}
\]
where $\sigma_{ij}$ is the stress tensor. Eliminating $\varepsilon_{ij}$ from Equation (11.7), we have the free energy as a function of stress,

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

where it is easy to see that an applied stress will affect the critical temperature (Equation (11.2)) since $\varepsilon_{ij}^o$ is a quadratic function of $\eta$ (Equation (11.4)). From the above discussion, it is shown that it is important to specify the mechanical boundary conditions for determining the coefficients in the free energy expression.

### 11.3 Phase-Field Model of Solid State Phase Transformations

To model the domain or microstructure evolution, one can simply extend the above phenomenological description of solid state phase transformations to inhomogeneous systems. In a heterogeneous microstructure, both the order parameters and the strain (or stress) are space-dependent. Using the same example as discussed in the last section, following the Cahn-Hilliard diffuse-interface theory (Cahn and Hilliard 1958), the total free energy for an inhomogeneous distribution of $\eta$ and $\varepsilon_{ij}$ can be written as

$$F = \int_v \left[ \frac{1}{2} \alpha_o (T - T_c) \eta^2 (r) + \frac{1}{4} \beta' \eta^4 (r) + \frac{k}{2} \left( \nabla \eta (r) \right)^2 + \frac{1}{2} \lambda_{ijkl} \left( \varepsilon_{ij} (r) - \varepsilon_{ij}^o (r) \right) \left( \varepsilon_{kl} (r) - \varepsilon_{kl}^o (r) \right) \right] dV$$

where the third term in the square bracket, called the gradient energy, is an extra term introduced for an inhomogeneous system. Its introduction automatically takes into account the contributions of domain-wall or interfacial energy to the thermodynamics. Therefore, the same free energy functional (11.9) can be used to determine both the bulk driving force for phase transformations and the interfacial energy force for domain or microstructure coarsening. The free energy in Equation (11.9) can be separated into two parts: chemical free energy and elastic energy, i.e.

$$F = F_{\text{chem}} + F_{\text{elast}}$$

where

$$F_{\text{chem}} = \int_v \left[ \frac{1}{2} \alpha_o (T - T_c) \eta^2 (r) + \frac{1}{4} \beta' \eta^4 (r) + \frac{k}{2} \left( \nabla \eta (r) \right)^2 \right] dV$$

and

$$F_{\text{elast}} = \int_v \frac{1}{2} \lambda_{ijkl} \varepsilon_{ij}^e (r) \varepsilon_{kl}^e (r) dV$$

$$= \int_v \left[ \frac{1}{2} \lambda_{ijkl} \left( \varepsilon_{ij} (r) - \varepsilon_{ij}^o (r) \right) \left( \varepsilon_{kl} (r) - \varepsilon_{kl}^o (r) \right) \right] dV$$

(11.12)
It is convenient to separate the strain, $\varepsilon_{ij}(r)$, to a sum of an average strain, $\bar{\varepsilon}_{ij}$, representing the macroscopic shape change of a crystal and the heterogeneous strain, $\delta\varepsilon_{ij}(r)$, whose volume average is zero (Khachaturyan 1983). In terms of $\bar{\varepsilon}_{ij}$ and $\delta\varepsilon_{ij}$, the elastic energy is given by

$$F_{\text{elast}} = \frac{1}{2} \int_V \lambda_{ijkl}(r) \left( \bar{\varepsilon}_{ij} + \delta\varepsilon_{ij}(r) \right) \left( \bar{\varepsilon}_{kl} + \delta\varepsilon_{kl}(r) \right) dV$$

$$= \frac{1}{2} \int_V \lambda_{ijkl} \bar{\varepsilon}_{ij} \bar{\varepsilon}_{kl} dV - \frac{1}{2} \int_V \lambda_{ijkl} \delta\varepsilon_{ij}(r) \bar{\varepsilon}_{kl} dV$$

$$+ \frac{1}{2} \int_V \lambda_{ijkl} \delta\varepsilon_{ij}(r) \delta\varepsilon_{kl}(r) dV$$

(11.13)

The homogeneous average strain is determined by the external strain or stress state (Khachatryan 1983; Li and Chen 1998). Last line in Equation (11.13) represents the elastic energy due to the heterogeneous deformation in a microstructure. Since not all strain components are independent, it is more convenient to express the free energy functional in terms of heterogeneous displacements, $u_i(r)$

$$F_{\text{hetero elast}} = \frac{1}{2} \int_V \lambda_{ijkl}(r) \left( \frac{\partial u_i(r)}{\partial r_j} \right) \left( \frac{\partial u_k(r)}{\partial r_l} \right) - 2 \left( \frac{\partial u_i(r)}{\partial r_j} \right) \varepsilon_{0}^{ij}(r) dV$$

(11.14)

It should be noted that from Equation (11.4), $\varepsilon_{0}^{ij}(r)$ is a function of $\eta(r)$, so it contains the information about the microstructure. During a microstructure evolution, both $\eta(r)$ and $u_i(r)$ vary with time. However, it can be generally assumed that the relaxation with respect to the elastic field is much faster than other kinetic processes governing the changes in the order parameter. Hence, the equilibrium elastic field is solved for a given distribution of the order parameter field, i.e. for a given microstructure. At equilibrium,

$$\frac{\delta F}{\delta u_i} = 0$$

(11.15)

Substituting Equation (11.14) into (11.15), we have

$$\partial \left[ \lambda_{ijkl}(r) \left( \frac{\partial u_i(r)}{\partial r_j} \right) - \varepsilon_{0}^{ij}(r) \right] = 0$$

(11.16)

or

$$\frac{\partial \left( \lambda_{ijkl}(r) \frac{\partial u_i(r)}{\partial r_j} \right)}{\partial r_j} = \frac{\partial [\lambda_{ijkl}(r) \varepsilon_{0}^{ij}(r)]}{\partial r_j}$$

(11.17)

For a given microstructure (described by $\varepsilon_{0}^{ij}(r)$), the calculation of elastic energy is reduced to finding the solution to Equation (11.17) with given mechanical boundary conditions.
11.4 Elastic Energy of a Microstructure

There exist many excellent references for describing methods of calculating elastic energy of single particles as well as microstructures, for example, the classical work of Eshelby (Eshelby 1956) and the excellent monograph by Khachaturyan (Khachaturyan 1983), just to name a couple. In this chapter, we will focus our discussions on the approaches that have been employed in phase-field simulations.

11.5 Bulk Microstructures with Periodic Boundary Conditions

Most of the existing phase-field simulations of coherent phase transformations have assumed an elastic homogeneity with periodic boundary conditions. In a homogeneous approximation, the elastic modulus tensor is a constant throughout a microstructure, then Equation (11.17) becomes

$$\lambda_{ijkl} \frac{\partial^2 u_k(r)}{\partial r_j \partial r_l} = \lambda_{ijkl} \frac{\partial \varepsilon^o_{kl}(r)}{\partial r_j}$$

(11.18)

which can be solved using the method of Fourier transforms as demonstrated by Khachaturyan more than thirty years ago (Khachaturyan and Shatalov 1969; Khachaturyan 1983),

$$\lambda_{ijkl} g_j g_l u_k(g) = -i \lambda_{ijkl} k_j \varepsilon^o_{kl}(g)$$

(11.19a)

or

$$u_k(g) = -i G_{ik}(n) \lambda_{ijkl} g_j \varepsilon^o_{kl}(g)$$

(11.19b)

where $i = \sqrt{-1}$, $G_{ik}(n)$ is the Green function tensor inverse to $G_{ik}(n) = \lambda_{ijkl} n_j n_l$, and $n = g / |g|$ is a unit directional vector in the reciprocal space. The inverse transform of $u_k(g)$ provides the heterogeneous displacement fields in real space from which the corresponding heterogeneous strain and stress can be obtained. Since the transformation strain, $\varepsilon^o_{ij}(r)$, is a function of field variables (Equation (11.4)), all the microstructure information is contained in $\varepsilon^o_{ij}(g)$, i.e., we have the elastic solution, $u_k(g)$, for any given arbitrary microstructure described by $\varepsilon^o_{ij}(r)$.

For systems with an inhomogeneous modulus, Equation (11.17) is non-linear and generally requires numerical solutions. In the case of weak elastic inhomogeneity, Onuki et al. solved the elasticity equation using a first-order approximation (Onuki 1989). For more general cases, Khachaturyan et al. developed an analytical solution for the elastic field in an elastically inhomogeneous system using a perturbation method and sharp-interface description (Khachaturyan, Semenovskaya et al. 1995). The strain energy is expressed as a sum of multiparticle interactions between finite elements of the constituent phases, pairwise, triplet, quadruplet and so on, the $n$-particle interaction energy being related to the $(n-2)^{th}$ order term in the Taylor expansion of the Green function with respect to the elastic modulus misfit. The order of approximation required for a given system depends on the desired accuracy and the
degree of elastic inhomogeneity. However, direct application of the analytical elastic energy expression to numerical simulation of coherent microstructure evolution is difficult since the elastic strain energy involves multi-dimensional integrals in both the real and Fourier spaces.

Hu and Chen proposed an efficient diffuse-interface phase-field model for elastically inhomogeneous systems (Hu and Chen 2001). They used an iterative approach for numerically solving the elastic equilibrium equation. To illustrate the method, the local elastic modulus tensor is expressed as

$$\lambda_{ijkl}(r) = \lambda_{ijkl}^0 - \lambda_{ijkl}^\prime(r)$$

(11.20)

where \( \lambda_{ijkl}^\prime \) is measured with respect to the reference value \( \lambda_{ijkl}^0 \). Higher order solutions for the elastic displacement, \( u_k^n(r) \), can be obtained by an iteration method.

$$\lambda_{ijkl}^0 \frac{\partial^2 u_k^n(r)}{\partial r_j \partial r_l} = \frac{\partial}{\partial r_j} \left( \lambda_{ijkl}^0 - \lambda_{ijkl}^\prime(r) \right) \bar{\varepsilon}_{kl}^0(r) + \frac{\partial}{\partial r_j} \left[ \lambda_{ijkl}^0 \frac{\partial u_k^{n-1}(r)}{\partial r_l} \right]$$

(11.21)

where \( u_k^{n-1}(r) \) is the solution from a lower-order approximation. The zero order solution \( u_k^0(r) \) is calculated using Equation (19). Therefore, the heterogeneous displacement fields in n-order approximation are given in Fourier space

$$u_k^n(g) = -iG_{ik}(n) g_j \left[ \lambda_{ijkl}^0 - \lambda_{ijkl}^\prime(r) \right] \bar{\varepsilon}_{lm}^0(r) + \lambda_{ijkl}^\prime(r) \frac{\partial u_k^{n-1}(r)}{\partial r_m} \right]$$

(11.22)

A comparison of the results obtained by the iteration method (Hu and Chen 2001) and the conjugate gradient method (Zhu, Chen et al. 2001) shows that the first-order approximation is reasonable if the elastic inhomogeneity is less than about 10%. When the stiffness of the harder phase is chosen as the reference, the iteration method is efficient even for systems with strongly elastic inhomogeneity. For example, only three iterations are required for systems with about 30% elastic inhomogeneity. Even in cases of voids and cracks where the elastic constants are zero, elastic solutions can also be obtained by the iteration method.

More recently, Wang et al. (Wang, Jin et al. 2002) developed a microelasticity theory of an elastically inhomogeneous solid. According to the Eshelby’s equivalent inclusion theory (Eshelby 1956), an elastically inhomogeneous solid can be viewed as an elastically homogeneous solid with an eigenstrain \( \bar{\varepsilon}_{ij}^0(r) \). As mentioned above, it is straightforward to obtain the elastic solution in an elastically homogeneous solid with an eigenstrain \( \bar{\varepsilon}_{ij}^0(r) \). Therefore, the problem to find the elastic solution in an elastically inhomogeneous solid is reduced to finding the equivalent eigenstrain \( \bar{\varepsilon}_{ij}^0(r) \). Notice the fact that the two equivalent elastic solids should have the same strain and stress fields, the elastic energy in the elastically inhomogeneous solid can be described as

$$E^{\text{inhom}} = \frac{1}{2} \int_v \left[ \lambda_{ijkl}^0 S_{klmn} \lambda_{pq}^0 - \lambda_{ijkl}^0 \right] \left[ \bar{\varepsilon}_{ij}^0(r) - \bar{\varepsilon}_{ij}^\prime(r) \right] \left[ \bar{\varepsilon}_{pq}^0(r) - \bar{\varepsilon}_{pq}^\prime(r) \right] dV$$

$$\frac{1}{2} \int_v \lambda_{ijkl}^0 \bar{\varepsilon}_{ij}^0(r) \bar{\varepsilon}_{kl}^0(r) dV - \bar{\varepsilon}_{ij} \int_v \lambda_{ijkl}^0 \bar{\varepsilon}_{kl}^0(r) dV + \frac{V}{2} \lambda_{ijkl}^0 \bar{\varepsilon}_{ij} \bar{\varepsilon}_{kl}$$

$$- \frac{1}{2} \int_v \frac{d^2 g}{(2\pi)^3} n_i \sigma_{ij}^0(g) G_{jk}(n) \sigma_{kl}^0(g)^* n_l$$
where $\varepsilon^{*}_{ij}(r)$ is related to the lattice mismatch, $(\sigma^{0}_{ij}(g))^{*}$ indicates the complex conjugate of $\sigma^{0}_{ij}(g) = \lambda^{0}_{ijkl} \varepsilon^{0}_{kl}(g)$, and $\varepsilon^{0}_{kl}(g)$ is the Fourier transform of $\varepsilon_{kl}(r)$. It is shown that the eigenstrain $\varepsilon^{0}_{kl}(r)$ minimizing the strain energy (11.21) determines the equilibrium strain of an elastically inhomogeneous solid. The minimizer $\varepsilon^{0}_{ij}(r)$ can be regarded as a variational parameter of the process whose driving force is $\delta E_{\text{inhom}} / \delta \varepsilon^{0}_{ij}(r)$. Therefore, the following evolution equation of $\varepsilon^{0}_{ij}(r)$ can be used to obtain the equivalent eigenstrain $\varepsilon^{0}_{ij}(r)$.

$$\frac{\partial \varepsilon^{0}_{ij}(r, t)}{\partial t} = -L_{ijkl} \frac{\delta E_{\text{inhom}}}{\delta \varepsilon^{0}_{kl}(r, t)}$$  \hfill (11.24)

where $L_{ijkl}$ is the positively defined kinetic coefficient, and $t$ is ‘time’ or the number of iterations. The stress fields of inclusions, voids and cracks obtained by the iteration equation (Hu and Chen 2001) and Equation (11.24) (Wang, Jin et al. 2002) are compared (Hu and Chen 2002), and the results show that these two methods yield similar efficiency and accuracy.

### 11.6 A Single Crystal Film with Surface and Substrate Constraint

Obtaining elasticity solutions for microstructures or domain structures in thin films are 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 constraint $\bar{\varepsilon}_{ij}(r)$, Li et al. showed that elastic solutions can be obtained analytically or at least semi-analytically (Li, Hu et al. 2001; Li, Hu et al. 2002) by assuming the same elastic constants between the film and substrate. As in the bulk, the distribution of eigenstrain or stress-free strain $\varepsilon^{0}_{ij}(r)$ includes the information on the film domain structure. $r = (r_1, r_2, r_3)$ is used as a rectangular coordinate originating at the film/surface interface with $r_3$ outward to film surface. The strain is separated into a sum of an average strain $\bar{\varepsilon}_{ij}$ representing the macroscopic shape change of a film and the heterogeneous strain $\delta \varepsilon_{ij}$. Let $\bar{\varepsilon}_{\alpha\beta}$ $(\alpha, \beta = 1, 2)$ represent the macroscopic shape deformation in the film plane, which are totally controlled by the sufficiently thick substrate. The other homogeneous strain, $\bar{\varepsilon}_{33}$, is assumed to satisfy $\bar{\sigma}_{33} = c_{33} \bar{\varepsilon}_{33} \delta E_{\text{inhom}}$. It should be pointed out that $\bar{\varepsilon}_{33}$ thus obtained is only part of the total shape deformation of the film.

The stress-free boundary condition at the top surface is given by

$$\sigma_{33}|_{r_3 = h_f} = 0,$$  \hfill (11.25)

where $h_f$ is the film thickness (Figure 11.2).

Since the elastic perturbation resulted from the heterogeneous strain $\delta \varepsilon_{ij}$ $(r)$ disappears in the substrate far from the film/substrate interface, one can use the following condition

$$u_i|_{r_3 = -h_s} = 0$$  \hfill (11.26)

to replace the constraint of the substrate. $h_s$ is the distance from the film/substrate interface into the substrate, beyond which the elastic deformation is ignored (Figure 11.2).
Therefore, the elastic solution can be obtained in two steps. First, Khachaturyan’s elasticity theory (Khachaturyan and Shatalov 1969; Khachaturyan 1983) is employed to solve Equation (11.18) in a three-dimensional (3-D) space. The corresponding displacement solution is labeled by \( u_A^{(r)} \). The next step is to find a solution, labeled by \( u_B^{(r)} \), in an infinite plate of thickness \( h = h_f + h_s \) satisfying the mechanical equilibrium equation

\[
\lambda_{ijkl} u_{B}^{k,l} = 0,
\]

with the following boundary conditions:

\[
\lambda_{ijkl} u_{B}^{k,l} \bigg|_{r_3 = h_f} = - \lambda_{ijkl} (u_{A}^{k,l} - \varepsilon_{0}^{k,l}) \bigg|_{r_3 = h_f},
\]

\[
u_{B}^{i} \bigg|_{r_3 = -h_s} = - u_{A}^{i} \bigg|_{r_3 = -h_s}.
\]

The Equations (11.27)–(11.29) can be solved by employing 2D Fourier transforms with the help of the Stroh formalism of anisotropic elasticity (Stroh 1962; Ting 1996). \( u_i (r) = u_i^{A} (r) + u_i^{B} (r) \) yields the solution for the boundary value problem of equations (18, 25-26). The total elastic energy can be calculated by Equation (11.13). For an extreme case, a monolayer with the constraint of an anisotropic substrate and an arbitrary distribution of eigenstrains, the elastic solution was derived by Suo (Suo and Lu 2000).

### 11.7 Elastic Coupling of Structural Defects and Phase Transformations

Recently, it is shown that the interaction between precipitates and structural defects such as dislocations can be described using the same approach as discussed above (Hu and Chen 2001;
Wang, Jin et al. 2001; Wang, Jin et al. 2001; Hu and Chen 2002). Assuming that the stress-free strain or eigen-strain caused by a structural defect is given by $\varepsilon_{ij}^d(r)$, its contribution to the elastic energy can be evaluated in exactly the same way as the transformation eigenstrains. The defect distribution can be rather arbitrary and there can be any number of different defects represented by different eigenstrains. Let’s consider a generic phase transformation with a transformation eigenstrain, $\varepsilon_{ij}^0(r)$. The elastic strain is given by

$$\varepsilon_{ij}^e(r) = \varepsilon_{ij}(r) - \varepsilon_{ij}^0(r) - \varepsilon_{ij}^d(r)$$

where $\varepsilon_{ij}(r)$ is the total strain. Therefore, one can obtain the elastic energy of a microstructure in the presence of structural defects by replacing $\varepsilon_{ij}^0(r)$ in Equation (11.12) using $\varepsilon_{ij}^0(r) + \varepsilon_{ij}^d(r)$. 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 (Hu and Chen 2001; Hu and Chen 2002) as well as dislocation motion (Wang, Jin et al. 2001; Wang, Jin et al. 2001)

### 11.8 Phase-Field Model Applied to Solid State Phase Transformations

Phase-field method has been to a wide variety of diffusional and diffusionless solid-state phase transformations (for previous brief reviews, see (Chen, Wolverton et al. 2001; Chen 2002)). Examples include isostructural phase separation, precipitation of an ordered intermetallic phase from a disordered matrix, cubic to tetragonal transformations, hexagonal to orthorhombic transformations, proper and improper martensitic transformations in single and polycrystals, ferroelectric transformations, and phase transformations on surfaces and in thin films.

### 11.9 Isostructural Phase Separation

The most studied type of coherent transformation process is the isostructural decomposition in a binary system. In this case, a two-phase microstructure can be described using only a single compositional field. Due to the compositional dependence of the lattice parameter, there is a lattice mismatch between regions with different compositions within an inhomogeneous solid solution. Most of the existing simulations assumed the Vegard’s law for the composition dependence of lattice parameter. In this case, the local transformation strain is given by

$$\varepsilon_{ij}^e(r) = \varepsilon_o \delta c(r) \delta_{ij}$$

where $\varepsilon_o = (1/a) (da/dc)$ is the composition expansion coefficient. The total free energy is given by

$$F = \int_V \left[ f(c) + \frac{\kappa}{2} (\nabla c)^2 + \frac{1}{2} \lambda_{ijkl}(r) \varepsilon_{ij}^e \varepsilon_{kl}^e \right] dV$$

(11.31)
where the first two terms are the local bulk free energy density, $f(c)$, and the composition gradient energy, respectively. The last term is the familiar elastic energy density. Existing phase-field simulations employed either a double-well potential for the free energy density function or a free energy directly from databases. For example, for a regular solution, $f(c)$ has the form,

$$f(c) = RT \left[ \varphi c(1 - c) + c \ln(c) + (1 - c) \ln(1 - c) \right]$$  \hspace{1cm} \text{(11.32)}$$

where $R$ is the ideal gas constant, $T$ is the absolute temperature, and $\varphi$ is a material constant.

For an ideal solution, $\varphi$ is equal to zero.

The temporal evolution of solute composition $c$ is described by the Cahn-Hilliard equation (Cahn 1961),

$$\frac{\partial c(r, t)}{\partial t} = \nabla M \frac{\delta F}{\delta c(r, t)} + \xi(r, t)$$  \hspace{1cm} \text{(11.33)}$$

where $M$ is a kinetic coefficient related to the solute diffusional mobility, and $\xi(r, t)$ is the noise term reflecting the thermal fluctuation.

With Cahn-Hilliard equation (11.33) or its microscopic version of diffusion equation (Khachaturyan 1983), the effects of elastic energy, elastic anisotropy, and elastic inhomogeneity on morphological changes during phase separation have been investigated. For example, assuming homogeneous modulus, Wang et al. (Wang, Chen et al. 1992; Wang, Chen et al. 1993) used a microscopic mean-field model to simulate the particle shape changes, particle splitting, as well as particle spatial alignment due to the long-range elastic interactions in two-dimensional (2-D). Zhang et al. (Zhang, Li et al. 1998) extended the work to 3-D using the continuum Cahn-Hilliard equation, and it was shown that it was possible for a single coherent particle to split to eight particles to relax the elastic energy at the expense of increasing the total interfacial energy between the precipitate and matrix. Very recently, the strain-induced morphological evolution of two-phase cubic alloys in 3-D was modeled (Ni, He et al. 2002). Lochte et al. studied the Guinier-Preston (GP) zone formation kinetics in Al-Cu alloys using the Cahn-Hilliard equation (Lochte et al. 2000). The morphological evolution during spinodal decomposition in Al-Zn alloys was investigated by Miyazaki et al. (Miyazaki, Koyama et al. 2001). Employing a first-order approximation, Onuki (Onuki 1989; Nishimori and Onuki 1990) modeled the effect of elastic inhomogeneity. They found that the precipitate shapes were strongly affected by both the elastic anisotropy and inhomogeneity. As the precipitate sizes grow, precipitates that are harder than the matrix change their shapes from spheres to other shapes to cancel the anisotropic elastic field produced by other precipitates. Shape changes of softer precipitates are much more drastic and take place even in the single precipitate case. The elastically softer phase forms a percolated network wrapping the harder precipitates, and the coarsening process is decelerated due to the lattice mismatch. Sagui et al. (Sagui, Orlikowski et al. 1998; Orlikowski, Sagui et al. 1999; Orlikowski, Sagui et al. 2000) carried out three-dimensional simulations of Ostwald ripening using Onuki’s model with or without an applied stress. For systems with a strong elastic inhomogeneity, Hu et al. (Hu and Chen 2001) employed the iteration method discussed in the last section for examining the shape dependence of isolated particles as well as the morphological dependence of a phase-separated multi-particle system on the degree of elastic inhomogeneity in elastically
11.10 Precipitation of Cubic Intermetallic Precipitates in a Cubic Matrix

Precipitation of ordered intermetallic phase from a disordered matrix is a fundamental process underlying the development of many advanced alloys such as Ni-base superalloys. It involves both ordering and compositional clustering, and thus both composition and order parameter fields are required to physically describe the precipitation process and the resulted two-phase microstructures. In this case, the free energy functional is typically written as

\[
E = \int_V \left[ f_0(c, \eta_i) + \frac{\kappa_0}{2} (\nabla c)^2 + \frac{3}{2} \sum_{i,j=1}^3 \kappa_{ij} (\nabla \eta_i \nabla \eta_j) + \frac{1}{2} \lambda_{ijkl}(\mathbf{r}) \varepsilon_{ij} \varepsilon_{kl} \right] dV \tag{11.34}
\]

While the composition fields evolve according to the Cahn-Hilliard equation, the long-range order parameter fields follow the Allen-Cahn equation (Allen and Cahn 1977),

\[
\frac{\partial \eta_i(\mathbf{r}, t)}{\partial t} = -L \frac{\delta E}{\delta \eta_i(\mathbf{r}, t)} + \zeta_i(\mathbf{r}, t) \tag{11.35}
\]

where \( L \) is a kinetic coefficient related to the atomic jump probability between two nearest neighbors and \( \zeta_i \) represents the thermal noise.

Using the phase-field equations (11.33) and (11.35) or their microscopic versions, early works on the effect of elastic strain on the precipitation of ordered intermetallic phases were
focused on simple model binary systems with two-types of antiphase domains (Chen, Wang et al. 1991; Sagui, Somoza et al. 1994). The transformation strain is dilatational, similar to the isostructural phase separation case, and contains the information about a microstructure through its dependence on composition, $c(r, t)$, or both $c(r, t)$ and $\eta_i(r, t)$. It was shown that for certain compositions and temperatures the precipitation process is preceded with a transient ordered phase, and the elastic energy contribution resulted in the alignment of ordered precipitates along the elastically soft directions during decomposition of the transient ordered phase. The aligned morphology finally formed the modulated structure found in many alloys with ordered precipitates separated by the disordered phase. The effect of elastic strain on the precipitate morphological evolution was also studied in a number of real alloys including Fe-Mo and Fe-Al-Co systems (Miyazaki 2001; Miyazaki 2002).

Figure 11.4: An example of microstructure evolution during precipitation of $\gamma'$ particles in a $\gamma$ matrix obtained from a phase-field simulation. The composition is Ni-13.6at%Al, and the temperature is 1023 K. The total system size is $640 \text{ nm} \times 640 \text{ nm} \times 640 \text{ nm}$. From left to right: $t = 2100 \text{ s}$; $t = 7500 \text{ s}$; $t = 14700 \text{ s}$; $t = 28200 \text{ s}$. 
One of the most studied examples using phase-field simulations is the precipitation process of a cubic intermetallic phase ($\gamma'$) from a cubic disordered matrix ($\gamma$) in Ni-Al binary alloys. (Wang and Khachaturyan 1995; Li and Chen 1997; Li and Chen 1998; Wang, Banerjee et al. 1998; Banerjee, Banerjee et al. 1999; Rubin and Khachaturyan 1999; Venugopalan and Chen 2000; Vaithyanathan and Chen 2002; Zhu, Liu et al. 2002). The precipitation process is described by a compositional field and a three-component order-parameter field, $\eta_i$. The local free energy function with expansion terms up to the fourth order is given by Equation (11.10) in Chapter Chen. Both 2D and 3D simulations have been carried out. An example of a $\gamma/\gamma'$ microstructure from a three-dimensional phase-field simulation is shown in Figure 11.4 (Zhu 2003). The initial state is a homogeneous disordered Ni-Al alloy, $\gamma$. Upon annealing within the two-phase ($\gamma + \gamma'$) field, the $\gamma'$ ordered phase particles nucleate (though the thermal noise at the initial stages of the precipitation process) and grow in the disordered matrix. In a given microstructure, the particle shapes of relatively small particles are nearly spherical. As the particle sizes increase, their shapes gradually change to cuboidal and subsequently plate-like for relatively large particles. Since the interfacial energy is assumed to be isotropic in this particular simulation for obtaining the microstructure, the cuboidal and plate-like shapes are entirely due to the anisotropic long-range elastic interactions. The particles tend to align along the crystallographically-soft (minimum elastic modulus) directions during the precipitation process and the degree of alignment increases as particle coarsening goes on, essentially the same phenomenon as in isostructural phase separation. The coarsening kinetics of ordered intermetallic precipitates with coherency stress was studied using a phase-field model in two dimensions (2D) (Vaithyanathan and Chen 2002). The average aspect ratio of the precipitates in the microstructure is found to increase with time and decrease with volume fraction. Contrary to all existing coarsening theories but consistent with experimental measurements on the coarsening kinetics of ordered $\gamma'$ precipitates in Ni-base superalloys (Maheshwari and Ardell 1992; Kim and Ardell 2000), it was found that the coarsening rate constant from the cubic growth law decreases as a function of volume fraction for small volume fractions (less than about 20%) and is constant for intermediate volume fractions (20–50%). It was also demonstrated that the two length scales in a stress-dominated coherent two-phase microstructure, the average precipitate size and average spacing between arrays of aligned precipitates, follow different growth exponents. Very recently, the morphological evolution of $\gamma'$ precipitates during a continuous cooling was modeled using the phase-field method (Wen, Simmons et al. 2003).

11.11  Structural Transformations Resulting in a Point Group Symmetry Reduction

A coherent phase transformation resulting in a reduction in the point-group symmetry produces several orientation variants. For example, in a cubic-tetragonal transformation, there are three possible orientation variants with the tetragonal axes along the [001], [010], and [001] directions. In the absence of any external field, all of them have the same probability to form in the parent cubic phase. The number of orientation variants increases to six if the crystalline point-group symmetry of the product phase is orthorhombic. A crystal lattice mis-
match between the parent and product phases as well as between different orientation domains of the product phase usually generates a significant elastic energy contribution to the total free energy of the system. On the other hand, the existence of several orientation variants of the product phase provides a configuration freedom to reduce the elastic energy. The elasticity theory for multidomain structures was developed by Khachaturyan (Khachaturyan and Shatalov 1969). In a phase-field model, different orientational variants are described by a set of structural order parameters. For example, for the cubic to tetragonal transformations, three structural parameter fields, \( \eta_1, \eta_2, \) and \( \eta_3 \) are required to represent the three possible orientation variants. There are two types of structural transformations, the so-called “improper” or “proper” ferroelastic transformations. For a proper ferroelastic transformation, the transformation strain is the primary order parameter. In the case of improper transformations, the strain is a secondary order parameter, and the eigenstrain, \( \varepsilon_{ij}^0(r) \), is usually proportional to the square of structural order parameters,

\[
\varepsilon_{ij}^0(r) = \sum_p \varepsilon_{ij}^0(p) \eta_p^2(r)
\]  

(11.36)

where \( \varepsilon_{ij}^0(p) \) is the eigenstrain associated with lattice mismatch of \( p \)th variant. Examples include many ferroelectric and order-disorder transformations. The elastic solutions as a function of structural parameter fields can easily be obtained by substituting the above eigenstrain expression (11.36) into the Equation (11.19a, (11.19b).

Chen et al. conducted a simulation of domain structure evolution during a cubic to tetragonal ordering transformation in a model 2-D binary alloy using the eigenstrain (Chen, Wang et al. 1992). It successfully predicted for the first time the detailed kinetics of formation of tweed morphologies and the tweed \( \rightarrow \) twin coarsening process driven by the transformation-induced elastic strains. A number of real examples have since been modeled using the phase-field approach. For example, the cubic-tetragonal phase transformation as well as the precipitation of tetragonal particles in partially stabilized \( \text{ZrO}_2 \) systems were studied (Wang, Wang et al. 1993; Fan and Chen 1995; Fan and Chen 1995; Wang, Wang et al. 1995; Wang, Chen et al. 1996). Other examples include the precipitation of dislike \( \theta' \) (\( \text{Al}_2\text{Cu} \)) precipitates in AlCu alloys (Li and Chen 1998; Vaithyanathan, Wulverton et al. 2002), and the effect of applied stress on coherent precipitation of multi-variant \( \text{Ti}_11\text{Ni}_{14} \) precipitates in TiNi shape memory alloys (Li and Chen 1997; Li and Chen 1998). It was demonstrated that an applied stress leads to a selective variant nucleation and growth. The morphological evolution during hydride precipitation was recently model by Ma et al. (Ma, Shi et al. 2002; Ma, Shi et al. 2002; Ma, Shi et al. 2002). Proville and Finel (Proville and Finel 2001) considered the precipitation reaction in \( \text{Al}_2\text{Zr} \) alloy in which a microstructure with cubic \( \text{Li}_2 \) precipitates appears as a transient state during the decomposition of a homogeneous disordered solid solution into a microstructure with tetragonal \( \text{DO}_23 \) precipitates embedded into a disordered matrix. La Boar and Khachaturyan (Le Bouar, Loiseau et al. 1998) showed that the chessboard-like microstructures observed in both \( \text{CuAu}_{(1-x)}\text{Pt}_x \) and \( \text{Co}_{39.5}\text{Pt}_{60.5} \) alloys were produced by strain-driven self-assembling of orientation variants of the precipitate phase, faceted by the planes normal to the elastically soft directions.

A multiple-order-parameter model for ordering on a binary hexagonal-close-packed (HCP) crystal structure was formulated, taking into account the underlying symmetries of
the HCP crystal in both the bulk and gradient energy terms of the free energy (Cahn, Han et al. 1999). The model was used to compute the orientation dependence of the structure and energy of antiphase boundaries in ordered CdMg structures using the diffuse-interface description. The effect of elastic interaction on the formation and dynamic evolution of multi-domain microstructures during a hexagonal to orthorhombic transformation in the absence and presence of an externally applied strain field was investigated by Wen et al. (Wen, Wang et al. 1999; Wen, Wang et al. 1999; Wen, Wang et al. 2000; Wen, Wang et al. 2002) using the phase-field approach. It is shown that a hexagonal to orthorhombic transformation results in a number of unique multi-domain structures such as the star patterns, compound star patterns, fan patterns, etc., which have been frequently observed experimentally in systems undergoing hexagonal to orthorhombic or similar transformations. The complex microstructural development during precipitation of a coherent orthorhombic phase (O-phase) from a $\alpha_{2}$ matrix (DO_{19}) in a Ti-Al-Nb system was also modeled using the phase-field approach (Wen, Wang et al. 2000; Wen, Wang et al. 2001).

Wang et al. conducted the first phase-field simulation of the improper cubic to tetragonal martensitic transformation in a single crystal (Wang and Khachaturyan 1997). Their simulations predicted that the nucleation of martensites in a perfect crystal occurs collectively to accommodate the coherency strain, e.g., the critical nuclei are formed by two internally twinned orientation variants, and the equilibrium martensite plate has its habit as an invariant plane. Artemev et al. (Artemev, Jin et al. 2001) extended the phase-field model to cubic to tetragonal proper martensitic transformations. In particular, they simulated the development of the martensitic transformation through nucleation, growth and coarsening of orientation variants in Fe-31%Ni alloy at different levels of undercooling. The effect of polycrystal (Jin, Artemev et al. 2001; Artemev, Jin et al. 2002), multilayer (Artemev, Wang et al. 2000), and applied stress (Artemev and Khachaturyan 2000) on the martensitic transformation were also studied. The simulations demonstrated that the martensitic transformation in polycrystals drastically differs from the transformation in single crystals and that the deformation curve of a polycrystalline material has a hysteretic effect. Figure 11.5 shows an example of domain structure evolution during a cubic $\rightarrow$ trigonal martensitic transformation in a polycrystalline Au-Cd alloy.

11.12 Ferroelectric Phase Transformations

In a ferroelectric phase transformation, the net local polarization in the high-temperature phase is zero and the low-temperature ferroelectric state has a spontaneous polarization. To describe a proper ferroelectric transition, the spontaneous polarization $P = (P_1, P_2, P_3)$ is chosen as the order parameter and strain is considered as a secondary order parameter. Therefore, in a phase-field model, a ferroelectric domain structure is described by the spatial distribution of the polarization with its temporal evolution described by the Time Dependent Ginzburg-Landau (TDGL) equations. For an ferroelectric crystal, in addition to the bulk chemical free energy, domain wall energy, and elastic energy considered in other structural transformations, electrostatic energy arising from the dipole-dipole interactions should be incorporated in the thermodynamics.
Figure 11.5: The hysteresis loop and the 3D microstructures obtained at different stresses for a cubic→trigonal martensitic transformation in a Au-Cd polycrystal.
The eigenstrain for a given ferroelectric transformation is related to the spontaneous polarization. For example, for a central symmetric parent cubic phase, the eigenstrain can be written as

$$\varepsilon_{ij}^0 = Q_{ijkl}P_kP_l,$$

(11.37)

where $Q_{ijkl}$ is the electrostrictive coefficient tensor. Using the eigenstrain (11.37) in Equation (11.19b) and (11.13), the elastic energy is then a functional of the polarization field.

Both 2- and 3-D computer simulations of ferroelectric domain formation and evolution have been performed for a cubic-to-tetragonal ferroelectric phase transition (Nambu and Sagala 1994; Hu and Chen 1997; Hu and Chen 1998). It is shown that the initial stage of the transition during the annealing of a quenched cubic paraelectric phase involves the nucleation and growth of the ferroelectric domains, followed by the domain coarsening leading to the formation of 90 degrees and 180 degrees domain structures. It is shown that the non-local elastic interaction is critical to the formation of twin structure and the dipole-dipole interaction is responsible for the head-to-tail arrangements of dipoles at twin boundaries. It is found that when there are no surface charges to compensate the Lorentz field due to the polarization charges, and if the system is mechanically clamped, both 90 degrees and 180 degrees domains are thermodynamically stable. The effect of randomly distributed defects on ferroelectric phase transition was considered by Semenovskaya et al. (Semenovskaya and Khachaturyan 1998; Semenovskaya and Khachaturyan 1998) using a 2D phase-field model. They introduced the effect of defects on the critical phase transition temperatures by modifying the second-order Landau coefficients. It is shown that the electrostatic dipole-dipole interaction may produce the same “martensite-like” effects as the strain-induced interaction in the martensitic transition. These effects are the formation of the mixed two-phase equilibrium between the ferroelectric and paraelectric phases, the thermal hysteresis, and the loss of ergodicity. Effect of dipolar defects on domain nucleation during domain reversal was recently modeled by Ahluwalia and Cao (Ahluwalia and Cao 2001).

11.13 Phase Transformation in a Reduced Dimensions: Thin Films and Surfaces

The presence of a stress-free surface and a substrate constraint for a thin film can dramatically change its phase transformation behavior as well as the stability of resulted domain structures and their evolution compared to those in the corresponding bulk. A well studied example, both experimentally and theoretically, is the ferroelectric phase transformations and domain structures in thin films. Recently, Li et al. (Li, Hu et al. 2001; Li, Hu et al. 2002) employed the elastic solutions discussed in the last section to study the stability and evolution of 3-D ferroelectric domain structures in a thin film constrained by a substrate. Without making any a priori assumptions with regard to the possible domain structures that might appear under a given temperature and substrate constraint, it is able to predict not only the effect of substrate constraint on phase transition temperatures and the volume fractions of orientation domains, but also the detailed domain structures and their temporal evolution during a ferroelectric transition. A number of thin film systems have been studied, including PbTiO$_3$ (Li, Hu et al. 2001;
In all these examples, the high-temperature paraelectric phase is cubic and heteroepitaxially grown on a cubic substrate along the (001) orientation. The film undergoes ferroelectric phase transitions from the cubic paraelectric phase when it is cooled below the Curie temperature. The type of ferroelectric phases and the number of domain variants obtained depend on the temperature and substrate constraint, and in the case of PZT, also on composition. Ferroelectric domain stability in constrained films was also recently studied by Slutsker and Roytburd using the phase-field approach (Slutsker, Artemev et al. 2002).

As an example, a phase stability map of temperature, $T$, versus the mole fraction, $x$, with a substrate constraint, $e_0 = 0.005$, for a PZT film is shown in Figure 11.6. For comparison, the phase diagram for the bulk is displayed as dashed lines in the same figure. The material constants are taken from the works of Haun et al. The coefficients depend on the mole fraction $x$ of PbTiO$_3$ in PZT. It is seen that the phase stability map for the PZT film constrained by a substrate is drastically different from that for the corresponding bulk obtained under the stress-free conditions. For instance, there are only two stable ferroelectric phases, i.e. tetragonal phase and rhombohedral phase in the bulk whereas in the tensile constrained film, an orthorhombic phase of $|P_1| = |P_2| \neq 0$, $P_3 = 0$ becomes stable in the region with small $x$ at high temperatures.

The rhombohedral phase ($|P_1| = |P_2| = |P_3| \neq 0$) is distorted, i.e., $|P_1| = |P_2| \neq 0$, $P_3 = 0$, in the film due to the biaxial substrate constraint. An example of a distorted rhombohedral domain structure is shown in Figure 11.7. Very recently, it is demonstrated that one can also incorporate the effect of misfit dislocations on the domain nucleation and spatial distribution in the phase-field model for thin films (Hu, Li et al. 2003) as well as the effect of electric boundary conditions (Li, Hu et al. 2002).

There are a number of other applications of the phase-field method to the morphological pattern formation in thin films and on surfaces. For example, Leo et al. (Leo and Johnson 2001) proposed a 2-D phase-field model using the Cahn-Hilliard equation to study the spinodal decomposition and long-time coarsening behavior in a thin film attached to a compliant substrate. Seol et al. studied the spinodal decomposition (Seol, Hu et al. 2003) and the cubic to tetragonal martensitic transformation (Seol, Hu et al. 2002) in constrained films with a three dimensional phase-field model. The simulations demonstrated that the substrate constraint and surface stress free remarkably affect the nucleation process as well as final domain structure (Li, Hu et al. 2002).

There is a number of other applications of the phase-field method to the morphological pattern formation in thin films and on surfaces. For example, Leo et al. (Leo and Johnson 2001) proposed a 2-D phase-field model using the Cahn-Hilliard equation to study the spinodal decomposition and long-time coarsening behavior in a thin film attached to a compliant substrate. Seol et al. studied the spinodal decomposition (Seol, Hu et al. 2003) and the cubic to tetragonal martensitic transformation (Seol, Hu et al. 2002) in constrained films with a three dimensional phase-field model. The simulations demonstrated that the substrate constraint and surface stress free remarkably affect the nucleation process as well as final domain structure (Suo et al. 2001; Suo et al. 2002; Suo et al. 2002) investigated the self-assembly dynamics of a two-phase monolayer on an elastic substrate by incorporating an anisotropic surface stress into a phase-field model of spinodal decomposition. The competition between short-range attractive interactions and long-range repulsive elastic interactions results in nanoscale morphological patterns such as interwoven stripes, parallel stripes, triangular lattice of dots, and herringbone structures. The spatial scale of the patterns may be in the range 1–100 nm, and stable against coarsening on annealing. Recently, a phase-field model for the dynamics of steps and island on crystal surfaces are developed by Yeon et al. (Yeon, Cha et al. 2003). Their model is based on the Ginzburg-Landau approach and reduces to the Gibbs-Thompson equation at the step edges. The elastic field, which is generated by the surface force, is calculated using elastic surface Green function under the assumption that the atomic displacement field decays exponentially from the surface.
11.14 Summary

This Chapter reviews the applications of the phase-field approach to solid state phase transformations. The focus is on the strain contributions to coherent domain and phase microstructure evolution. It is shown that for phase transformations in bulk single crystals, it is possible to include the elastic strain effect on microstructure evolution for any arbitrary elastic anisotropy and large elastic inhomogeneity. It is also rather straightforward to include the effect of the elastic fields produced from structural defects such as dislocations. Various domain wall orientations and habit planes in coherent microstructures can be automatically predicted from phase-field simulations. These models can be extended to polycrystalline materials. For thin films, current phase-field models for phase transformations are limited to flat surface and interfaces with the film subject to a substrate constraint. For the more general cases with simultaneous evolving microstructures in the film and evolving surface topologies, phase-field models are yet to be developed. Recently, it was shown that the same models for phase trans-
11 Phase-Field Method Applied to Strain-dominated Microstructure Evolution

![Image](image.png)

**Figure 11.7:** (a) Ferroelectric domain structure obtained from a phase-field simulation for a PbZr$_{1-x}$Ti$_x$O$_3$ ($x=0.10$) thin film at room temperature under a substrate tensile constraint $e_{01}=0.005$. The ferroelectric phases shown are distorted rhombohedral phases with polarization $\mathbf{P} = (P_1, P_2, P_3)$. (b) Schematic illustration of the corresponding polarization orientations and domain wall orientations.

Formations can be extended to modeling dislocation dynamics (Hu and Chen 2001; Jin and Khachaturyan 2001; Wang, Jin et al. 2001; Wang, Jin et al. 2001; Koslowski, Cuitino et al. 2002; Hu, Li et al. 2003; Rodney, Le Bouar et al. 2003; Shen and Wang 2003) and crack propagation (Aranson, Kalatsky et al. 2000; Jin, Wang et al. 2001; Karma, Kessler et al. 2001; Jin, Wang et al. 2003). Quantitative study of nucleation kinetics has also recently been attempted using the phase-field approach (Roy, Rickman et al. 1998; Simmons, Shen et al. 2000; Granasy, Borzsonyi et al. 2002; Granasy, Borzsonyi et al. 2002). One of the limitations for current phase-field models for solid phase transformations is the fact that the length scale
is limited to nanoscale system size due to the physical nature of the field variables and the actual physical width of interfaces. It is recently demonstrated that it is possible to increase the length scale by introducing artificial phase fields (Cha, Yeon et al. 2003) who performed a phase-field simulation of particle splitting in nickel-based super-alloys with the elastic solution obtained by the iteration method (Hu and Chen 2001), and by Zhu et al. (Zhu, Wang et al. 2003) and Hu et al. (Hu 2003) for precipitate microstructure evolution in Ni- and Al-alloys.

References


References


References


12 Irregular Cellular Automata Modeling of Grain Growth

Koenraad G. F. Janssens

12.1 Introduction

Soon after the first discovery and development of cellular automata their potential for modeling the evolution of the microstructure of polycrystalline materials was realized by many different researchers (Hesselbarth and Göbel (1991), Pezzee and Dunand (1994), Davies (1995)). A detailed overview and history can be found in Chapter 3. An essential aspect of the cellular automata approach is its discretization of space by mapping the grains in the microstructure onto a lattice of points. Conventional lattice grids usually have a square or hexagonal distribution in the plane, and a cube or tetrahedral distribution in three dimensions. A known drawback (e.g. Rollet (1997)) of such symmetrical cell lattices geometries is that the local morphology of the grain boundaries is determined by the choice of the neighborhood of each cell. Even in the case of probabilistic cellular automata, where the interaction between cells is energetically controlled as is the case in Monte Carlo methods (See Chapter 3), the problem is not entirely solved (E. A. Holm and C. C. Battaile, 2001). The issue is that grain boundary energy and mobility depends on crystallographic misorientation between neighboring grains. A symmetrical grid also discretizes the grain boundary surface in such a way that only a very limited number of orientations are represented. Irregular or random grid cellular automata solve this problem by local variation of the cell lattice geometry. The first publications on irregular cellular automata known to the author are by Thieme-Marti (1999a) in technical science and by Flache (2001) in general. Other research groups have published alternative approaches to solving this problem, e.g. by predefining the geometry of the grains (Q. Y. Xu and B. C. Liu, 2001) or by implementing a hybrid Monte Carlo-cellular automata model (A. D. Rollet and D. Raabe, 2001).

12.2 Irregular Cellular Automata

12.2.1 The Concept

The concept behind irregular cellular automata is given away by its name: instead of using a periodic distribution, cells are distributed randomly (Janssens, 2003a). Apart from this random distribution irregular cellular automata do not need to be much different from conventional ones, at least not if one uses the Voronoi cells computed starting from a random distribution of points as implemented by A. Flache (2001) and illustrated in Figure 12.1. The neighborhood of a cell can then for example be defined to include each cell with which it shares an interfacial
12.3 Irregular Shapeless Cellular Automata for Grain Growth

plane or edge, or including additional cells that are further away using another definition. Once the definition of neighborhood is clear, irregular cellular automata can be used in exactly the same way as regular ones (see Chapter 3).

Figure 12.1: An example of a two dimensional irregular cellular automata distribution based on the Voronoi cells computed from a random distribution of points. The neighborhood for one cell using the definition as given in the text is indicated.

12.2.2 Shapeless or Point Cellular Automata

Shapeless or point cells are alternatives to Voronoi based cell shapes. Shapeless cell shapes are not computed explicitly; only a point position is used by the cellular automata algorithm. Besides not having to compute the Voronoi cells, which can become computationally intensive in three dimensions, shapeless cells have the advantage of being much more flexible in their definition of neighborhood. A feasible instantiation of shapeless cellular automata could go like this:

- The starting point is a distribution of points with an average area or volume density. Note that the distribution may be equidistant, like in conventional automata, or random/irregular.

- Each point represents a cell. The volume of a cell can either be an average area or volume that follows directly from the average density, or a more precise definition depending on the local cell density.

- The neighborhood definition consists of a shape and its according size specification, e.g. a sphere and its radius. The neighborhood of a cell includes all point cells which are contained within this sphere. As only point coordinates are used, shapes and sizes can easily be varied, even locally. Position dependent variation of the neighborhood is essential for its application to grain growth, as will become clear in the rest of this chapter.

With this definition, shapeless cellular automata can be used as conventional ones.

12.3 Irregular Shapeless Cellular Automata for Grain Growth

This section presents a detailed description that clarifies how an irregular shapeless cellular automaton can be used to simulate grain growth (Janssens et al., 2002). The reader is assumed to be familiar with the materials science of grain growth; see for example F. J. Humphreys and M. Hatherly (1996), G. Gottstein and L. S. Shvindlerman (1999) and J. W. Martin et al.
In a broad sense, grain growth can be defined as the motion of grain boundary surfaces under the influence of a driving pressure. As originally published by J. E. Burke and D. Turnbull (1952), the velocity $v$ of a grain boundary is usually assumed to depend on the driving force $p$ and the grain boundary mobility $m$ in the following simple equation:

$$v = mp$$  \hspace{1cm} (12.1)

In the context of modeling grain growth using cellular automata it is now necessary to differentiate between the three different grain growth modes listed below.

- **Curvature driven grain growth**, in which it is the energy stored in grain boundaries and grain boundary junctions that drives their motion.
- **Non-curvature driven grain growth**, in which energy is present in another form (stored or externally induced) in the microstructure and the energy of the grain boundaries is negligible.
- **Finally a mixed mode**, in which grain boundary and additional energy lead to comparable driving pressures, and must therefore both be taken into account.

Depending on the grain growth mode, a different modification of cellular automata can be used to construct a model based on Equation (12.1).

### 12.3.1 Curvature Driven Grain Growth

When the motion of grain boundaries is governed by their curvature the driving pressure can be mathematically represented in the most simplified manner by:

$$p = \frac{2\gamma}{R}$$  \hspace{1cm} (12.2)

with $\gamma$ the grain boundary energy and $R$ the local radius of the grain boundary curvature.

The issue here is the representation of surfaces in the cellular automata model: in automata where the shape of the cells is used, a highly discontinuous and faceted grain boundary is the result; in shapeless automata no surface is explicitly defined. One way to solve this is to compute the surface and derive its local curvature, but irregular shapeless automata allow the implementation of a much more elegant solution, similar to an approach published by J. W. Bullard et al. (1995).

### 12.3.1.1 Cell State and Neighborhood Definition and the Space-Time Connection

Making an analogy to an atom on a grain boundary, it is clear that like the cell, the atom has no knowledge of any surface, let alone any surface tension. The key question to pose here is why atoms preferentially jump to the convex side of a curved grain boundary. The combination of atom movement due to thermal vibration and a simple geometric consideration leads to a plausible answer to this question. When an atom jumps from one lattice position to
a neighboring one (independent of how exactly such a jump is accomplished), it has a limited number of possible positions where it can end up. Let us assume that it can end up in any lattice position within a radius $r$ of its original position. An atom near a flat grain boundary has an equal probability of ending up in either grain at the boundary. However, in the case of a curved boundary, the probability for the atom to end up on the convex side of the boundary is higher, and therefore the grain boundary will tend to move towards its concave side. This pseudo-physics thought-experiment can be straightforwardly transposed to the rules by which the cellular automaton changes state, which can be formulated like this:

1. Determine the cells which are located within the neighborhood.
2. Determine which grains are present in the neighborhood and how many neighborhood cells belong to each of these grains.
3. The probability that the cell changes state belonging to grain $X$ is proportional to the number of cells of grain $X$ that are located in its neighborhood. This is algorithmically very efficient as it is the equivalent of randomly picking a cell in its neighborhood and changing state to the grain to which the picked cell belongs.

![Figure 12.2: A shapeless cell on a curved grain boundary and its spherical neighborhood.](image)

As an example, consider a point cell on a curved grain boundary and its neighborhood as depicted in Figure 12.2. The cells above the grain boundary belong to grain $A$, the cells below to grain $B$. Assuming the center cell, located on the grain boundary also belongs to grain $B$, its probability to end up in grain $A$ is $6/11$, so on average the grain boundary will move down as expected (statistically more cells should be found in the larger volume partition).

Does this state change rule mimic the physics in Equation (12.2)? The answer to these questions is surprisingly simple. Start with abstracting the cells in Figure 12.2 as volume partitions $V_1$ and $V_2$ of the neighborhood of a cell. A point cell at the grain boundary will jump to grain $A$ with a probability $P_A = V_1/(V_1 + V_2)$, and to grain $B$ with a probability equal to $P_B = V_2/(V_1 + V_2)$. This means that at every increment the grain boundary moves one layer of cells into grain $B$, and that with a probability

$$P = \frac{V_1 - V_2}{V}$$

with $V = V_1 + V_2$. What is interesting is that using $r_{nbh}$ as the radius of the neighborhood, Equation (12.3) computes to

$$P = \frac{3r_{nbh}}{8R}$$

(12.4)
in which the grain boundary curvature radius $R$ shows up again. Given a volume cell density, the average volume that a cell represents can be computed. Approximating the cell volume with an equivalent spherical volume, an equivalent radius $r_c$ can be computed, from which in turn the velocity of the grain boundary in the automaton model can be derived, namely $v = 2P r_c / \Delta t$ with $\Delta t$ the time step for a computational increment. Combined with Equations (12.2) and (12.4) this results in

$$
\frac{2\gamma}{R} = \frac{3r_{nhb} r_c}{4R \Delta t}
$$

Equation (12.5)

Given a time step $\Delta t$ for the computation one must thus set

$$
r_{nhb} = \frac{8m\gamma \Delta t}{3r_c}
$$

Equation (12.6)

to comply with the theoretical grain growth equation (12.2). In other words the local variation of the grain boundary curvature is intrinsically accounted for by the cellular automaton, and the curvature need not be calculated. In addition, local variation of the grain boundary mobility $m$ and energy $\gamma$ can be straightforwardly introduced by locally varying the neighborhood radius $r_{nhb}$.

Equation (12.6) constitutes a relation between $r_{nhb}$, $\Delta t$ and $r_c$ ($r_c$ fixes the spatial resolution); note that this link forces the model to be an explicit one by connecting the time increment to the spatial resolution. Also, $r_{nhb}$ cannot be chosen entirely freely. Its maximum value is limited by considerations of computing resources: a large neighborhood increases the number of cells that one must keep track of. The minimum value of $r_{nhb}$ is also limited, as for too small a value the neighborhood contains no cells. Though a burden to the programmer, the latter limitation can be overcome by setting a minimal value for $r_{nhb}$. In case the value needed to satisfy Equation (12.6) is smaller, then only perform the state change with a proportional probability.

### 12.3.1.2 Triple Junctions

Equation (12.6) also works for the case of triple or multiple junctions. Consider the triple junctions depicted in Figure 12.3. The vertical grain boundary constitutes a driving force for the triple junction to move down. The cellular automaton now determines for each cell to which grain boundary it belongs and then computes the neighborhood size using Equation (12.6). It should be clear from the figure that cells close to the triple junction have a higher probability to end up in grain $A$ than in grain $B$ or $C$. Variable grain boundary energy and mobility can also be taken into account as shown in the right part of Figure 12.3: a lower energy for the boundary between grains $C$ and $D$ leads to a smaller neighborhood and thus to a smaller number of cells that are able to change to grain $A$ in a time increment (i.e. they must be closer to the triple junction to be able to make the jump), hence a lower grain boundary velocity.

Figure 12.4 shows a simple example of two triple junctions in an evolving microstructure. Their velocity is constant as the driving pressure, namely the energy stored in the horizontal grain boundaries, is constant.

---

1 $r_{nhb}$ is systematically underestimated in a discrete spatial representation as used in cellular automata, which introduces an additional correction factor into Equation (12.6).
12.3 Irregular Shapeless Cellular Automata for Grain Growth

![Diagram of grain growth](image)

**Figure 12.3:** An unstable triple junction configuration, for (left) equal grain boundary energy and (right) a relatively lower energy for the boundary between grains C and D. Neighborhoods for cells at different positions are shown.

![Diagram of triple junction motion](image)

**Figure 12.4:** The evolution of a microstructure (symmetrical boundary conditions) showing the motion of triple junctions. The diagram on the right illustrates the velocity of the triple junctions is constant (arbitrary physical constants were used).

12.3.2 In the Presence of Additional Driving Forces

12.3.2.1 Large Driving Forces and an Inverse Update Scheme

In conventional cellular automata all the cell states are updated in every step of the computation. This computing scheme requires that in each increment a new status must be computed for all cells. In the presence of a large driving pressure, compared to which the driving pressure from the grain boundary curvature can be neglected, the direction of motion of the grain boundaries is predetermined and unchanging. This is the case for transformations where nucleated grains grow into an existing microstructure, including recrystallization and certain phase transformations, i.e. in those microstructure transformations concerning the growth of nucleated grains into an existing microstructure. In the presence of such a large driving pressure updating all cells is not necessary. To optimize the algorithm the cell state can differentiate between its location being behind (or in a newly nucleated grain, i.e. in-nucleated) or in front the moving grain boundary. The most straightforward algorithm optimization that can be made is a simple preselection of cells: those having a state labeled behind the moving boundary are excluded from the updating scheme performing the recrystallization. An alterna-
tive update rule (Janssens, 2003a) approaches the update problem from another direction: if a cell has a in-nucleated state, then annex all the cells in its neighborhood to the grain to which the current cell belongs. By introducing this rule the numerical method is taken outside the normal definition of cellular automata and actually shows more similarities with an Eden-type model (M. Eden (1961) and R. Jullien and R. Botet (1985)). Looking at the modified cellular automata more closely, it becomes clear that it is no longer possible to update all the cells in parallel: a cell with state not-in-nucleated in the neighborhood of two different nucleated grains could be annexed by the first or by the second grain. Randomizing the order in which the cells are evaluated for neighborhood updating averages out any influence on the statistical outcome of a simulation. An additional problem could be that, again depending on the evaluation order, a cell annexed in an update step could in turn annex the cells in its own neighborhood within the same step. Fortunately this kind of cascading can be circumvented, again at the expense of the programmer’s nerves, by excluding all cells that were not annexed before the current step from being used. To do so an additional state is introduced: completed. At all times cells now have one of three states: in-existing-grain, in-nucleated-grain or completed; the updating rule is now: to consider all cells with state in-nucleated-grain (one after the other in a random order), to change their state to completed and to change the state of all in-existing-grain cells in their neighborhood to in-nucleated-grain.

12.3.2.2 Small Driving Forces and Drag Forces

The combination of a grain boundary driving pressure with an additional driving or drag pressure can be solved using two different approaches. The first option is to combine the inverse and the conventional update rule within the same time step, more precisely by performing one after the other. Such sequencing has been implemented with success by many other authors, e.g. D. P. Landau and K. Binder (2001) or A. D. Rollet and D. Raabe (2001) to name just two, more recent implementations. The second option is to modify Equation (12.3) to include an additional driving force (Janssens, 2003a):

\[ P = \frac{V_1 - V_2}{V(1 + f)} \pm \frac{f}{1 + f} \]  

(12.7)

This modification adds an imaginary volume to the existing neighborhood of size \( fV \), which represents an additional driving force \( p_a \) and can be straightforwardly introduced into the algorithm at the time the neighborhood is evaluated. Equation (12.7) guarantees that the absolute value of \( P \) is between zero and one, independent of the value chosen for \( f \). The sign in front of the second term depends on the direction of the additional driving force \( p_a \) and is positive if acting in the same direction as the curvature pressure does. Equation (12.5) then becomes:

\[ m \left( \frac{2\gamma}{R} \pm p_a \right) = v = \frac{2r_c}{\Delta t} \left( \frac{3r_{nbh}}{8R(1 + f)} \pm \frac{f}{1 + f} \right) \]  

(12.8)

which for a given time step \( \Delta t \) holds when

\[ r_{nbh} = \frac{8m\gamma(1 + f)\Delta t}{3r_c} \quad \text{and} \quad f = \frac{3p_ar_{nbh}}{16\gamma} \]  

(12.9)
12.4 A Qualitative Example: Static Annealing of a Cold Rolled Steel

Consider a common, low carbon, Al-killed type steel, containing a known distribution of AlN particles and some remaining N in solution. The steel is cold rolled with a reduction of 70%, which textures the steel and introduces deformation energy in the form of dislocations. Subsequently the material is annealed to obtain deformable metal that is useful for e.g. deep drawing applications (Janssens et al., 2003b).

Of interest is a model that predicts (or more modestly: helps us understand) the evolution of the microstructure, starting from specific knowledge concerning the microstructure in its pre-rolled condition. The model should include information on the evolution of the following microstructure aspects:

- the size and shape distribution of the grains,
- the stored deformation energy, which for simplicity is here reduced to the dislocation density,
- and finally the texture.

Such a model can be split up in two sequences, the deformation phase and the annealing phase, the latter involving concurrent recovery, recrystallization, grain growth and precipitation in the ferrite phase of iron.

12.4.1 The Deformation Model

Though the deformation model is outside the scope of this chapter, some information is given here as it is essential to understand the annealing part of the model. While other authors have published research combining finite element modeling of the deformation and combining it with Potts or cellular automata (B. Radhakrishnan et al. (1998), C. C. Battaile et al. (1999), O. Grong and H.R. Shercifff (2002) and K. Mori et al. (2003)), the one used in this example is a statistical, modified Taylor-type model (S. Thieme-Marti (1999b) and F. Vanini et al. (2003)). Starting from average values for grain size and shape, an experimentally measured orientation distribution function (ODF) and a collection of tensile curves, this model sets up an initial microstructure model. The microstructure is partitioned into sub-volumes, each representing a constant crystallographic orientation and a constant dislocation density, the...
values of which are fitted to the experimental information. Given the deformation tensor, the model then computes the evolution of the microstructure, including:

- the orientation distribution of the deformed material,
- a per-orientation-volume dislocation density distribution,
- the average size and shape of the deformed grains.

This data is the starting point for the annealing phase of the model.

12.4.2 The Annealing Model

In addition to the data obtained from the deformation simulation, before the annealing model can be set up further information and approximations are necessary:

- In this example recovery of the microstructure is reduced to an exponentially decreasing, time and temperature dependent dislocation annihilation model which predicts the average dislocation density only. Important is that the dislocation density in the material diminishes in time, which influences the kinetics of the other microstructure processes involved. See for example Section 4.2.1 in Martin et al. (1997) and references therein for more details.

- Recrystallization starts with nucleation of new grains. The crystallographic orientation of these grains is taken from the ones present in the texture of the deformed microstructure. The kinetics (incubation time and a non-constant nucleation rate) of the nucleation process are introduced with a purely phenomenological model, which is dependent on the temperature and the local dislocation density. In addition the size-distribution of AlN particles that was present in the pre-deformed microstructure is taken into account by assuming an increased local dislocation density (but no lattice reorientation) at the location of the particle. The incubation time and nucleation rate are nonlinear functions with an inverse response to the dislocation density and the temperature.

- Grain boundary energy and mobility are assumed constant and thus independent of crystallographic misorientation. Solute drag is not taken into account explicitly, but assumed to be included in the values for grain boundary mobility (the values of which are not known quantitatively and must be calibrated to the experiment).

- The evolving distribution of AlN particles precipitating during the annealing process are simulated using an additional statistical model developed by E. Kozeschnik and B. Buchmayr (2001). The presence of particles is taken into account as a Zener-type drag force on the grain boundary motion (C. Zener (1948)). Particles are not introduced directly into the cellular automaton grid as usually (M. A. Miodownik et al., 1999), but through a separate, parallel data structure, thereby allowing to represent particles with a volume smaller as that of a cell.

Figure 12.5 illustrates the outcome of a simulation.
12.4 A Qualitative Example: Static Annealing of a Cold Rolled Steel

(a) Orientation distribution function of the recrystallized steel as computed. This ODF compares very well to typical experimental data.

(b) Two dimensional cross sections of the evolving microstructure. The left column shows the recrystallization phase, the right column the subsequent grain growth in the recrystallized microstructure.

(c) Evolution of the normalized grain size distribution. A scaling regime is soon reached as expected from theory.

Figure 12.5: Simulation results for a typical Al-killed steel that is first cold rolled and subsequently annealed. More details in the text.
12.5 Conclusion

A modified cellular automaton is described, allowing the simulation of grain growth phenomena under various conditions, including the combination of volume-stored energy and grain boundary curvature driving pressures. The nature of the model and the level of abstraction give the model a high versatility, which is simultaneously an advantage and a drawback. The drawbacks are its computational cost, which is at the very limits of today’s computing power, and the fact that the multitude of microstructural processes that can concurrently be taken into account can lead to a proliferation of variables that makes the overall model difficult to handle. The advantage is its flexibility, which allows material scientists to gain insight in the complex, three-dimensional nature of microstructure evolution, a manifold of different interacting processes. The potential of this kind of overall model strongly depends on a quantitative understanding of all the sub-processes taken into account, which is not the case for nucleation or for the details of the grain boundary mobility dependence on lattice misorientation in the current state of the art in the field.

Acknowledgements

The research presented in this chapter was largely performed at the Swiss Federal Institute of Technology in the Institute of Virtual Manufacturing of Prof. J. Reissner, and final corrections were added at Sandia National Laboratories, a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy under Contract DE-AC04-94AL85000. Special thanks to Dr. Fabiano Vanini and Dr. Ernst Kozeschnik for a pleasant cooperation, and to Dr. Elizabeth Holm for helping out with the final details.

References


References

13 Topological Relationships in 2D Trivalent Mosaics and Their Application to Normal Grain Growth

Robert Brandt, Kurt Lücke†, Guiseppe Abbruzzese, and Jiri Svoboda

13.1 Introduction

Grain growth (GG) shall here be understood as the coarsening of a polycrystalline microstructure caused by the tendency of reducing the total grain boundary (GB) energy of the specimen by means of GB motion and subsequent disappearance of the GBs. Here the principal progress in theoretical understanding of GG kinetics shall briefly be reviewed. As to be seen in Table 13.1 these investigations are based on the original work by Smith (1952) who clarified the role of the specific GB energy $\gamma$ and specified the GB curvature $K$ as driving force for the GB motion:

$$v = -m\gamma K = -MK.$$  (13.1)

Here is $v$ the velocity, $m$ the mobility and $M = m\gamma$ the diffusivity of the GB. According to Lücke et al. (1999) the progress mainly rests on improvements in the characterization of the individual grains by increasing the number of the considered parameters.

In the first, rather rough attempts by Burke and Turnbull (1952) and Beck (1954) the individual grains were simply disregarded and only the average grain size was applied (Table 13.1). The mean grain radius $R$ was considered as a kind of effective radius of curvature for the GB motion so that, according to Equation (13.1), these authors arrived at the so called $t^{0.5}$-law:

$$v \propto \frac{dR}{dt} \propto \frac{M}{R} \Rightarrow R^2(t) - R_0^2 \propto Mt.$$  (13.2)

In the following pioneering work by Hillert (1965) – who only considered the case of uniform GBs – each grain was characterized by its radius $R$ and, consequently, the microstructure by a 1-parametric grain size distribution function (SDF) $\phi(R, t)$. Here is $\phi(R, t)dR$ the fraction of the number of grains being at the time $t$ in the interval between $R$ and $R + dR$. With $N(t)$ being the total number of grains the author then assumed a continuity equation

$$\frac{\partial[N(t)\phi(R, t)]}{\partial t} = -\frac{\partial[N(t)\phi(R, t)G(R, t)]}{\partial R}$$  (13.3)

describing the flow of the number of grains from the interval at $R$ to the next higher (growing grains) or next lower (shrinking grains) interval (Figure 13.1). For the average growth rate
Table 13.1: Main steps in developing the statistical theory of GG in homogeneous Materials.

<table>
<thead>
<tr>
<th>Number</th>
<th>Characterisation of Polycrystalline Samples</th>
<th>Authors</th>
<th>Applied Means</th>
<th>Main Achievements</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Mean radius: R</td>
<td>Burke 1952 Beck 1954</td>
<td>Rough approach: $\frac{dR}{dt} = \frac{1}{R}$</td>
<td>Normal GG $R'^2(1) - R_0^2 = t$ (t$^n$ - law)</td>
</tr>
<tr>
<td>1</td>
<td>Radius: R, Size distribution - function SDF: $\phi(R)$</td>
<td>Hillert 1965</td>
<td>Uniform GBs Continuity equation: $\frac{d}{dt} \left[ N\phi(R) \right] = - \frac{d}{dR} \left[ N\phi(R)G(R) \right]$ For $G(R)$ heuristic assumptions</td>
<td>Normal GG $t^n$ - law Hillert self similar distribution $\phi_{Hillert}(R/R)$</td>
</tr>
<tr>
<td>2</td>
<td>Radius: R, Coordination: n Size distribution - function SCDF: $\phi(R,n)$</td>
<td>Abbruzzese - Lücke - Heckelmann - Brandt 1984 - 1999</td>
<td>Uniform GBs 120° angles at vertices Continuity equation: $\frac{d}{dt} \left[ N\phi(R) \right] = - \frac{d}{dR} \left[ N\phi(R)G(R,n) \right]$ For $G(R,n)$ first principles: Topolog. relationship $n_i = n(R_i)$</td>
<td>Individual Grains $\frac{dA}{dt} = \frac{m\gamma}{2} (n - 6)$</td>
</tr>
<tr>
<td>3</td>
<td>Radius: R, Coordination: n Orientation: $\gamma^h, \gamma^k, ...$ Size - coordination distribution - function SCODF: $\phi^n(R,n)$</td>
<td>Abbruzzese - Lücke - Eichelknut Brandt 1986 - 1999 (present state)</td>
<td>Non uniform GBs Equilibrium angles $\beta$ at vertices Set of continuity equations: $\frac{d}{dt} \left[ N\phi^n(R) \right] = - \frac{d}{dR} \left[ N\phi^n(R)G(R) \right]$ For $G(R)$ first principles: Topological functions $k_{i,j}, P_{i,j}$ (Two-Grain Model)</td>
<td>Abnormal GG Time evolution of SDF and ODF: $\phi(R,t), R^2(t)$ $t^n$ - law No $t^n$ - law No self similar distribution</td>
</tr>
</tbody>
</table>
G(R, t) := dR/dt of the grains of size R he made the heuristic assumption that grains being larger than a critical size $R_{cr}$ grow and smaller than $R_{cr}$ shrink. For the special case of statistical self-similarity (SSS) as the final result the $t^{0.5}$-law (Equation (13.2)) was recovered and a self-similar distribution, the so called Hillert distribution $\phi_H(R)$, was derived.

Experimentally, however, the $t^{0.5}$-law was observed only very seldomly and the Hillert distribution never. For this reason several authors tried to modify the model. The most important modification is based on some work by von Neumann (1952) and by Mullins (1956) who showed for uniform GBs that the rate of area change of an individual grain is not determined by its radius $R$ but by the number $n$ of sides of the grain (Table 13.1), the so called coordination. Under the assumption of uniform GBs and equilibrium at the vertices (triple points, TPs) with dihedral angles of $2\pi/3$ these authors derived for the rate of change of the grain area $A$ by integrating the GB velocity $v$ along the GBs around this grain

$$\frac{dA}{dt} = \int v \, ds = \frac{M\pi}{3} (n - 6). \tag{13.4}$$

Beside the above mentioned contributions a lot of interesting works of different authors exist which, however, will not be discussed here in detail.

On this basis Abbruzzese et al. (1992) and Lücke et al. (1992) developed for the case of uniform GBs a statistical theory of GG where the grains are characterized by the parameters $R$ and $n$ and the whole microstructure by the 2-parametric size – coordination distribution function $\phi(R, n)$. This theory was derived from first principles not needing any further heuristic assumptions. Instead, beside the above rate equation (Equation (13.4)), topological relationships guaranteeing space filling (first principle) during GG were introduced.

However, because of the crystalline nature of the grains, the above case of uniform GBs does not exist in reality. Instead one has to assume that the dihedral angles at the TPs are no longer $2\pi/3$ since the properties of the GBs (e.g. $\gamma$) depend on their crystallographic structure and thus also on the crystallographic orientations $\gamma^H$ and $\gamma^K$ of the neighboring grains (Table 13.1). This means each TP and its dihedral angles must be at least characterized by three parameters $H$, $K$ and $L$ and a GB and its properties by two parameters $H$ and $K$ (e.g. GB energy by $\gamma^{HK}$), each grain by the three parameters $R$, $n$ and $g$ and the whole specimen by a 3-parametric size – coordination – orientation distribution function $\phi^H(R, n)$. Finally, instead of one continuity equation (Equation (13.3)), here a whole set of continuity equations
is needed, one for each orientation class. All this will lead to dramatic changes of the results with respect to cases of only uniform GBs.

![Figure 13.2: GB networks: Voronoi mosaic M-I with straight GBs, M-II after simulation of GG, M-IV after simulation of GG, and M-V, a 2D section of a 3D Al3%Mg specimen after GG.](image)

Up to now the topological aspects were mostly taken care of by some heuristic assumptions or rather rough models, but, here for the 2D case these neighborhood relations are considered correctly. The present lecture gives a brief overview of this new field here being called “Topological Metallography” for only 2D and only trivalent mosaics where at each TP three grains, i.e. also three GBs, meet. For the description of each grain 4 parameters are considered (grain size $R$, grain shape $f$, number of sides $n$, crystallographic orientation $g$) and five examples of 2D microstructures are used (see Figure 13.2):

- M-I: A Voronoi mosaic of $N = 10000$ grains.
- M-II: A GG mosaic of $N = 2488$ grains starting from $N_0 = 9000$ grains.
- M-III: A hard-core mosaic of $N = 4900$ grains. (No figures are given.)
- M-IV: A hard-core GG mosaic of $N = 3039$ grains with $N_0 = 4900$ grains.
- M-V: A 2D section through a 3D Al3%Mg sample after GG ($N = 1988$).

### 13.2 Individual Grains and their Distributions (One-Grain Model)

#### 13.2.1 Definition of Parameters

The main parameter of a grain $\mu$ is its perimeter $L_\mu$ equivalent grain size $R_\mu$. The shape of the grain shall be described by a second parameter $f_\mu$ giving the ratio of the grain size $R_\mu$ and the area $A_\mu$ equivalent radius $R^A_\mu$. Related to the mean radius $\overline{R}$ a reduced radius $r$ is defined. Thus one obtains

$$R_\mu := R^L_\mu = \frac{L_\mu}{2\pi}; \quad f_\mu := \frac{R^L_\mu}{R^A_\mu}; \quad R^A_\mu = \sqrt{\frac{A_\mu}{\pi}} = \frac{R_\mu}{f_\mu}; \quad r_\mu = \frac{R_\mu}{\overline{R}} = f_\mu r^A_\mu. \quad (13.5)$$

The value $f_\mu = 1$ characterizes a circular grain shape. A third parameter is the coordination number $n_\mu$. $R$ will be represented in a discrete way by the size class indices $i, j, k, \ldots$
The whole microstructure is described in a 2D phase space \((r, n)\) by the two-parametric distribution function \(\varphi_{in}\) which denotes the fraction of the number of grains in size class \(i\) having the number of sides \(n\). One thus has

\[
\sum_{i,n} \varphi_{in} = \sum_{i} \varphi_i = \sum_{n} \varphi_n = \sum_{H} \varphi^H = 1. \tag{13.6}
\]

\(\varphi_i\) represents the grain size distribution function (SDF) and \(\varphi_n\) the coordination distribution function (CDF) with the average values \(\overline{r} = 1\) and \(\overline{n} = 6\).

### 13.2.2 The Grain Sizes and Shapes and their Distributions

It is little be said about the shape factors \(f\). In Figure 13.3 values \(f\) of the grains are plotted as dots. The open circles are representing the average values \(f_i\). For the simulated mosaics M-II and M-IV \(f_i\) is approximately constant given by \(f_i = 1.08\). The lower limits of \(f\) are plotted as solid lines showing almost free zones between the lower limit and \(f = 1\) for circular grain shapes.

Practically most often applied for describing microstructures is the SDF. In Figure 13.4 the measured values are plotted as full circular symbols and they are compared with best fits of 4 different types of distributions often applied in literature.

\[
\begin{align*}
\text{M-I} &\quad N = 10^4 \\
\text{M-II} &\quad N = 2488 \\
\text{M-IV} &\quad N = 3039 \\
\text{M-V} &\quad N = 1988
\end{align*}
\]
13.2.3 The Coordination and its Distributions

The frequency of grains possessing a size \( r \) and the coordination number \( n \) is given by a two-parametric distribution \( \varphi_{in} \) plotted into a 2D phase space \((r, n)\). In Figure 13.5 for mosaics M-I and M-V the individual grains are plotted as dots (for mosaics M-II and M-IV the dots are omitted). For the sake of visibility, they are not positioned at \( n = 3, 4, 5 \ldots \), but they are randomly distributed between these integer values. The partial average number of sides \( n_i \) (closed circles) and \( r_n \) (open squares) were plotted, too, and one recognizes that these average values are situated near a solid and a dashed line

\[
\bar{n}_i = 3 + 3r_i \quad \text{and} \quad n = 6 + \frac{3}{\rho^2} (\bar{r}_n - 1)
\]

which are also called “Abbruzzese–Lücke equations” (ALEs). \( \rho \) is the correlation coefficient. These lines describe rather well the partial averages. Although even the partial averages of the simulated GG mosaics show particular deviations of it the ALEs are representing a kind of idealized behavior being well suited for giving a general description.

13.3 Topological Relationships of Trivalent Mosaics

13.3.1 Grain Boundaries (GBs) and Triple Points (TPs)

While in Sec. 13.2, as a first approximation for describing a polycrystalline microstructure, a one-grain model, i.e. only individual grains, were considered, now, in order to include the topology, a two-grain model is applied. This means that we take care of the arrangement of grains with respect to their neighbors and thus of the properties of GBs and TPs. For this purpose we consider now the average length \( \ell_{ij} \) of the GBs between grains of the size \( R_i \) and others of size \( R_j \) and we consider the average values \( \theta_{ij} = -\theta_{ij} \) of the total change of inclination along these GBs. Furthermore, we consider the average angle \( \beta_{ijk} \) at the TP \( i, j, k \) which becomes by averaging over all \( k \) for all GBs \( i, j \) the angle \( \beta_{ij} \) (see Figure 13.6).
13.3.2 The Geometry of the GB (Function $p_{ij}$)

Here use is made of the fact that the individual GBs of a grain $\mu$ always form a closed loop leading to a total change of inclination of $2\pi$ by going once around clockwise (by convention) along its perimeter $L$. The perimeter $L$ is composed of the lengths $\ell_\lambda$ of the $\lambda = 1, \ldots, n_\mu$ GBs and the total change of inclination of $2\pi$ by the changes of inclination $\vartheta_\lambda$ (by convention negative sign) of the $n_\mu$ GBs and $\beta_\lambda$ of the $n_\mu$ TPs at the end of the GBs, i.e.

$$\sum_{\lambda=1}^{n_\mu} \ell_\lambda = L ; \quad \sum_{\lambda=1}^{n_\mu} \beta_\lambda - \vartheta_\lambda = 2\pi.$$  \hspace{1cm} (13.8)

These equations lead to the fraction $w_{ij}$ of the perimeters $N_i L_i$ of all grains $i$ which are covered by grains $j$ and to the fraction $w^*_ij$ of the total change of inclination $2\pi N_i$ of all grains $i$ which is made up by grains $j$

$$w_{ij} := \frac{E_{ij}}{N_i n_{ij}} \Rightarrow \sum_j w_{ij} = 1 ; \quad w^*_ij := \frac{E_{ij}}{N_i n^*_{ij}} \Rightarrow \sum_j w^*_ij = 1.$$  \hspace{1cm} (13.9)

$E_{ij}$ is the total number of contacts between the number $N_i$ grains of size $R_i$ and others of size $R_j$ and $n_{ij} = L_i/\ell_{ij}$ and $n^*_{ij} = 2\pi/(\beta_{ij} + \vartheta_{ij})$ are coordination numbers with respect to the length $\ell_{ij}$ and the total change of inclination $\beta_{ij} + \vartheta_{ij}$, respectively. The ratio between these quantities gives the topological function $p_{ij}$

$$p_{ij} := \frac{n_{ij}}{n^*_{ij}} = \frac{w^*_ij}{w_{ij}}.$$  \hspace{1cm} (13.10)

The function $p_{ij}$ therefore indicates whether a number of $n_{ij}$ averaged GBs $ij$ together with a number of $n_{ij}$ averaged TPs $ij$ can form a closed loop of total inclination $2\pi$.

13.3.3 Size Correlations of Nearest Neighbor Grains (Function $k_{ij}$)

For a completely random covering of the perimeter of the grains $i$, the quantity $w_{ij}$ would be equal to the fraction $w_j$ of the perimeter of the grains $j$ and the perimeter of all grains:

$$w_j = \frac{\varphi_j L_j}{\sum_j \varphi_j L_j} = \varphi_j r_j.$$  \hspace{1cm} (13.11)
13.3 Topological Relationships of Trivalent Mosaics

For the non random cases correlation functions (coefficients) must be introduced which characterize the deviations from the random behavior:

\[ k_{ij} := \frac{w_{ij}}{w_j}; \quad k^*_{ij} := \frac{w^*_{ij}}{w^*_j}; \quad (13.12) \]

Since the part \( w_{ij} \) of the perimeters \( N_i L_i \) of the \( N_i \) grains of class \( i \) which is covered by grains \( j \) is equal to that part \( w_{ji} \) of the perimeters \( N_j L_j \) of the \( N_j \) grains \( j \) which is covered by grains \( i \) one obtains the symmetric relationship \( k_{ij} = k_{ji} \). A generally valid relationship \( k^*_{ij} = k^*_{ji} \), however, does not exist.

13.3.4 Space Filling (Function \( q_{ij} \))

Space filling along the GBs requires that the pair of boundaries confining two adjacent grains \( \mu \) and \( \nu \) are identical, i.e. one obtains the principal GB symmetry conditions:

\[ \ell_{ij} = \ell_{ji}; \quad \vartheta_{ij} = -\vartheta_{ji}. \quad (13.13) \]

This leads to

\[ \frac{n_{ij}}{n_{ji}} = \frac{L_i/\ell_{ij}}{L_j/\ell_{ji}} = \frac{r_i}{r_j} \quad \text{and} \quad \vartheta_{ij} = \frac{\beta_{ij} - 2\pi p_{ij}/n_{ij}}{\beta_{ji} - 2\pi p_{ji}/n_{ji}} = -1 \quad (13.14) \]

and with a further topological function \( q_{ij} \)

\[ q_{ij} := \frac{\beta_{ij} + \beta_{ji}}{2\pi/3} \quad \text{where} \quad q = \frac{\sum_{ij} \varphi_i w_{ij} n_{ij} q_{ij}}{\sum_{ij} \varphi_i w_{ij} n_{ij}} = 1 \quad (13.15) \]

to

\[ n_{ij} = 3 \left( p_{ij} + \frac{r_i}{r_j} p_{ji} \right). \quad (13.16) \]

By averaging the relative number of contacts \( E_{ij}/N_i = w_{ij} n_{ij} \) over all \( j \) one gets

\[ n_i = \sum_j w_{ij} n_{ij} = 3 \left( \sum_j \varphi_j r_j \frac{k_{ij} p_{ij}}{q_{ij}} + r_i \sum_j \varphi_j \frac{k^*_{ij} p_{ji}}{q_{ji}} \right). \quad (13.17) \]

This relationship which only depends on the function \( k_{ij} p_{ij}/q_{ij} \) of the above topological functions is very important since \( n_i \) which gives the average number of sides of grains of the size \( r_i \), for example, is absolutely needed for the derivation of the statistical theory of GG kinetics.
13.4 Cases of Randomness

13.4.1 “Abbruzzese–Lücke Equations” (ALE, Full Randomness)

One recognizes that the 2D networks are characterized by the above topological functions $k_{ij}$, $k_{ij}^*$, $p_{ij}$ and $q_{ij}$ (Secs. 13.3.2, 13.3.3 and 13.3.4) which, in general, have to be experimentally determined. Of special interest, however, is the case that all four topological functions are constant and that they are replaced by their “expectation values”, i.e. by their average values $\bar{k} = \bar{p} = \bar{q} = 1$. These as well as the functions which are derived from it are always indicated by a tilde, e.g. $\tilde{k}_{ij}$ or $\tilde{n}_{ij}$. By substitution of $\tilde{k}_{ij} = \tilde{p}_{ij} = \tilde{q}_{ij} = 1^1$ into Equation (13.17) and by averaging over all $j$ one finally gets

$$\tilde{n}_i = 3 \left( \sum_j \varphi_j r_j + r_i \sum_j \varphi_j \right) = 3 (1 + r_i).$$  \hspace{1cm} (13.18)

This is denoted as “Special Linear Relationship” (SLR). For a linearly correlated set of pairs of data $(r, n)$ a second, corresponding equation for $\tilde{r}_n$ can be found by linear regression analysis. This leads together with the SLR (Equation (13.18)) again to the “Abbruzzese–Lücke Equations” (ALE)

$$\tilde{n}_i = 3 (1 + r_i), \quad \tilde{r}_n = 1 + \frac{\rho^2}{3} (n - 6)$$  \hspace{1cm} (13.19)

which were so far only experimentally found. The ALEs can thus be considered as the idealized case which proves that most mosaics are nearly random.

13.4.2 “Weaire–Aboav Equation” (WAE, Partial Randomness)

Beside the fully random case also cases of partial randomness are existing where only some of the functions are replaced by their “expectation values”, e.g. as observed for simulated GG mosaics (Lücke et al., 1999). Here that case of partial randomness shall be discussed which leads to the correct derivation of the well known “Weaire–Aboav equation” (Aboav, 1970) and it is characterized by a hat over the corresponding topological function, i.e. $\hat{k}_{ij} = 1$ and $\hat{q}_{ij} = 1$, and the functions which are derived from it, e.g. $\hat{w}_{ij}$.

Starting from Equation (13.12) one gets by use of $E_{ij} = E_{ji}$ and the corresponding terms of Equations (13.9), (13.13), (13.15)

$$\hat{w}_{ij}^* = w_j^* = \frac{\sum_i E_{ij} (\hat{\beta}_{ij} - \hat{\vartheta}_{ij})}{\sum_i 2\pi N_i} = \frac{\varphi_j}{3} \left( \sum_i w_{ji} n_{ji} q_{ji} - 3 \right)$$  \hspace{1cm} (13.20)

since

$$\sum_m E_{ij} (\hat{\beta}_{ij} - \hat{\vartheta}_{ij}) = \sum_m E_{ji} [\hat{\beta}_{ij} + \hat{\beta}_{ji} - (\hat{\beta}_{ji} - \hat{\vartheta}_{ji})] = 2\pi/3N_j \sum_i w_{ji} n_{ji} q_{ji} - 2\pi N_j.$$

$^1 \hat{k}_{ij}^* = 1$ follows automatically by use of Equation (13.10) leading to full randomness
13.4 Cases of Randomness

All angles are \( \beta_{\mu\nu\lambda} = \pi/3 \) by applying the mechanical equilibrium condition so that one obtains from Equation (13.20) with \( q_{ij} = \hat{q}_{ij} = 1 \)

\[
\hat{w}^*_j = \varphi \frac{n_j - 3}{3} = w_j \frac{n_j - 3}{3r_j} \Rightarrow \frac{\hat{w}^*_j}{w_j} = \frac{n_j - 3}{3r_j}.
\] (13.21)

![Figure 13.7: The product \( k_{ij}p_{ij} \) as function of \( r_j \). The measured values of mosaic M-II (symbols) are fitted by \( \hat{w}^*_j/w_j \).](image)

Since the fraction \( w^*_{ij} \) of the total inclination \( 2\pi N\varphi_i \) of the grains \( i \) is made up by the number \( E_{ij} = E_{ji} = N\varphi_j w^*_{ji} n^*_{ji} \) of GBs \( ij \) and their TPs, i.e.

\[
2\pi\varphi_i w^*_{ij} = \varphi_j w^*_{ji} n^*_{ji} (\beta_{ij} + \vartheta_{ji}),
\] (13.22)

one gets for the coordination number \( n^*_{ij} \) of this case of partial randomness by use of Equation (13.15) and \( \hat{q}_{ij} = 1 \), the definition of \( n^*_{ij} \) (Equation (13.10)). Substituting \( w^*_{ij} \) by \( \hat{w}^*_j \) finally yields

\[
\hat{n}^*_{ij} = 3 \left( 1 + \frac{n_i - 3}{n_j - 3} \right).
\] (13.23)

This case of partial randomness has been observed for mosaics generated by simulation of GG. The product \( k_{ij}p_{ij} = \hat{w}^*_j/w_j \) is found to be independent of \( r_i \) (Figure 13.7) and it is fitting well to \( \hat{w}^*_j/w_j \) which is given by Equation (13.21) and which is indicated by a full line in Figure 13.7. This yields for the example of mosaic M-II by use of Equation (13.10) a correlation coefficient \( k^*_{ij} = \hat{k}^*_{ij} = 1 \) which means that one observes here completely random covering of the circumference of a grain of total inclination \( 2\pi \) by other grains \( j \). A similar case of partial randomness is found for mosaic M-IV, too. The Voronoi mosaics M-I and M-III show a non random behavior with respect to this definition.

For the derivation of the WAE, instead of the grain size \( r_i \), the coordination number \( n \) is of main interest. By application of \( \hat{k}^*_{mn} = \hat{q}_{mn} = 1 \) one gets

\[
\hat{w}^*_{mn} = w^*_n = \varphi_n \frac{n - 3}{3} \quad \text{and} \quad \hat{n}^*_{mn} = 3 \left( 1 + \frac{m - 3}{n - 3} \right).
\] (13.24)

The result of this special case of randomness can be applied for the correct derivation of the “Weaire–Aboav equation” which gives the average number \( \bar{m}_n \) of sides for the grains which are next neighbor to the \( n \)-sided grains by

\[
\bar{m}_n = 5 + \frac{6 + \sigma_n^2}{n} \quad \text{with} \quad \sigma_n = \pi n^{-}\kappa_n.
\] (13.25)


13.5 Curvature Driven GG

13.5.1 Direct Simulations

For demonstrating the effect of the initial microstructure on GG three different series of different initial microstructures are presented for direct simulations with uniform GBs.

S-I Here 10 Voronoi mosaics of 900 grains each were chosen as the initial microstructures of series S-I. From each set of 900 randomly distributed sites situated in quadratic areas, cells were formed by assigning each part of the area to that of the sites to which it is closest.

S-II Here a hard core mosaic of 4900 grains was chosen as the initial microstructure of series S-II. From a set of 4900 randomly distributed sites situated in a quadratic area, where the distance between each two sites was chosen to be at least $1/(2\sqrt{4900})$ of the edge of the quadratic interval of periodicity, cells were formed again by assigning each part of the area to that of the sites to which it is closest.

S-III Here a mosaic of 1600 grains of uniform size was chosen as the initial microstructure of series S-III. From a set of 1600 almost regularly distributed sites situated in a quadratic area, cells of almost quadratic shape and of nearly equal size were formed again by assigning each part of the area to that of the sites to which it is closest.

These mosaics are trivalent mosaics where at each TP three grains, i.e. also three GBs, meet with an average number of sides $\pi = 6$ (Lücke et al., 1999; Graustein, 1931). The cells of these series were exposed to numerical simulation of curvature driven GG including switching and vanishing of grains (see Figure 13.8) by applying uniform GB energy $\gamma_0$ and GB mobility $m_0$.

The closed symbols in Figure 13.9 indicate the growth of the grains. One recognizes at small times the grain size $R/R_0 = 1$. Due to initial adjustments of the shape of the grains leading to a reduction of the total perimeter length and of the average shape factor $f$ the perimeter equivalent grain size even decreases. This range is the longest for series S-III followed by S-II and S-I. The reason for this different initial behavior is that in series S-I one has already relatively small grains in the beginning whereas in S-II and even S-III no such small grains are available in the initial microstructures. However, at long times when a sufficient number of grains has been vanished the well known $t^{0.5}$ law occurs. The transient regime between the initial stage of minor shrinkage of the mean perimeter and the range of the $t^{0.5}$ law becomes relatively small for series S-III and it is the longest for series S-I (as to be seen in Figure 13.9).

These difference show up still more pronounced in the variation coefficients $\kappa$. In Figure 13.9 the relative width $\kappa$ of the SDFs is indicated by open symbols. The initial values of $\kappa$ increase from $\kappa_V = 0.24$ of a Voronoi mosaic for series S-I to $\kappa = 0.43$, from $\kappa_{hc} = 0.15$ of a hard core mosaic for S-II to $\kappa = 0.44$, and from $\kappa_{III} = 0.01$ of the initial mosaic for S-III to $\kappa = 0.41$ finally leading to a plateau value of $\kappa \approx 0.42$ (Figures 13.9) independent of the initial values. However, the width of the SDFs is significantly larger at small times for series S-I than for S-III. This shows also up by the Hillert value $\kappa_{Hillert} = 0.331$ (Hillert, 1965), which is gained by series S-I at $t \approx 20$, by S-II at $t \approx 100$ and by S-III at $t \approx 200$. The
13.5 Curvature Driven GG

Figure 13.8: Evolution of the microstructures of series S-I, S-II and S-III during GG.

Figure 13.9: The relative grain sizes $R/R_0$ and the variation coefficients $\kappa$ during GG of series S-I, S-II and S-III. Results by the statistical method are printed as lines.

The transient regime between the initial stage and the range of plateau $\kappa \approx 0.42$ becomes here the smallest for series S-III and it is the longest for series S-I.

13.5.2 Simulations by the Statistical Theory

The statistical theory which accurately describes curvature driven GG shall now be applied. One gets for the growth rate $G_i$ by use of the averaged von Neumann-Mullins law $dA_i/dt =$
mγπ(n_i - 6)/3 and the assumption of an uniform shape factor \( f_i = f \approx 1.08 \)

\[
G_i := G(r_i) = \frac{dR_i}{dt} = \frac{f_i^2}{2\pi R_i} \frac{dA_i}{dt} = \frac{mγf_i^2}{6R_i}(n_i - 6).
\] (13.26)

The shape factor \( f_i = f_i(t) \) and the topological relationship \( n_i \) which are introduced into Equation (13.26) for the calculation of the growth rate \( G_i \) shall be approximately given according to the results of the direct simulations.

Let us first consider the results for the shape factor \( f \) of series S-I. Figure 13.10 shows the evolution of the shape factors \( f \) of the individual grains as dots and the averaged shape factors \( f_i \) for the grains of size \( r_i \) as open circles.

It is clearly seen that except for small times \( t \) the average shape factors \( f_i \) are very well approximated by a value \( f(t) \) which is independent of the grain size \( r \). The assumption of a size independent average shape factor \( f_i \) which is approximated by a function \( f = f(t) \) is therefore fully justified. Here the value \( f(t) \) shall be approximated by the measured total averages \( \bar{f}(t) \) in the simulated microstructures of series S-I indicated as closed circles. At small times a value of \( \bar{f}(0) \approx 1.17 \) is obtained. Then the values are decreasing and for large times the average shape factor becomes a constant value at about \( \bar{f}(\infty) \approx 1.08 \). The measured points are approximated very well by the fit

\[
\bar{f}(t) = 1.08 + [1.17 - 1.08] \exp \left( -\frac{t}{10} \right)
\] (13.27)

which is plotted in Figure 13.10 as a solid line. The shape factor \( f \) of Equation (13.26) is then to be substituted by this expression \( \bar{f}(t) \) of Equation (13.27).

Secondly, in the left hand diagram of Figure 13.11 the closed circles are representing the measured partial averages \( n_i \) for the initial Voronoi mosaic of series S-I and the open squares the corresponding partial averages \( r_n \). In the right hand diagram of Figure 13.11 the symbols are representing the partial averages \( n_i \) and \( r_n \) for a GG microstructure.

For the Voronoi mosaic at small GG times the Abbruzzese-Lücke Equations (ALEs) (Lücke et al., 1999)

\[
\bar{n}_i = 6 + a(r_i - 1) \quad \text{and} \quad n = 6 + \frac{a}{\rho}(\bar{r}_n - 1) \quad \text{with} \quad a = 3
\] (13.28)
indicated as lines in both diagrams of Figure 13.11 do agree well with the measured partial averages $n_i$ and $r_n$, but after some GG the ALEs fit them only roughly and particularly at small grain sizes a hump appears. As was shown in Brandt and Lücke (1999) the 4-sided grains are showing great stability by which their transition to 3-sided grains is impeded and thus the relative number of 4-sided grains is strongly increased. In this way for $r \to 0$ a value $n$ close to 4 instead of 3 is obtained and the relationship is better described by a parabola than by a straight line. Thus for the fit of the topological relationship $n_i$ to the data of series S-I a parabola $n(r, t)$ is used

$$n(r, t) = n_0(t) + n_1(t)r + n_2(t)r^2$$  \hspace{1cm} (13.29)

with

$$n_0(t) = 4.0 - 0.8 \exp \left( -\frac{t}{41} \right); \quad n_1(t) = 1.1 + 1.6 \exp \left( -\frac{t}{39} \right);$$

$$n_2(t) = \frac{6 - n_0(t) - n_1(t)}{r^2}. \hspace{1cm} (13.30)$$

The topological relationship $n_i$ of Equation (13.26) is then to be substituted by this expression $n(r, t)$ of Equation (13.29).

For the remaining series S-II and S-III very similar relationships are introduced. For all simulations by the statistical method and the direct method the same GB diffusivity $M = m_0 \gamma_0$ is applied.

In terms of this approach (see Sec. 13.2.1) the perimeter equivalent growth rate $G_i$ (Equation (13.26)) is proportional to the area equivalent growth rate $G_i^A = dR_i^A/dt$ since the shape of the grains is assumed to be uniform. It is approximated just by a function $f = f(t)$. A rigorous distinction of perimeter and area equivalent grains sizes by a more general function $f = f(r, t)$ in the case of series S-I seems to be inevitable only in the very early stage of GG. There the shapes of the grains given by the average shape factor $f_i$ are obviously depending on the grain size $r_i$ (Figure 13.10). Just after a short time of GG the shape of the grains is fully adjusted to an about uniform shape factor $f = f(t)$ leading to $r = fr^A$ for all grain size classes. This is due to fast adjustments of the perimeters $L$ of the grains without significant changes of their grain areas $A$ in the beginning. These fast adjustments of the perimeter equivalent SDF show only a minor effect on the area equivalent SDF. Here, for the simulations the
shape of the initial perimeter equivalent SDF is set to be identical to that of the initial area equivalent SDF of the initial microstructure of the considered series. They are given for the three series by:

S-I A normal distribution of initial width $\kappa = 0.27$.

S-II A log-normal distribution of initial width $\kappa = 0.16$.

S-III A normal distribution of initial width $\kappa = 0.05$. For numerical reasons a larger value than the measured value of $\kappa_{III} = 0.01$ is chosen.

The resulting SDFs of the simulations by the statistical method are shown in Figure 13.12 by lines. They agree well with the SDFs of the directly simulated microstructures which are indicated by symbols and by perpendicular columns in the case of series S-III. Furthermore, the grain sizes $R$ and the corresponding values of $\kappa$ are correctly simulated by this statistical approach (see Figure 13.9). This means that both approaches of describing curvature driven GG are equivalent. However, deviations are obtained for the intermediate stage of GG in series S-III. Here the direct simulation method predicts a multimodal SDF whereas the statistical method predicts a unimodal SDF. This is due to the very special initial microstructure of series S-III: the grains are of about equal size $R \approx \overline{R}$ and of nearly quadratic shape. A closer look, however, shows that the grains are not exactly quadratic but the microstructures are of trivalent nature with a mean coordination $\overline{n} = 6$. Depending upon $n > 6$, $n = 6$ and $n < 6$ the grains will be able to carry out GG leading to the multimodal SDF.

### 13.6 Summarizing Remarks

The main parameters ($R$, $f$ and $n$) have been introduced for a topological description of homogeneous 2D trivalent microstructures (see Sec. 13.2.1). Although these microstructures seem to look rather random it could be shown that there exist well-defined relationships between these parameters of nearest neighbor grains (one-grain model). This model has been extended by a two-grain and a three-grain model where two or three next neighbor grains and thus GBs and TPs are included into the considerations. The points are thoroughly discussed at the example of four different types of mosaics (see Sec. 13.1).
13.6 Summarizing Remarks

By this way a statistical theory of microstructures has been developed which is well suited for deriving various types of topological relationships. The topological functions $k_{ij}, k^*_{ij}, p_{ij}$, and $q_{ij}$ have been introduced where $k_{ij}$ represents the fraction of the perimeters of grains of the size class $i$ which is covered by other grains of size class $j$. $k^*_{ij}$ represents the analogous relationship for the total changes of inclination. (For $p_{ij}$ and $q_{ij}$ see Sec. 13.3.)

This theory is leading to a new and elegant derivation of the fundamental topological law $n = 6$ which had been derived so far mainly by geometrical means (Graustein, 1931). By substitutions of Equations (13.16), (13.12), (13.15), and $k_{ij} = k_{ji}$ and finally by averaging the terms of Equation (13.17) over all $i$ one gets

$$\bar{n} = \sum_{i,j} \varphi_i w_{ij} n_{ij} = \sum_{i,j} \varphi_i w_{ij} n_{ij} q_{ij} = 6. \quad (13.31)$$

By means of these functions $k_{ij}, k^*_{ij}, p_{ij}$ and $q_{ij}$ a well founded definition of randomness has been established. Since the total average values characterizes the randomness for the special case of fully randomness one gets $k_{ij} = k^*_{ij} = p_{ij} = q_{ij} = 1$. These equations yield the “Abbruzzese–Lücke Equations” (ALEs)

$$\bar{n}_i = \bar{n}(r_i) = 3(1 + r_i), \quad \bar{r}_n = \bar{r}(n) = 1 + \frac{\rho^2}{3}(n - 6). \quad (13.32)$$

which are shown to be a reasonable approximation for most 2D mosaics (e.g. Figure 13.5). This means that these equations are a kind of ideal topological relationships. However, the exact fully random case is not found for any of the mosaics. The mosaic M-V comes close to it since the 2D section intersects the 3D body randomly so that $k_{ij}, k^*_{ij}, p_{ij}$ and $q_{ij}$ are given by their average values.

However, there do not only exist correlations between grains of different size classes but also between grains of different coordination classes. The topological function $k^*_{mn} = k^*_{nm}$ represents thus also a quantity giving randomness. This provides a new interpretation and derivation of the “Weaire–Aboav equation”:

$$\bar{m}_n = 5 + \frac{6 + \sigma^2}{n}. \quad (13.33)$$

We calculated first an expectation value for the mean number of sides $\bar{m}_{n\mu}$ of a grain $\mu$ which is nearest neighbor to a grain $n$ and, finally, in a second step the average number of sides of the neighbors of grain $n$. This is the result of only applying the principle of random covering of the total inclination of $2\pi$ along the circumference of a grain $n$ by other grains $m$. The partial mean values $\bar{m}_n$ of simulated GG mosaics, e.g. M-II, and of 2D sections of microstructures after GG, e.g. mosaic M-V, are described in good approximation by this equation. The corresponding values $m_n$ of a Voronoi mosaic M-I, however, are badly fitted by this equation since strong correlations occur leading to $k^*_{mn} \neq k^*_{nm}$. A derivation of the “Weaire–Aboav equation” has also been given in Blanc (1979), but, there it is based on the heuristic assumption that there exist no correlation between the coordination number of a grain and any one of its second nearest neighbors which, however, does not agree to the measurements (Heckelmann, 1992).

The statistical theory of microstructures allows an understanding of the mechanisms of actual formation of the microstructure which is by C. S. Smith (Smith, 1952) “a result of the
References

interplay between the requirements of the physical forces operating between the individual parts and the mathematical requirements of space filling”. This theory has been applied also to numerical simulations of curvature driven GG. In contrast to other statistical models on GG, here no heuristic assumptions nor assumptions on the shape of grains have been made. This theory has been derived by means of first principles including all geometrical and topological requirements.

Acknowledgements

The authors gratefully acknowledge the financial support from the MaTech program of the “Bundesministerium für Bildung, Wissenschaft, Forschung und Technologie” (BMBF) under number 03K8008 2.

References

14 Motion of Multiple Interfaces: Grain Growth and Coarsening

Britta Nestler

A new diffuse interface model (phase-field model) for non-isothermal solidification in alloy systems with multiple components and multiple phases is introduced. The model is capable to describe phase transitions, microstructure formations and interfacial motion in polycrystalline grain structures. The model is derived from an entropy functional in a thermodynamically consistent way and is related to classical sharp interface models in the limit of zero interface thickness. Anisotropic and faceted growth characteristics are included in both the kinetic coefficients and the surface energies. The set of governing equations is discretized and implemented. 2D and 3D simulations are performed showing the temporal evolution of phase fractions combined with the heat and mass diffusion during solidification and melting processes. To improve the computational efficiency for large domains, two strategies are followed. Firstly, a 3D parallel simulator based on a heterogeneous network of workstations is set up. This network solves the governing equations using a finite difference discretization on a uniform grid with explicit time update. Secondly, an adaptive finite element differential equation analysis library is used. Simulation results are presented showing the motion of multiple interfaces (phase and grain boundaries). In particular, phenomena such as anisotropic and crystalline curvature flow, grain growth, coarsening, wetting, symmetry properties and stability conditions of adjacent multiple junctions are described. In order to describe phase transformations and solidification processes in multi-component alloys, the specific phase diagrams are incorporated in the diffuse interface model via the free energies. Within this context, ternary eutectic structures are simulated. The effect of anisotropy on eutectic structures, such as the growth of tilted eutectic lamellae, and the formation of eutectic grains at different length scales: the grain structure on a larger scale and the eutectic substructure on a smaller scale are investigated. The stability of ternary eutectic growth fronts and the characteristic spacings depending on process conditions are studied by iterated simulations.

14.1 Introduction

The phase-field method has become a powerful methodology to describe phase transition phenomena in alloys with reliable qualitative results in comparison with experimental data. The method has successfully been used to describe solidification processes (e.g. Caginalp et al. (1989); Penrose and Fife (1990); Karma and Rappel (1998); Karma (2001)) as well as microstructure evolution in solids (e.g. Fried and Gurtin (1994)) and liquid-liquid interfaces (e.g. Lowengrub and Truskinovsky (1998)). Diffuse interface models have been formulated
for pure substances (e.g. Caginalp et al. (1989); Penrose and Fife (1990)), binary alloys (e.g. Caginalp and Xie (1998); Tiaden et al. (1998)) for eutectic, peritectic and monotectic systems (e.g. Wheeler et al. (1996); Nestler and Wheeler (2000); Nestler et al. (2003, 2000); Plapp and Karma (2002)). Furthermore the evolution of grain boundaries can be modelled by phase-field approaches or order parameter models (e.g. Chen and Yang (1994); Fan and Chen (1997); Garcke and Nestler (2000); Nestler and Wheeler (1998)). The concept of describing interface evolution by a phase-field model was originally introduced by Langer (1986) using ideas of Halperin et al. (1974) who studied critical exponents for transport coefficients. Diffuse interface models such as the phase-field model have the characteristic property that the interfaces in the system have a non-zero thickness. First approaches of diffuse interfaces were already used by van der Waals (1893); Cahn and Hilliard (1958) and Landau and Ginzburg (1965). In papers by Caginalp and Fife (1988); Caginalp et al. (1989), the authors carried out a formal asymptotic expansion for phase-field models of solidification. By taking the limit of vanishing interfacial thickness, the diffuse interface profile relates to classical sharp interface models, also known as Stefan problems where the interface is described by a sharp step function profile. In these sharp interface models, boundary conditions for the physical quantities have to be prescribed at the interface between two bulk phases. In recent years, phase-field models have been derived from thermodynamic principles being now referred to as thermodynamically consistent (see Penrose and Fife (1990); Wang et al. (1993)).

Traditionally, the evolution of interfaces has been modelled by moving free boundary problems separating the regions of pure phases by a sharp interface contour line. In these sharp interface models, partial differential equations for e.g. mass and heat diffusion are solved within the bulk phases. These equations are coupled by boundary conditions on the interface, such as the Stefan condition demanding energy balance and the Gibbs-Thomson equation. Across the sharp interface, quantities such as the heat flux, the concentration field and the energy may show a discontinuous jump profile. The physics at the boundaries is prescribed by explicit tracking of the interfaces. In diffuse interface models the different bulk phases in the system are distinguished by a vector of phase-field variables $\phi$. The number of components of this vector depends on the number of phases or grains of different crystal structure or orientation in the system. The phase boundaries are modelled by a diffuse interface, where the phase fields and other physical quantities change smoothly on a thin transition layer. For a single solid-liquid phase transition, the phase-field variable $\phi$ may take values of one in the solid and zero in the liquid. Across the interface, the phase field varies smoothly from one to zero. From a computational point of view, it is required that the spatial resolution of the numerical method must be finer than the thickness of the diffuse phase boundary layer. The interfacial thickness itself must be less than the characteristic scale of the growing microstructure. In this case, a nonuniform grid with adaptive refinement can dramatically reduce computational resources compared with a uniform grid with the same spatial resolution.

In the following sections a non-isothermal phase-field model is derived in a thermodynamically consistent way that allows for an arbitrary number of phases/grains and components and that is defined solely via the bulk free energies of the individual phases, the surface energy densities of the interfaces and diffusion and mobility coefficients. Since the diffuse interface model yields classical moving boundary problems in the sharp interface limit, the full set of phase-field evolution equations is defined by quantities which can experimentally be measured. The physical effects occurring during the solidification such as heat and mass transfer,
14.2 The Diffuse Interface Model

The diffuse interface model for a general class of alloy systems is formulated consisting of $K$ components ($Al$, $Fe$, $Si$, $Cu$, ...) and $N$ different phases and grains of different crystal structures and orientations in a domain $\Omega \subset \mathbb{R}^3$. The concentrations and phase fractions or grains are represented by a vector $c(\vec{x}, t) = (c_1(\vec{x}, t), \ldots, c_K(\vec{x}, t))$ and by an order parameter $\phi(\vec{x}, t) = (\phi_1(\vec{x}, t), \ldots, \phi_N(\vec{x}, t))$, respectively. The phase-field model is based on an entropy functional of the form

$$S(e, c, \phi) = \int_\Omega \left( s(e, c, \phi) - (\varepsilon a(\phi, \nabla \phi) + \frac{1}{\varepsilon} w(\phi)) \right) dx. \quad (14.1)$$

The bulk entropy density $s$ is assumed to depend on the internal energy density $e$, on the concentrations of the components $c_i$, $i = 1, \ldots, K$, and on the phase-field variable $\phi_\alpha$, $\alpha = 1, \ldots, N$. The additional contributions $a(\phi, \nabla \phi)$ and $w(\phi)$ of the entropy functional reflect the thermodynamics of the interfaces (e.g. Visintin (1996)). In diffuse interface models, $\varepsilon$ is a small length scale parameter related to the thickness of the diffuse interface. The surface entropy contributions are expressed by a Ginzburg-Landau type functional of the form

$$-\int_\Omega \left( \varepsilon a(\phi, \nabla \phi) + \frac{1}{\varepsilon} w(\phi) \right) dx,$$

where $a(\phi, \nabla \phi)$ is a gradient energy density and $w(\phi)$ is an energy of multi-well or multi-obstacle type. Examples of possible choices will be given in the next section. The variable $\phi_\alpha$ denotes the local fraction of phase $\alpha$ and it is required that the concentrations of the components and the phase-field variables fulfill the constraints

$$\sum_{i=1}^K c_i = 1, \quad \sum_{\alpha=1}^N \phi_\alpha = 1. \quad (14.2)$$

It will be convenient to use the free energy as a thermodynamical potential. Therefore the Gibbs relation is postulated

$$df = -sdT + \sum_i \mu_i dc_i + \sum_\alpha r_\alpha d\phi_\alpha,$$
where $T$ is the temperature, $\mu_i = f_{c_i}$ are the chemical potentials and $r_\alpha = f_{\phi_\alpha}$ are potentials due to the appearance of different phases. The notations $f_{c_i}$ and $f_{\phi_\alpha}$ denote the partial derivatives of the free energy $f$ with respect to $c_i$ and $\phi_\alpha$. Using the thermodynamical relation $e = f + sT$ it follows that

$$de = Tds + \sum_i \mu_idc_i + \sum_\alpha r_\alpha d\phi_\alpha,$$

$$ds = \frac{1}{T}de - \sum_i \frac{\mu_i}{T}dc_i - \sum_\alpha \frac{r_\alpha}{T} d\phi_\alpha.$$

Interpreting $s$ as function of $(e, c, \phi)$ one gets

$$s, e = \frac{1}{T}, \quad s, c_i = \frac{-\mu_i}{T}, \quad s, \phi_\alpha = \frac{-r_\alpha}{T}.$$

Knowing the free energy densities of the pure phases $f_\alpha$, the total free energy $f$ is obtained as a suitable interpolation of $f_\alpha$.

The governing equations for the conserved order parameters $e$ and $c_i, i = 1, \ldots, K$ are derived from the entropy functional in Equation (14.1)

$$\frac{\partial e}{\partial t} = -\nabla \cdot J_0 \quad \text{(energy balance)}, \quad (14.3a)$$

$$\frac{\partial c_i}{\partial t} = -\nabla \cdot J_i \quad \text{(mass balances, } i = 1, \ldots, K) \quad (14.3b)$$

and they are coupled to the evolution equations of the non-conserved order parameters $\phi_\alpha, \alpha = 1, \ldots, N$ via

$$\omega \frac{\partial \phi_\alpha}{\partial t} = \frac{\delta S}{\delta \phi_\alpha} - \lambda, \quad \alpha = 1, \ldots, N, \quad (14.3c)$$

in such a way that the second law of thermodynamics is fulfilled in an appropriate local version. For the non-conserved phase-field variables $\phi_\alpha$, it is assumed that the system locally tends to maximize entropy, while, at the same time, concentration and energy is locally conserved. The variational derivative $\frac{\delta S}{\delta \phi_\alpha}$ of Equation (14.3c) gives

$$\omega \frac{\partial \phi_\alpha}{\partial t} = \varepsilon \left( \nabla \cdot a_{c\phi_\alpha} - a_{\phi_\alpha} \right) - \frac{1}{\varepsilon} w_{\phi_\alpha} - \frac{f_{\phi_\alpha}}{T} - \lambda,$$

where $a_{\phi_\alpha}, w_{\phi_\alpha}, f_{\phi_\alpha}$ and $a_{c\phi_\alpha}$ denote the derivatives of the energy contributions with respect to $\phi_\alpha$ and $\nabla \phi_\alpha$, respectively. $\omega$ is a constant kinetic coefficient and $\lambda$ is an appropriate Lagrange multiplier such that the constraint $\sum_\alpha^N \phi_\alpha = 1$ in Equation (14.2) is satisfied, i.e.

$$\lambda = \frac{1}{N} \sum_\alpha \left[ \varepsilon \left( \nabla \cdot a_{c\phi_\alpha} - a_{\phi_\alpha} \right) - \frac{1}{\varepsilon} w_{\phi_\alpha} - \frac{f_{\phi_\alpha}}{T} \right]. \quad (14.4)$$

The quantities on the right hand sides of the Equations (14.3a) and (14.3b) are the energy flux $J_0$ and fluxes $J_1, \ldots, J_K$ of the components $c_1, \ldots, c_K$. Appropriate expressions are
derived from generalized thermodynamic potentials \( \frac{\delta S}{c_i} = \frac{1}{T} \) and \( \frac{\delta S}{c_i T} = \left( \frac{\delta s}{\delta \mu} \right) \) driving the evolution. Appealing to non-equilibrium thermodynamics it is postulated that the fluxes are linear functions of the thermodynamic driving forces \( \nabla \frac{\delta S}{c_i}, \nabla \frac{\delta S}{c_i T}, \ldots, \nabla \frac{\delta S}{c_i T^N} \) to obtain

\[
J_0 = L_{00}(T, c, \phi) \nabla \frac{\delta S}{c_i} + \sum_{j=1}^{K} L_{0j}(T, c, \phi) \nabla \frac{\delta S}{c_j}
\]

\[
= L_{00}(T, c, \phi) \nabla \frac{1}{T} + \sum_{j=1}^{K} L_{0j}(T, c, \phi) \nabla \frac{-\mu_j}{T}, \tag{14.5a}
\]

\[
J_i = L_{i0}(T, c, \phi) \nabla \frac{\delta S}{c_i} + \sum_{j=1}^{K} L_{ij}(T, c, \phi) \nabla \frac{\delta S}{c_j}
\]

\[
= L_{i0}(T, c, \phi) \nabla \frac{1}{T} + \sum_{j=1}^{K} L_{ij}(T, c, \phi) \nabla \frac{-\mu_j}{T} \tag{14.5b}
\]

with mobility coefficients \( (L_{ij})_{i,j=0,\ldots,K} \) relating to heat and mass diffusion coefficients. To fulfill the constraint \( \sum_{i=1}^{K} \mu_i = 1 \) in Equation (14.2) during the evolution, is required that \( \sum_{i=1}^{K} L_{ij} = 0, j = 0, \ldots, K \) implying \( \sum_{i=1}^{K} J_i = 0 \) and hence \( \partial_t \left( \sum_{i=1}^{K} c_i \right) = \nabla \cdot \left( \sum_{i=1}^{K} J_i \right) = 0 \). Further, it is assumed that the \( L_{ij} \) are positive semi-definite and symmetric according to the Onsager relations. In Garcke, Nestler and Stinner (2002) it is shown that this condition leads to an entropy inequality ensuring positive local entropy production. Cross effects between mass and energy diffusion can be neglected by setting \( L_{i0} = 0 \) and \( L_{0j} = 0 \) for all \( i,j \in \{1, \ldots, K\} \).

### 14.3 Free Energies

The diffuse interface model is capable of describing alloy systems with a very general class of multiphase multicomponent phase diagrams by specifying the free energies \( f_\alpha(T, c) \). The model allows for systems with concave entropies \( s_\alpha(c, c) \) in the pure phases corresponding to free energies \( f_\alpha(T, c) \) being convex in \( c \) and concave in \( T \). If \( f(T, c, \phi) \) is not convex in the variable \( c \), the free energy needs to contain gradients of the concentrations as e.g. in a Cahn-Hilliard type model. Choosing the liquid phase to be the last component \( \phi_N \) of the phase-field vector \( \phi \), the bulk free energies are defined for the individual phases by

\[
f_\alpha(T, c) = \sum_{i=1}^{K} \left( c_i L_i^\alpha \frac{T - T_i^\alpha}{T} + k_B T c_i \ln(c_i) \right) - c_N T (\ln(T) - 1)
\]

with \( L_i^N = 0 \) and \( L_i^\alpha, i = 1, \ldots, K, \alpha = 1, \ldots, N - 1 \), being the latent heat per unit volume of the phase transition from phase \( \alpha \) to the liquid phase and pure component \( i \). Furthermore, \( T_i^\alpha, i = 1, \ldots, K, \alpha = 1, \ldots, N - 1 \) is the melting temperature of the \( i \)-th component in the
phase $\alpha$, $c_\alpha$ is the specific heat which is for simplicity assumed to be independent of $c$ and $\phi$; $k_B$ is the Boltzmann constant. Then the total free energy density follows:

$$f(T, c, \phi) := \sum_{\alpha=1}^{M} \sum_{i=1}^{N} \left( c_i L_i^\alpha \frac{T - T^\alpha_c}{T_c} h(\phi_\alpha) \right)$$

$$+ \sum_{i=1}^{N} \left( k_B T c_i \ln(c_i) \right) - c_v T (\ln(T) - 1).$$  \hspace{1cm} (14.6)

With a suitable choice of the function $h(\phi)$ satisfying $h(0) = 0$ and $h(1) = 1$, e.g. $h(\phi_\alpha) = \phi_\alpha^2$ or $h(\phi_\alpha) = \phi_\alpha^2 (3 - 2\phi_\alpha)$, the free energy density $f$ is an interpolation of the individual free energy densities $f_\alpha$. The entropy density $s(e, c, \phi)$, the inner energy density $e$ as well as the chemical potentials $\mu_i$ follow from the free energy density in Equation (14.6):

$$s = -f_T = - \sum_{\alpha=1}^{M} \sum_{i=1}^{N} \left( c_i L_i^\alpha \frac{T^\alpha}{T^\alpha_c} h(\phi_\alpha) \right) - \sum_{i=1}^{N} \left( k_B c_i \ln(c_i) \right) + c_v \ln(T),$$

$$e = f + Ts = - \sum_{\alpha=1}^{M} \sum_{i=1}^{N} \left( c_i L_i^\alpha h(\phi_\alpha) \right) + c_v T,$$

$$\mu_i(T, c, \phi) = f_{c_i}(T, c, \phi) = \sum_{\alpha=1}^{M} \left( L_i^\alpha \frac{T - T^\alpha_c}{T_c} h(\phi_\alpha) \right) + k_B T \ln(c_i) + 1.$$  \hspace{1cm} (14.7a)

Next, the terms modelling interfacial contributions to the free energy will be defined. In Garcke, Nestler and Stoth (1998, 1999b); Steinbach et al. (1996) it has been shown that gradient energies of the form

$$a(\phi, \nabla \phi) = \sum_{\alpha=1}^{M} \sum_{\beta, \alpha<\beta} A_{\alpha\beta}(\phi_\alpha \nabla \phi_\beta - \phi_\beta \nabla \phi_\alpha)$$

have very good properties with respect to calibrating parameters in the phase-field model to the surface terms in the sharp interface model. A choice that leads to isotropic surface terms is

$$a(\phi, \nabla \phi) = \sum_{\alpha<\beta} \frac{\gamma_{\alpha\beta}}{m_{\alpha\beta}} |\phi_\alpha \nabla \phi_\beta - \phi_\beta \nabla \phi_\alpha|^2$$  \hspace{1cm} (14.8)

with constants $\gamma_{\alpha\beta}$ and $m_{\alpha\beta}$ that can be related to surface entropy densities $\gamma_{\alpha\beta}$ and kinetic coefficients $m_{\alpha\beta}$. In the case of surface energy anisotropy, the $\gamma_{\alpha\beta}$ parameters depend on the orientation of the interface. The antisymmetric formulation $\phi_\alpha \nabla \phi_\beta - \phi_\beta \nabla \phi_\alpha$ allows to treat the physics of each interface individually.

Expressions for the potential $w(\phi)$ can be formulated as direct extensions of the standard double well or double obstacle potential for solid-liquid phase-field models to a multi well $w_{st}(\phi)$ or multi obstacle potential $w_{ob}(\phi)$ for the multi phase model:

$$w_{st}(\phi) = 9 \sum_{\alpha<\beta} \tilde{m}_{\alpha\beta} \gamma_{\alpha\beta} \phi_\alpha^2 \phi_\beta^2$$ and $$w_{ob}(\phi) = \frac{16}{\pi} \sum_{\alpha<\beta} \tilde{m}_{\alpha\beta} \gamma_{\alpha\beta} \phi_\alpha \phi_\beta.$$  \hspace{1cm} (14.9)
By numerical experiments comparing different choices for the potential $w(\phi)$, it is found that the multi obstacle potential with additional higher order variants

$$\tilde{w}_{ob}(\phi) = w_{ob}(\phi) + \sum_{\alpha < \beta < \delta} \gamma_{\alpha \beta \delta} \phi_{\alpha} \phi_{\beta} \phi_{\delta}$$

yields the best calibration properties. It is generally difficult in multi phase-field models to compute the surface free energy densities (or surface entropy densities). In studies by Garcke, Nestler and Stoth (1999a) free energies for phase-field methods with good calibration properties have been developed. This means that for experimentally given surface free energies one can calibrate the parameters in the free energies of the phase-field model in such a way that the sharp interface limit is defined via the given surface tensions. Another advantage of using an obstacle type potential for numerical simulations is that the potentials $w_{ob}(\phi)$ or $\tilde{w}_{ob}(\phi)$ are defined to be infinity whenever $\phi$ is not on the Gibbs-Simplex. Therefore, the equations for the phase-fields $\phi_{\alpha}$, $\alpha = 1, \ldots, N$ only need to be solved in the region of the diffuse interface layer. It is referred to Garcke, Nestler and Stoth (1999a,b) for a further discussion of the properties of the surface terms.

To complete the explicit definitions of quantities in the diffuse interface model, an example for the mobility matrix $(L_{ij})_{i,j=0,\ldots,N}$ is given defining heat and mass diffusion as well as cross effects between them. By assuming that the mobilities of the pure components are linear in $c_i$ and by referring to Garcke, Nestler and Stinner (2002) the following expressions are obtained

$$L_{ij} = D(T, c, \phi)c_i(\delta_{ij} - c_j), \quad (14.10)$$

$$L_{0j} = -\sum_{i,\alpha} D(T, c, \phi)c_j(\delta_{ij} - c_i)h(\phi_{\alpha})L_{i\alpha}, \quad (14.11)$$

$$L_{00} = k(T, c, \phi)T^2 + \sum_{i,j,\alpha,\beta} D(T, c, \phi)h(\phi_{\alpha})L_{i\alpha}c_j(\delta_{ij} - c_i)h(\phi_{\beta})L_{j\beta}, \quad (14.12)$$

where $D(T, c, \phi)$ and $k(T, c, \phi)$ contain mass and heat diffusion coefficients and $\delta_{ij}$ denotes the Kronecker delta.

### 14.4 Numerical Simulations

Inserting the relations for $\lambda$, $J_0$, $J_i$ in Equations (14.4,14.5a,14.5b), the free energies $f$, $a$, $w$ in Equations (14.6,14.8,14.9), the mobility matrix $(L_{ij})$ in Equations (14.10,14.11,14.12) and the resulting derivatives for $e, \mu_i$ in Equations (14.7a, 14.7b) into the governing equations in Equations (14.3a,14.3b,14.3c), a full set of evolution equations for the inner energy, the concentrations and the phase-fields is set up and discretized. Numerical simulations of the motion of multiple interfaces in grain structures, of solidification processes and phase transitions in multi component alloy systems are performed. For simulating microstructure evolutions in large domains at acceptable computation times, two different strategies are used: A 3D parallel simulator based on a heterogeneous network and on a server-client mechanism is set up for parallel solving of the partial differential equations with a finite difference discretization and
an explicit time scheme. Alternatively, an adaptive finite element method with a semi-implicit time discretization is applied to solve the equations. Diffuse interface simulations require that the spatial resolution of the numerical method is greater than the thickness of the diffusive phase boundary layer. The interfacial thickness itself must be less than the characteristic scale of the growing microstructure. In this case, a nonuniform grid with adaptive refinement can dramatically reduce the use of computational resources compared to a uniform grid with the same spatial resolution. Both approaches may converge to a hybrid solution at some future stage.

14.4.1 Grain Growth and Coarsening

The first simulation in Figure 14.1 is motivated by an experiment of heteroepitaxial thin film growth of a cubic tricrystalline aluminum grain structure on a silicon substrate, described in Dahmen and Thangaraj (1993). According to the experimental setup, the focus lies on a tricrystalline grain formation with three allowed orientational variants. The variants are rotated about a common \(<001>\) axis by 30° with respect to each other. Two superimposed cubic grains with a relative rotation of 30° between their crystallographic orientation exhibit eight symmetry/mirror axes. The symmetry/mirror axes are preferred directions for grain boundaries to be formed, because in these directions the boundaries are in states of minimal energy. This can be explained using the criterion of symmetry dictated-extrema, Cahn and Kalonji (1994). Therefore, an eightfold convex and crystalline (facetted) surface energy anisotropy is formulated having a typical cusp-like structure at the eight preferred directions (see Garcke, Nestler and Stoth (1999b) for the anisotropy formulation).

As in experimental observations, the grain boundaries in polycrystalline structures are forced by a minimization of energy criterion to evolve at certain symmetry conditions. Force balance laws lead to grain configurations with adjacent triple junctions which must always belong to different symmetry classes and hence adjust at different angle conditions. The evolution of the triple junctions is driven by the classical Young’s law. This law relates the surface energies meeting at the junctions to the angles formed at the junction. In the case of anisotropic surface energies, in particular crystalline (facetted) energies, additional shear forces in tangential direction to the grain boundaries occur in the Young’s equation. A more detailed description of the crystalline modification of Young’s law was first derived in Garcke and Nestler (2000).

In numerical computations, triple junctions and the influence of surface energy anisotropy on the angle condition at the triple junction are investigated. In accordance with the classical equilibrium force balance law by Young with shear forces due to surface energy anisotropy, angle conditions differing from the equal 120° case are obtained as a result of additional torque terms in this anisotropic force balance equation. Wetting behavior that occurs in physical systems, if the surface energies violate the stability condition \(\gamma_{ij} \leq \gamma_{ik} + \gamma_{kj}\) can also be recovered.

The simulation in Figure 14.2 shows the process of two-dimensional grain growth and coarsening in a system of multiple grains. In this context, the components \(\phi_\alpha\) of the multiple order parameter model describe different orientational variants in a crystalline material. In a system where all the surface energies are equal as in Figure 14.2, all triple junctions have a 120° angle condition. Characteristic features of the von Neumann law can be observed.
14.4 Numerical Simulations

postulating that grains with less than six neighbors shrink, grains with more than six neighbors grow and six-sided grains keep their area in time. A region which shrinks is marked by a cross whereas a region which keeps the area is marked by a circle.

Next, the phase-field model is applied to three-dimensional grain growth phenomena in an isotropic and anisotropic grain system in Figures 14.3 and 14.4, respectively. The simulations are performed with five order parameters. For the computation in Figure 14.4, a crystalline hexagonal anisotropy with grains of different crystallographic orientations is chosen. The evolution of the grain boundaries support the fact that facets form in certain preferred directions and that anisotropy changes the equilibrium angle condition of 120° at triple junctions.

14.4.2 Multicomponent Multiphase Solidification

The objective of this section is to show the utility of the diffuse interface model to simulate a wide variety of realistic growth structures and morphologies in multicomponent multiphase systems as reported in Kurz and Fisher (1992); Sahm and Kurz (1975); Akamatsu and Faivre (2000). To describe the phase transitions and solidification processes in metallic alloys, the
Figure 14.3: Phase-field simulation of three-dimensional grain growth with equal and isotropic surface energies. The grains are colored by different gray scales.

Figure 14.4: Phase-field simulation of three-dimensional grain growth with crystalline (facetted) hexagonal surface energies and grains of different orientation.

specific physical parameters and the phase diagram are taken into account via the free energies in the model. A collection of binary peritectic, eutectic and monotectic microstructures is presented in the snapshots, Figure 14.5.

Figure 14.5: Snapshots of a simulated binary alloy solidification of a peritectic (left), a eutectic (middle) and monotectic system (right). The growing bulk phases are illustrated in white and black. The concentration profile in the undercooled melt with depleted and enriched regions is shown in continuously varying gray scale.

Examples for three-dimensional binary eutectic alloy structures are displayed in Figures 14.6 and 14.7. Once the parent liquid phase (metallic melt) is cooled below the eutectic temperature, the liquid phase $L$ transforms into two new solid phases $\alpha$ and $\beta$ via a eutectic reaction: $L \rightarrow \alpha + \beta$. The phase-field simulations were carried out taking a symmetric phase diagram and periodic boundary conditions.
14.4 Numerical Simulations

Figure 14.6: Three-dimensional simulation of a binary eutectic alloy establishing a regular rod-like structure of the alternating two solid phases growing into the undercooled melt (transparent gray).

Figure 14.7: Irregular three-dimensional growth mode of a binary eutectic due to the antisymmetric initial phase configuration differing from the symmetric eutectic phase diagram chosen for the simulation.

An application of the diffuse interface modelling technique to the multi-scale phenomena of competitive evolution of eutectic grains is illustrated in Figure 14.8. The two involved length scales are the grain structure on a large scale and the finer lamellar eutectic substructure. The simulation was set up with different anisotropies for the two eutectic grains.

Figure 14.8: Two eutectic grains (white/black and light/dark-gray) of a binary alloy with different crystal orientations growing into the melt (continuous gray scale). The images visualize the phase evolution and concentration profiles of the alloy composition in the liquid ahead of the growing solid phases at different time steps.
Ternary alloy solidification involves phase changes of four different phases and diffusion of three components. A ternary eutectic phase diagram is constructed by common tangents and the according free energies are used for the following simulations. Aiming to investigate the stability of ternary lamellar fronts, different permutations of the three solid phases $\alpha, \beta, \gamma$ are considered in Figure 14.9 and Figure 14.10.

The effect of crystalline surface energy anisotropy on ternary eutectic microstructures is studied in the phase-field simulations of Figures 14.11 and 14.12. The anisotropy leads to tilted growth typically observed in experiments. A phase shift within the lamellae occurs for crystal orientations aligned in the growth direction.

Figure 14.9: Ternary eutectic lamellae with a configuration $\alpha$ (gray), $\beta$ (black), $\gamma$ (white), showing the concentration of one of the components in the melt (continuous gray scale) at three different times.

Figure 14.10: Ternary eutectic lamellae with a configuration $\alpha$ (gray), $\beta$ (black), $\alpha$ (gray), $\gamma$ (white). The three pictures illuminate the concentration fields (continuous gray scale) ahead of the solid front for all three components of the $A - B - C$ alloy at a medium time step.

Figure 14.11: Tilted solidification front in a ternary eutectic alloy with three solidifying solids (gray, black, white) as a result of the influence of surface energy anisotropy.
Finally, ternary alloy solidification is computed in three space dimensions. Regular lamellae of the $\alpha, \beta, \gamma$ solids in diagonal space direction are formed for the initial condition in Figure 14.13.

![Figure 14.12: Phase shifts in ternary eutectic solid phases with a fourfold crystalline anisotropy.](image)

![Figure 14.13: Three-dimensional computation of a ternary eutectic alloy with three different solid phases (gray, black, white) evolving with steady state lamellar growth behavior in diagonal space direction.](image)

All three solid phases grow from the liquid phase. While simultaneously growing, the solid phases mutually enhance each other’s growth conditions as they reject opposite components of the alloy into the liquid. This leads to the establishment of a steady regular hexagonal shape in Fig 14.14 with phase boundaries having isotropic surface energies.

![Figure 14.14: The dynamic solidification process of ternary eutectic hexagonal rods corresponding to three different times.](image)
In a subsequent computation the phase evolution was simulated taking crystal anisotropy and differently rotated solid particles into account, Figure 14.15.

Figure 14.15: The influence of crystalline surface energies on the evolution of phase boundaries in a phase-field simulation of ternary eutectic phase transitions forming strong facets in the preferred growth directions.

14.5 Outlook

A thermodynamically consistent diffuse interface model (phase-field model) has been developed describing non-isothermal solidification in alloy systems with multiple components and phases. The new model was discretized and two- and three-dimensional simulations of phase transitions, of solidification processes, of complex microstructure formations and of interfacial motion in polycrystalline grain structures are performed. Based on these results, it is intended to apply the phase-field simulations to model complex multiscale growth phenomena involving different length (and time) scales. Examples are the growth of eutectic colonies resulting from small amounts of ternary impurities and dendritic networks with interdendritic eutectic substructures. In order to quantitatively describe phase transformations and solidification processes in multi-component alloys, it is planned to incorporate specific phase diagrams into the phase-field model via the free energies by linking the governing equations to a thermodynamics data base. Another challenge for future work is to include convection into the phase-field model, to investigate the interaction of different physical fields and their effect on the microstructure evolution.

Acknowledgements

The author gratefully thanks Harald Garcke and his group for collaboration over many years. The financial support through the German Research Foundation within the two priority research programs: Multiscale problems (1095) and phase transitions in multicomponent melts (1120) is gratefully acknowledged.
References


References

15 Deformation and Recrystallization of Particle-containing Aluminum Alloys

Bala Radhakrishnan and Gorti B. Sarma

The phenomenon of particle-stimulated nucleation (PSN) during recrystallization of commercial aluminum alloys has been widely investigated both from a scientific viewpoint and from the industrial perspective. The technological impact of this phenomenon is the finding that depending upon the thermo-mechanical processing conditions, the occurrence of PSN may strengthen or weaken the deformation texture during recrystallization, which may or may not be desirable for a specific application of the final product. From a scientific standpoint, the mechanics of formation of localized deformation zones in the vicinity of the hard particles and the development of deformation substructures within the deformation zone that eventually turns into a recrystallized nucleus are still not completely understood.

This paper describes the application of a coupled finite element-Monte Carlo technique to study the phenomenon of particle-stimulated nucleation (PSN) during recrystallization of aluminum and the subsequent growth of the nuclei. The deformation and recrystallization textures arising from initial microstructures with and without hard particles are compared to elucidate the particle effect on texture. The simulations have been applied to single crystals, bicrystals and tri-crystals of aluminum of specific crystallographic orientations with and without the hard particles at specific locations such as grain interior, grain boundary and triple line. The simulation results are compared with existing experimental data on the deformation and recrystallization of particle-containing aluminum alloys.

15.1 Background

When aluminum alloys containing micron-sized hard second-phase particles are subjected to thermo-mechanical processing, distinct deformation zones form around the hard particles where the stored energy of deformation and the crystallographic orientations are significantly different from those in the matrix far away from the particles. Recrystallized nuclei form in the deformation zones when the deformed material is subjected to annealing. This phenomenon, known as particle stimulated nucleation (PSN), results in nucleus orientations that can be significantly different from the orientations produced in the absence of hard particles.

---

1 Research sponsored by the Division of Materials Sciences and Engineering, U.S. Department of Energy, under contract DE-AC05-00OR22725 with UT-Battelle, LLC. The submitted manuscript has been authored by a contractor of the U.S. Government under contract No. DE-AC05-00OR22725. Accordingly, the U.S. Government retains a non-exclusive, royalty-free license to publish or reproduce the published form of this contribution, or allow others to do so, for U.S. Government purposes.
(Humphreys 1977, Humphreys 1979, Humphreys and Hatherly 1995). In addition to significant changes in texture, PSN also gives rise to marked enhancement in the recrystallization kinetics compared to that of particle-free alloys (Humphreys and Hatherly 1995).

A brief review of the progress in our understanding of the formation of the deformation zone and the evolution of recrystallization textures in particle containing aluminum alloys is provided in this section.

### 15.1.1 Formation of Deformation Zones

In the past there have been several investigations of the formation of deformation zones around coarse particles in single crystals and polycrystals of aluminum, and the development of recrystallization textures in these alloys. Humphreys (1979) investigated the formation of deformation zones during tensile deformation and showed that the occurrence of lattice rotations in the deformation zones was significant only when the particle sizes were in the micron range. In single crystals oriented for single slip, the lattice rotations were about an axis normal to the slip direction and the slip plane normal, in accordance with the suggestion of Ashby (1970). An orientation gradient was observed in the vicinity of the particle, with the rotation being the maximum near the particle-matrix interface and decreasing with increasing distance from the particle. Empirical equations were fitted to the rotation data for different particle sizes and different strains. In the case of polycrystals deformed by rolling (Humphreys 1977, Humphreys 1979, Hansen and Bay 1981) the deformation zone was elongated in the rolling direction with the formation of small, equiaxed and highly misoriented subgrains in the vicinity of the particle-matrix interface. The misorientation between these subgrains and the matrix was of the order of $40^\circ$.

Early attempts at modeling the formation of deformation zones near particles were limited to single crystals oriented for single slip. Humphreys and Kalu (1990) extended the idea of Ashby (1970) to the deformation of single crystals deforming in more than one slip system, and polycrystals containing particles to predict the size and the orientations in the deformation zone. In the case of polycrystals, it was assumed that each particle will produce several deformation zones based on the operation of different slip systems. The rotation axis for each deformation zone was related to the corresponding slip system according to Ashby (1970), or in cases where the different zones overlapped, the rotation axis was derived on the basis of overlap of the various rotation axes. One of the drawbacks of this model was that it was better applicable for single slip rather than poly-slip deformation, as well as for small and medium deformations rather than large deformations. When the deformation was heavy, the maximum rotations in the deformation zone were over-predicted by the model.

Humphreys and Ardakani (1994) developed an alternate approach for predicting the lattice rotations in the vicinity of the particles in deformed polycrystals based on the concept of “slip shadowing”. It was assumed that the matrix far away from the particles deformed on five slip systems. In the vicinity of the particles, slip on several systems was blocked because of the particle. The effective rotation in the matrix was assumed to depend on the extent to which slip was curbed on these slip systems. In regions very close to the particle, all slip systems were largely inactive and therefore the orientations within this zone were very close to the initial matrix orientations. The assumption of a deformation limited zone resulted in reasonable predictions of lattice rotation in the vicinity of the particle for large deformations. By using a
Taylor analysis modified by the blockage of slip systems, they were able to correctly predict the lattice rotations in polycrystals. The above model has recently been successfully applied to the deformation of Goss oriented single crystals of aluminum (Ferry and Humphreys 1996).

Although there have been some other attempts to model the orientations in the deformation zones (Sandström 1980, Ørsund and Nes 1988b), these models are not very satisfactory with regard to their ability to capture the orientation gradients and spatial distribution of orientations.

15.1.2 Formation and Growth of Particle Stimulated Nuclei

It was recognized in early experiments (Humphreys 1977) that the nuclei originated at pre-existing subgrains in the deformation zone, aided by the large orientation gradient and the small subgrain size. These experiments also showed that the nuclei need not originate right at the particle interface, but may form at a little distance from it, which has an important consequence on the recrystallized texture as described later. The time for the formation of nuclei is inversely proportional to the magnitude of the orientation gradient, which has a significant effect on the recrystallization kinetics of particle containing alloys.

The orientations of the particle stimulated nuclei were observed to be within the orientation spread existing in the deformation zone and were found to depend on the extent and type of deformation. In the case of single crystals subjected to small deformation in tension (Humphreys 1977), a sharp recrystallization texture corresponding to the rotation of the deformation texture by $30-40^\circ$ about a $\langle112\rangle$ axis was observed. In many instances (Jack et al. 1989, Humphreys and Ardakani 1994) the recrystallized texture was found to be quite close to the initial matrix orientation, indicating that the nuclei occurred at regions very close to the particle interface where the deformation was effectively blocked on all possible slip systems. A similar trend was also observed for lightly deformed polycrystalline Al-Si alloy (Kalu and Humphreys 1988) where the orientations of many particle stimulated nuclei were close to the initial matrix orientations, although there were a few nuclei that were much farther away from the matrix orientations. In the case of Al-Mn alloys the recrystallization texture appeared to depend upon the recrystallization temperature (Ørsund and Nes 1988a). At high annealing temperatures, a weak recrystallization texture was produced while at low temperatures a sharp recrystallization was produced. The authors proposed that the sharp recrystallization texture was due to nuclei originating at the outer edges of the deformation zone where the orientations were close to the matrix orientation. However, such a mechanism does not appear to be well documented through experimental evidence. In heavily deformed polycrystals, the nuclei orientations were random and recrystallization gave rise to weak deformation textures (Chan and Humphreys 1984, Habiby and Humphreys 1993).

While the above observations and predictions of nucleus orientations and recrystallization textures in particle-containing aluminum alloys were essentially due to Humphreys and co-workers, a different approach has been taken by Lücke and Engler (Lücke and Engler 1990) and later by Engler and co-workers (Engler et al. 1997, Engler 1997, Engler 2001, Engler et al. 2001a, 2001b). According to these authors, the orientation spectra in the particle deformation zone in single crystals as well as polycrystals are also described based on the Ashby (1970) model. However, in their nucleation and recrystallization model it is assumed that the nuclei that form from the spectrum of subgrain orientations in the deformation zone are
those that have a special misorientation relationship of $40^\circ\langle 111 \rangle$ with respect to the matrix. This so-called “micro-growth selection” of nucleus orientation is supposed to occur very early in the recrystallization process. By careful measurements of the subgrain orientations in the deformation zone and the recrystallization textures in particle-containing single crystals and polycrystals they showed that in all cases, in addition to certain random components there were significant volume fractions of a characteristic ND-rotated cube orientation and a P-orientation that were both indicative of nuclei that had the above special orientation relationships with the matrix and originated in the particle deformation zone. It must be mentioned that the Al-Si and Al-Mn single crystals investigated by these authors (Engler et al. 2001b) did show some minor recrystallization components that could not be traced back to the slip systems existing in the deformation zones, although it is claimed that these weak orientations were indeed due to microgrowth selection from a deformed orientation that existed at microstructural heterogeneities in the deformed state.

All of the above experimental data and analysis of textures in particle-containing aluminum single crystals and polycrystals were limited to cold deformation followed by recrystallization anneal. However, there are important differences between the deformation substructures produced by cold deformation and hot deformation. In addition to a critical Zener-Hollomon parameter that has to be exceeded for particle stimulated nucleation and growth to be significant (Humphreys and Hatherly 1995), the orientations of the subgrains that are produced in the particle deformation zone as well as in the matrix after hot deformation could be different from those of cold deformation. It has been shown that during hot deformation of aluminum, slip on $(110)\langle 111 \rangle$ non-octahedral slip systems can be activated (Maurice and Driver 1997). If the slip activity along the non-octahedral slip systems is significant, then direct application of the Ashby model (Ashby 1970), shows that additional rotations about non-$\langle 112 \rangle$ rotation axes may be possible. It is important to understand how these additional rotations will affect the spectrum of orientations that might exist in the deformation zones near hard particles. It will be interesting to see if the particle stimulated nucleation texture models for cold deformation based on the presence of nuclei with $40^\circ\langle 111 \rangle$ misorientation with the matrix are still applicable to nucleation after hot deformation. It should be noted that the models described above for cold deformation have been used directly to predict the deformation and recrystallization textures after hot deformation (Vatne et al. 1996, Engler and Vatne 1998).

Although the above slip-based theories can provide an estimate of the lattice rotations near the particle as a function of the overall strain, they do not provide a quantitative description of the spatial distribution of the lattice rotations, and therefore the size and shape of the localized deformation zones as a function of the particle shape and size. Such a predictive model is required in order to understand quantitatively the formation of the recrystallization nuclei in the deformation zones during subsequent annealing. Also, the gradients in the stored energy of deformation over a length-scale of several micrometers must be quantified. The length-scale of the problem is too small for macroscopic plasticity theory, and the problem is too complicated to be solved using dislocation dynamics simulations.

One of the approaches for modeling plastic deformation at the microstructural length scale is the use of crystal plasticity theory to develop a constitutive response for a volume element undergoing a given plastic deformation. The details of the crystal plasticity theory are given in chapter 5, and the interested reader is also directed to overview articles on the subject
15.1 Background

The constitutive model can be used in conjunction with a finite element technique to simulate microstructural deformation. The finite element discretization is applied directly to a polycrystalline aggregate, with many elements used to discretize each grain, so that topological effects are directly incorporated into the model. The microstructure is typically deformed by application of boundary conditions corresponding to a homogeneous deformation, and the constitutive response in each element is obtained based on crystal plasticity. The deformation of each element is computed from the discretized deformation or velocity field, based on its position within the domain, and based on its material state (characterized by its crystallographic orientation and its slip system strength). The simulations permit the development of non-uniform deformations within each grain, depending on its size, orientation, and interactions with neighboring grains (Sarma et al. 1998). Several investigators have used this approach to study the deformation of single crystals and polycrystals of FCC metals, using both two-dimensional and three-dimensional discretizations. Early applications of this approach were limited to planar meshes (Becker 1991, Bronkhorst et al. 1992, Anand and Kalidindi 1994, Becker and Panchanadeeswaran 1995, Bate 1999). With the recent availability and advances in parallel computing technology, it has become possible to handle fairly large three-dimensional meshes (Beaudoin 1996, Sarma 1998, Dawson 1998).

In recent years, the above mesoscale deformation models have been coupled with other mesoscale techniques for simulating the evolution of the deformed microstructure during subsequent annealing (Radhakrishnan et al. 1995, Solas et al. 1999, Radhakrishnan et al. 2000, Raabe and Becker 2000). Such a coupled model has recently been used to predict the formation and evolution of particle-stimulated nuclei during recrystallization of a Goss-oriented single crystal containing a cube-shaped Si particle (Radhakrishnan et al. 2000). The simulations were able to capture not only the TD rotations seen in the experimental data of Ferry and Humphreys (1996), but also the shape of the deformation zone, the stored energy of deformation and the misorientations existing in the deformation zone and their gradients from the particle to the matrix. The deformation substructure extracted from FE simulations was evolved using a Monte Carlo technique. The simulations clearly captured the formation and growth of particle stimulated nuclei. The orientations of the nuclei corresponded to the extremes of the orientation spread in the TD direction found in the deformation zone.

In a polycrystalline alloy containing coarse particles, the shape of the deformation zones and the spatial distribution of the orientations within the deformation zone will depend upon the grain orientation, the particle shape and size. Therefore, the above microstructural parameters will influence not only the recrystallization kinetics but also the relative contributions of various nucleation mechanisms to the overall recrystallization texture. Another important microstructural parameter that applies to polycrystals is the microstructural location of the particle. For example, a particle situated completely inside a grain would generate a deformation zone significantly different from a particle located at a triple line or a grain boundary. Although such microstructural effects have been noticed in an indirect way in the form of a change in the final recrystallized texture as a function of initial grain and particle structures, quantitative descriptions of the deformation zone and the evolution of the recrystallization texture are not yet available.

The objective of the current research is to apply the previously developed coupled finite element–Monte Carlo approach to predict the influence of particle shape, initial grain orientation and the microstructural location of the particle on the recrystallization kinetics, as
well as to predict the resulting recrystallization textures. More importantly, the nucleation and growth of crystal orientations during recrystallization following hot deformation of particle-containing aluminum alloys will be investigated by considering the effect of the operation of non-octahedral slip systems during hot deformation.

15.2 Computational Approach

The computational approach, as in previous investigations involved the following three steps: (a) simulating the deformation at the microstructural length scale using a crystal plasticity based finite element technique, (b) extracting the deformation substructure corresponding to each volume element based on certain assumptions, and (c) evolving the deformation substructure using a Monte Carlo approach.

In the finite element technique, the simulation domain in the form of a single crystal, bi-crystal or a tri-crystal was discretized such that there were a large number of elements in each grain/crystal. The hard particle in the form of a sphere, a cube or a rotated-cube was also discretized in a similar fashion. The deformation conditions corresponded to a temperature of 673 K and a strain rate of 1.0 s\(^{-1}\). The material constants and model parameters were obtained by fitting a constitutive model to experimental stress-strain data. The hot deformation was carried out in the plane strain compression mode to a true strain of 0.4. Slip along both (111) and (110) slip planes was allowed in accordance with the observations of Maurice and Driver (1997). The finite element discretization was applied to a cube-shaped domain with 40 elements along each side. Velocity boundary conditions were applied to the nodes on the top surface that contains the transverse direction (TD) and the rolling direction (RD). On the faces that contain the rolling direction and the normal direction (ND), the nodes on the right were constrained to keep the face flat. On the left end, the displacements in the RD were constrained to be zero. Since periodic boundary conditions are not used, there is a lack of symmetry in the ND-RD plane. The slip system strength of the elements belonging to the hard particle was assigned to be three orders of magnitude higher than that of the matrix elements (well above the saturation value for the slip system strength) in order to prevent the particle from deforming.

The deformation substructure was extracted from the output of the FE simulations by making the following assumptions.

1. The stored energy of deformation within each element is obtained from the FE simulations by assuming it to be proportional to\(^2\),

\[
H \propto \sum \Delta \tau_{cr} \Delta \gamma
\]  

(15.1)

where \(\Delta \tau_{cr}\) is the increase in the slip system strength and \(\Delta \gamma\) is the increase in the effective slip system accumulated shear in each element during each deformation step.

The basis for using Equation (15.1) as a measure of the stored energy is the assumption

\(^2\) In our previous work (see Radhakrishnan et al. 2000) we have calculated stored energy as proportional to \(\tau^2\). However, this does not provide the correct ratio of the stored energies between Cube and S in a deformed Cube-S bi-crystal (see subsequent text).
that the incremental area in the stress-strain curve of each element arising from the incremental stress and incremental strain is proportional to the hardening provided by the increased dislocation density in the element.

2. The deformation substructure in each volume element is in the form of a collection of subgrains with a mean subgrain size and a mean misorientation among subgrains that depend on the local stored energy in the element. The stored energy per unit volume is related to the subgrain boundary energy per unit volume according to

\[ H = \frac{\gamma_0}{2d} \left[ \frac{\Delta \theta}{\theta^*} \left( 1 - \ln \frac{\Delta \theta}{\theta^*} \right) \right] \]  

(15.2)

where \( d \) is the mean subgrain size, \( \Delta \theta \) is the mean subgrain misorientation, \( \gamma_0 \) is the energy per unit area of a high-angle boundary and \( \theta^* \) is the misorientation limit for low-angle boundaries, taken as \( 15^\circ \). The orientations of the subgrains within each element are assumed to be randomly distributed with respect to the orientation of the element obtained from the FE simulations. In Equation (15.2) the mean subgrain size, \( d \) is an unknown because there is no length scale associated with the FE technique. However, an assumption can be made regarding the subgrain size based on experimental observations. Each element is subdivided into a number of regularly spaced sites in the three principal directions and a constant subgrain size is associated with these sites. The corresponding value of \( \Delta \theta \) is calculated for each element based on its stored energy value.

The deformation substructure representative of each element can be calculated by using Equations (15.1) and (15.2) for each element. The evolution of the deformation substructure is carried out using a Monte Carlo technique applied to subgrain growth. The Monte Carlo methodology used here is described in detail elsewhere (Radhakrishnan and Zacharia 2003). Realistic simulations of recrystallization using an abnormal subgrain growth approach requires proper handling of the energy and mobility of grain boundaries, especially for those that are close to the \( \Sigma 7 \) boundaries. In the current simulations the boundary properties are calculated based on the experimental results of Huang and Humphreys (1999). According to these authors, the boundary velocity for exact \( \Sigma 7 \) boundaries is roughly 10 times that of general high angle boundaries. The boundary velocity decreases roughly in a linear fashion to that of general high angle boundaries (HAGB) when the misorientation from exact CSL reaches about 20° and stays constant at higher misorientations. In the current simulations, the deviation of the boundary with respect to 40° \(<111>\) is first calculated. If the deviation in terms of the rotation angle is greater than 20° then the boundary is considered to be a general HAGB, and a relative boundary energy and boundary mobility of 1.0 are used. If the misorientation is less than 20° then the boundary velocity is calculated by assuming a linear drop in velocity from the ideal CSL location. It is further assumed that when the boundary misorientation increases from the exact CSL orientation, the boundary energy increases monotonically according to the Read-Schockley formulation up to a misorientation of 20°, beyond which it becomes equal to the energy of a general HAGB. Since the boundary velocity is proportional to the product of energy and mobility, the mobility is calculated by dividing the velocity by the boundary energy.
15.3 Simulations

The hot deformation and recrystallization simulations were carried out for three different starting materials—single crystals, bi-crystals and tri-crystals of aluminum. The single crystal orientations were either Cube (001)[100] or S (123)[634] or Copper (112)[111]. The bi-crystals were made up of combinations of the above three orientations consisting of S-Cube, S-Copper and Copper-Cube. A tri-crystal consisting of Copper, S and Cube was also used in the simulations. In the current simulations the effect of particle shape was investigated by employing hard particles in the form of a cube, a sphere or a rotated cube. The effect of microstructural location was investigated by placing the hard particle at the grain interior or at the grain boundary or at a triple point. In the case of the single crystals the hard particle was kept in the body center of the crystal. For bi-crystals, the particle was positioned at the center of the grain boundary that formed at the mid-thickness plane. In the case of the tri-crystal the particle was at center of the triple line formed by the three crystal orientations. Deformation substructures were extracted for each of the above cases for the mid-section along the TD, and the evolution of the substructure during subsequent annealing was simulated using the Monte Carlo technique. The focus of the simulations was to quantify the influence of the hard particles on the recrystallization texture and the recrystallization kinetics in each of the above cases.

15.4 Results and Discussion

15.4.1 Microstructure and Kinetics

Single Crystals: The single crystals were initially at the Cube, S or Copper orientations and the hard particles situated at the center of the crystal had the shape of either a cube, or sphere or a rotated cube. The formation of localized deformation zones with increased stored energy of deformation is shown in Figure 15.1 for the case of the Cube-oriented crystal containing a spherical hard particle. The early stages of recrystallization that lead to the formation of particle stimulated nuclei in the localized deformation zone near the particle during the recrystallization of an S-oriented crystal containing a spherical particle are shown in Figure 15.2. In this and subsequent figures showing grain structures the contrast is given by the rotation angle of the axis-angle pair representing the grain orientation. It is possible that within each grey level there are several grains that have the same rotation angle but a different rotation axis, although such occurrences are very unlikely, especially in the recrystallized regions. The subgrains are seen to grow very rapidly in the element indicated by arrow situated close to the particle. Careful analysis of the element orientations showed that the subgrains within the marked element were misoriented roughly 40° about <111> with respect to the adjoining matrix elements. This proves clearly the orientation selection mechanism proposed by Engler (1997a, 1997b). The growth sequence of the nuclei to consume the rest of the deformed matrix is shown in Figure 15.3.

The two important features that determine the shape of the deformation zone and hence the number and spatial distribution of the particle-stimulated nuclei are the particle shape and the initial grain orientation. Figure 15.4 shows partially recrystallized grain structures of single crystals of different starting orientations and different particle shapes. The effect of changing...
15.4 Results and Discussion

Figure 15.1: The formation of a localized deformation zone in the presence of a spherical hard particle in a Cube-oriented crystal following deformation.

Figure 15.2: Early stages of recrystallization showing the formation of PSN in a deformed S-oriented crystal with spherical particle.

Figure 15.3: Growth sequence of the particle stimulated nuclei shown in Figure 15.2.

the precipitate shape from a sphere to a cube in a single crystal initially at S-orientation can be seen by comparing fig 4a and 4b. The effect of changing the crystal orientation from S to Cube with a spherical hard particle can be seen by comparing Figures 15.4a and 15.4c. Both the spatial distribution and the number of active nuclei are seen to be influenced by
the crystal orientation and the particle shape. Previous analytical models were not able to capture these subtle changes although such effects were known to be important and a few experimental results did suggest the effect of particle shape on the overall recrystallization kinetics in particle-containing aluminum alloys (Engler et al. 2001a).

Figure 15.4: Partially recrystallized grain structures in S-oriented crystal with spherical particle (left), S-oriented crystal with cube particle (middle) and Cube-oriented crystal with spherical particle (right).

The current simulations also showed that single crystals of S, Cube and Copper without the hard particles did not recrystallize following hot deformation and annealing. They merely showed a uniform growth of subgrain structure whose kinetics was significantly slower than that of the particle-containing systems. In the presence of the particles, these single crystals readily recrystallized because of the formation and growth of PSN.

The particle shape had a significant effect on the recrystallization kinetics of the S-oriented crystal as shown in Figure 15.5. The spherical particle appears to result in the fastest recrystallization kinetics, followed by the Cube particle. The rotated cube particle has the slowest kinetics. In the case of the Copper-oriented crystal, the spherical particle is again seen to induce the fastest kinetics, although the cube and the rotated cube particles have roughly the same kinetics. It must be noted that in the presence of the spherical particle special boundaries close to $\Sigma 7$ were observed and nuclei that had such boundaries with respect to the deformed matrix were seen to grow rapidly. However, such boundaries were either non-existent or significantly deviated from $\Sigma 7$ orientation in the case of the Cube and rotated Cube particles. In the case of the Cube oriented crystal, all particles induced roughly the same kinetics, and the active nuclei did not have $\Sigma 7$ orientation with respect to the matrix. The above finding is somewhat different from the predictions of Engler for recrystallization following cold deformation, according to which the recrystallization nuclei always had a misorientation very close to a $40^\circ <111>$ with respect to the matrix. However, there are many investigations where the nuclei orientations were found to be quite close to the matrix investigations (Ørsund and Nes 1988, Jack et al. 1989, Humphreys and Ardakani 1994) and the nuclei were assumed to have formed either very close to the particle-matrix interface where the rotation is highly restricted because of slip blockage, or in the outer regions of the deformation zone. The current simulations show that the formation of nuclei close to the matrix can occur because of the particle shape effect which may be due to the effect of deformation temperature.
15.4 Results and Discussion

**Bi-crystals:** The recrystallization aspects of S-Cube, S-Copper and Cube-Copper bicrystals with and without hard particle were investigated. In the S-Cube bicrystal, the Cube is situated at the top and the S is at the bottom. The Cube is characterized by a dark contrast while the S has lighter contrast. In the absence of the hard particle, the Cube grain is seen to grow and consume the S grain during recrystallization due to the operation of a strain induced boundary migration (SIBM) mechanism at the S-Cube interface, as shown in Figure 15.6. The migration of the interface into the S is triggered by the lower stored energy inside the Cube than in the S in agreement with published experimental observations (Theyssier and Driver 1999).

**Figure 15.6:** Recrystallization of deformed Cube-S bicrystal showing S fully consumed by Cube.

The effect of hard particle situated on the grain boundary on particle stimulated nucleation and recrystallization is shown in Figure 15.7. In the presence of the hard particle, a nucleus of random orientation forms on the S-side of the bi-crystal and grows into the S grain as shown in the left figure. Interestingly, the nuclei that form on the Cube side near the particle do not grow as much, indicating that these nuclei are not bounded by special boundaries unlike those that form on the S side of the bi-crystal. The growth of the particle stimulated nucleus into the S grain is eventually curtailed as it gets completely surrounded by the growing Cube grain, leading to the formation of island grains contributed by PSN. Similar island grains, although much smaller in size are also seen on the Cube side. Such island grains have been observed experimentally by Humphreys in lightly deformed and annealed single crystal of aluminum in the Goss orientation containing coarse silicon particles (Ferry and Humphreys 1996), where oriented growth contributed significantly to the final recrystallized texture.
Figure 15.7: Partially recrystallized grain structures of deformed S-Cube bi-crystal (left), S-Copper bi-crystal and Copper-Cube bicrystal (right) containing spherical hard particle on the boundary.

In the S-Copper bicrystal, the S grain is at the bottom and the Copper grain is at the top. The nuclei that form in the Copper side do not grow significantly while the nuclei on the S side grow extensively and consume the entire S grain. In the case of the Copper-Cube bicrystal where the Cube is on top, the nuclei on the Cube side do not grow, while those on the Copper side grow significantly, consuming the entire Copper grain and transforming into a recrystallized grain of a new orientation.

The recrystallization kinetics of bi-crystals with and without particles is shown in Figure 15.8. The presence of particles leads to significant increase in the recrystallization kinetics of bi-crystals, for the Copper-Cube and S-Copper bicrystals. For the S-Cube case the difference is only marginal. These observations are in agreement with the development of recrystallized grain structures as shown in Figure 15.7.

Figure 15.8: Recrystallization kinetics of particle-containing bi-crystals.

Tri-Crystal: A partially recrystallized microstructure of a deformed S-Cube-Copper tricrystal with a spherical hard particle situated at the triple point is shown in Figure 15.9. The formation of new orientations in the vicinity of the hard particle is quite evident. In the particle-free tri-crystal, the interface between S and Cube was observed to migrate towards the S as in the bi-crystal case. However, the interface between Cube and Copper appeared to
15.4 Results and Discussion

be quite stationary. The recrystallization kinetics of the tri-crystal with the hard particle at the triple point was found to be slightly higher than for the particle-free tri-crystal.

Figure 15.9: Partially recrystallized grain structure of a Cu-S-Cube tri-crystal containing a spherical particle at the triple point.

15.4.2 Texture

The most significant effect of the presence of hard particles on texture is the formation of “random” orientations that are not present in the particle-free systems. Out of all the random orientations that are present in the localized deformation zone, the nuclei that ultimately win out are those that are close to special low-energy, high-mobility boundary with the surrounding matrix. The as-deformed and recrystallized <100> pole figures of a single crystal of S orientation containing a spherical hard particle are shown in Figure 15.10. In this and subsequent pole figures, the as-deformed textures are obtained by only considering orientations with a misorientation greater than 5° from the initial orientations, so that the texture would correspond to the region in the vicinity of the particle, rather than including the entire deformed matrix. The formation of a few orientations that are significantly deviated from the initial S orientation is seen in the pole figure on the left. The contour plot shows the recrystallized texture, on which the orientations obtained by rotating the initial S orientation by 40° about all possible <111> axes are superimposed. The rotation axis corresponding to each numbered set is shown in Table 15.1. It is clear that the orientation of the recrystallized nuclei are close to a 40°[111] rotation of S in this case.

Figure 15.11 shows the as-deformed and recrystallized texture plots for the case of a Cube grain containing a spherical particle. Note that in this case, the recrystallized texture components are not close to any of the <111> rotations. This result seems to indicate, that there is a significant effect of initial grain orientation and particle shape on the ability to form nuclei that are specially oriented to form a Σ7 boundary with respect to the matrix. As described previously, this result is not in agreement with the predictions of Engler for cold deformation and seems to agree with some of the experimental observations where the nuclei orientations where quite close to the matrix orientations.

Figure 15.12 shows the deformed and recrystallized textures for the S-Cube bi-crystal containing a spherical particle. One of the recrystallized texture components is close to 40° [111] rotation of S. This is in agreement with the grain structure in Figure 15.7 that shows the
Table 15.1: Rotation nomenclature

<table>
<thead>
<tr>
<th>Number</th>
<th>Rotation axis</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>111</td>
</tr>
<tr>
<td>2</td>
<td>111</td>
</tr>
<tr>
<td>3</td>
<td>111</td>
</tr>
<tr>
<td>4</td>
<td>111</td>
</tr>
<tr>
<td>5</td>
<td>111</td>
</tr>
<tr>
<td>6</td>
<td>111</td>
</tr>
<tr>
<td>7</td>
<td>111</td>
</tr>
<tr>
<td>8</td>
<td>111</td>
</tr>
</tbody>
</table>

Figure 15.10: Deformation and recrystallization textures for S oriented crystal with spherical particle.

...consumption of the S by the Cube. The orientation relationship between and S and Cube is close to a 40° [111] rotation of S. The island recrystallized grain that nucleated in the particle deformation zone has an orientation which is significantly deviated from any of the 40°[111] rotations of S. Since the interface between S and Cube is close to a 40°[111] rotation of S, the migration velocity of this interface to consume S is quite significant in comparison with that of the particle stimulated nucleus and therefore, the particle stimulated grain is eventually surrounded by the migration of the S-Cube interface, as shown in Figure 15.7.

In the case of the S-Copper bi-crystal, shown in Figure 15.13, the dominant new texture component that forms after recrystallization is close to a 40° [111] rotation of the S grain, which is in agreement with the formation and growth of the nuclei on the S side of the S-Copper bi-crystal as shown in Figure 15.7. In the Cube-Copper bi-crystal, the recrystallization components are close to a 40° [111] rotation of Copper, as shown in Figure 15.14.

The above texture predictions are in agreement with some of the general observations of Engler regarding the orientation selection mechanism operating in the deformation zone. However, it appears that the ability to form initial orientations that are close to these special orientations depends on the particle shape and the initial grain orientation. In some cases,
the recrystallized texture does not result in the formation of these specially oriented texture components. In these instances the nuclei orientations are closer to the initial matrix orientations, in agreement with several experimental observations, although in these experiments such effects are not explicitly attributed to particle morphology or initial grain orientation.

15.5 Summary

A mesoscale simulation approach based on coupling a crystal plasticity based finite element model with a Monte Carlo simulation of substructure evolution was used to investigate the deformation and recrystallization behavior of single crystals, bi-crystals and tri-crystals of aluminum containing coarse, non-deformable particles. The deformation simulations captured the size, shape and the spatial distribution of orientations in the localized deformation
Deformation and recrystallization of Particle-containing Aluminum Alloys

Figure 15.13: Deformation and recrystallization texture of S-Copper bicrystal with spherical particle. <111> rotations of Copper (middle) and S (right) are superimposed on the recrystallization texture.

Figure 15.14: Deformation and recrystallization textures of a Copper-Cube bi-crystal containing a spherical particle.

zone as a function of the particle shape, initial grain orientation in single crystals and microstructural location in the case of bi-crystals and tri-crystals. The Monte Carlo evolution of the hot deformation substructure extracted from the finite element results captured the formation and growth of nuclei in the deformation zone. While in many instances, the orientation of the particle stimulated nuclei were close to a 40°[111] rotation of one of the initial matrix orientations surrounding the particle, in some cases the nuclei orientations were quite close to the initial matrix orientations, thus indicating the effect of either particle shape or the initial matrix orientation or the deformation temperature itself on the orientations of the active nuclei stimulated by the particle. It is interesting to note that there are experimental observations that support the above nucleation modes after cold deformation, although such observations are not available for nucleation and growth after hot deformation. Also, the effect of particle shape or initial matrix orientation is not explicitly captured in these experiments. Further research on particle stimulated nucleation and growth involving specific grain orientations and
particle morphologies and deformation temperatures have to be carried out in order to validate the predictions of the current mesoscale simulations.

Acknowledgements

Research sponsored by the Division of Materials Sciences and Engineering, U.S. Department of Energy, under contract DE-AC05-00OR22725 with UT-Battelle, LLC. The authors would like to thank Drs. John Vitek and Judy Pang for a critical review of the manuscript.

References

Ashby M.F. 1970. Deformation of plastically non-homogeneous materials. Phil Mag. 21, 399
Becker, R., Panchanadeeswaran, S., 1995. Effects of grain interactions on deformation and local texture in polycrys-
Met Sci. 18, 527-529.
Engler, O., 1997. Influence of particle stimulated nucleation on the recrystallization textures in cold deformed Al-
Ferry, M., Humphreys, F. J., 1996. The deformation and recrystallization of particle-containing [011][100] alu-
Habiby, F., Humphreys, F.J., 1993. A microstructure investigation of the deformation and recrystallization of particle-
containing aluminum alloys. Texture Microstruct 20, 125-140.
Hansen, N., Bay, B., 1981. Initial stages of recrystallization in aluminum containing both large and small particles.
Acta Met. 29, 65-77.
Huang, Y., Humphreys, F.J., 1999. Measurements of grain boundary mobility and during recrystallization of a single-
References


16 Mesoscale Simulation of Grain Growth

David Kinderlehrer, Jeehyun Lee, Irene Livshits, and Shlomo Ta’asan

16.1 Introduction

The mesoscale simulation of grain growth consists in resolving a large coupled system of nonlinear evolution equations with appropriate boundary conditions which represents a network of interfaces. We present here an approach based on a dissipation principle and flexible enough to exploit the energy and mobility functions derived from recent experiments. Our objective is to simulate with sufficient accuracy to predict texture and at a sufficient scale to yield meaningful statistics. Since, generally, the result of such a simulation consists of the statistics it provides, we are led to the companion issue of coarse graining in mesoscale simulations. We seek to understand what statistics are reliable and how we may interrogate the dynamics of their distribution functions. There are many challenges to the successful execution of this program, from the algorithmic level where we are confronted with a highly nonlinear coupled system with many topological events or critical events, to the coarse grain level. For reasons of space, we defer discussion of this latter subject to a future work.

A number of collateral references are listed in the references. We would like to take this opportunity to thank A. D. Rollett and G. Rohrer and also G. Leoni, C. Liu, and P. Yu for their help. Supported by the MRSEC program of the NSF under award DMR 0079996 and NSF DMS 0072194, NSF DMS 9805582, NSF DMS 0305794 and the DoE Computational Materials Science Network.

Here we are concerned with the mesoscale simulation of large networks of grains or interfaces in two and three dimensions. The evolution is governed by the Mullins Equation of curvature driven growth. These equations, discussed below, are a system of evolution parabolic partial differential equations for each grain boundary curve, in two dimensions, or facet, in three dimensions. Grain boundaries meet, typically, at triple junctions, in two dimensions, or on triple lines, in three dimensions, where an additional boundary condition is required. We enforce the Herring Condition, a force balance, and below we show that it is the natural boundary condition for the Mullins Equation in equilibrium. Other boundary conditions are admissible but the Herring Condition is the simplest. The resulting system is then dissipative for the energy and the evolution can be viewed as a modified steepest descent for the total grain boundary energy. In addition, we must treat certain critical events. During evolution, a grain boundary or a grain may shrink and disappear. This creates unstable multiple junctions which split into triple junctions which we treat in a way that results in maximum energy reduction.

Our first objective is to review the Mullins Equation and to establish that the Herring Condition is its natural boundary condition. The two dimensional version of this has been
presented in Kinderlehrer et al. (2001). Consider a network of grains with facets which meet in triple lines. The most direct way to proceed is to begin with three facets represented as graphs over an \( x = (x_1, x_2) \) plane meeting along a triple line \( \Gamma' \),

\[
\begin{align*}
S^{(1)} : z = u^{(1)}(x), \quad x \in \Omega_+,
S^{(3)} : z = u^{(3)}(x), \quad x \in \Omega_-
\end{align*}
\]

\( u^{(1)} = u^{(2)} = u^{(3)} \) on \( \Gamma' \)

(16.1)

where \( \Gamma' \) denotes the projection of \( \Gamma \) onto the \( x \)-plane. Consider first a single facet, say \( S : z = u(x), \quad x \in \Omega (= \Omega_+) \).

The energy density of the facet \( S \) is given by

\[
\sigma(n), \quad n = \frac{1}{W}(-p_1, -p_2, 1), \quad \text{the normal to } S
\]

\( p_i = \frac{\partial u}{\partial x_i}, \quad W = \sqrt{1 + |p|^2} \) \hspace{1cm} (16.2)

\( \sigma \) is positively homogeneous of degree 0, and thus \( \nabla_n \sigma \cdot n = 0 \). The energy of \( S \) is

\[
E = \int_{\Omega} \sigma(n)W dx
\]

(16.3)

and equilibrium is determined by

\[
\delta E = 0.
\]

(16.4)

We give a brief derivation of the equilibrium equation which enables us to verify that the discrete version determined in the next section is an approximation of it. Let \( \eta \) be a test variation with compact support in \( \Omega \) and

\[
\begin{align*}
z &= u + \epsilon \eta, \\
E_\epsilon &= \int_{\Omega} \sigma(n_\epsilon)W_\epsilon dx, \\
n_\epsilon &= \frac{1}{W_\epsilon}(-\epsilon q_1, -\epsilon q_2, 1), \quad q = \nabla \eta \\
W_\epsilon &= \sqrt{1 + |p + \epsilon q|^2}
\end{align*}
\]

and

\[
\frac{dn_\epsilon}{d\epsilon} = -\frac{1}{W}(1 - n \otimes n)q \quad \frac{dW_\epsilon}{d\epsilon} = p \cdot q \quad \frac{dW}{d\epsilon} = \frac{p \cdot q}{W}.
\]

(16.5)

This leads to

\[
\begin{align*}
\frac{d}{d\epsilon} E_\epsilon &= \int_{\Omega} \left( \nabla_n \sigma(n) \frac{dn_\epsilon}{d\epsilon} W_\epsilon + \sigma \frac{dW_\epsilon}{d\epsilon} \right) dx \\
&= \int_{\Omega} \left( \nabla_n \sigma(n) \cdot (-q_1, -q_2, 0) + \sigma \frac{p \cdot q}{W} \right) dx \\
&= \int_{\Omega} \text{div} \left( \nabla_n \sigma(n) - \sigma \frac{p}{W} \right) \eta dx.
\end{align*}
\]

(16.6)
So equilibrium is determined by the equation
\[
\operatorname{div} \left( \nabla_n \sigma(n) - \sigma \frac{p}{W} \right) = 0 \quad \text{in} \quad \Omega
\] (16.7)

If
\[ S_t : z = u(x,t), \quad x \in \Omega, \quad t > 0, \]
is an evolving family of interfaces, its normal velocity is given by
\[
v_n = n \cdot \frac{\partial}{\partial t}(x_1, x_2, u) = \frac{1}{W} \frac{\partial u}{\partial t}
\]
and the governing equation of motion is the Mullins Equation
\[
v_n = \frac{1}{W} \frac{\partial u}{\partial t} = -\mu \operatorname{div} \left( \nabla_n \sigma(n) - \sigma \frac{p}{W} \right) \quad \text{in} \quad \Omega, \quad t > 0,
\] (16.8)
where \( \mu \) is a given mobility function.

To understand the boundary condition on the triple line \( \Gamma \), we must consider complete three dimensional variations \( \zeta = (\zeta_1, \zeta_2, \zeta_3) \) of the surface \( S \), rather than simply vertical ones. Here it is useful to introduce the variations separately as
\[
x = \xi \\
u = z + \epsilon \zeta_3
\] (16.9)
\[
x_1 = \xi_1 + \epsilon \zeta_1 \\
x_2 = \xi_2
\] (16.10)
\[
x_1 = \zeta_1 \\
x_2 = \zeta_2 + \epsilon \zeta_2 \\
u = z
\] (16.11)

The last two are variations of domain. We caution that although at first glance this procedure appears to permit three different equilibrium equations, they are, in fact, all the same. The result of the computation is a linear functional of \( \zeta \) given by, with the abbreviation \( F = \sigma W \),
\[
L(\zeta) = \int_\Omega \zeta \cdot n \ \operatorname{div} F_p \ W dx + \int_{\Gamma'} \zeta \cdot T(\nu) \ ds
\] (16.12)
where \( T(\nu) \) is the linear form on the normal \( \nu \) of \( \Gamma' \) given by
\[
T(\nu) = \begin{pmatrix}
F - p_1 F_{p_1} \\
-p_2 F_p & -p_1 F_{p_1} \\
F_{p_1} & F_{p_2}
\end{pmatrix} \nu
\]
Recalling that \( F = \sigma W \), we can decompose \( T(\nu) \) as
\[
T(\nu) = \frac{\sigma}{W} \begin{pmatrix}
1 + (p_2)^2 & -p_1 p_2 \\
-p_1 p_2 & 1 + (p_1)^2
\end{pmatrix} \nu + W \begin{pmatrix}
-p_1 \sigma_{p_1} & -p_1 \sigma_{p_2} \\
-p_2 \sigma_{p_1} & -p_2 \sigma_{p_2}
\end{pmatrix} \nu
\]
\[
= \sigma T_{iso}(\nu) + T_{an}(\nu),
\]
with \( T_{iso} \) the isotropic term corresponding to the \( T \) which arises for constant surface energy and \( T_{an} \) containing the torques. Let \( l = (l_1, l_2, l_3) = (-\nu_2, -\nu_1, -\nu_2 p_1 + \nu_1 p_2) \) denote the tangent direction to the curve \( \Gamma \). We then find that

\[
T_{iso} = \sigma n \times l \quad \text{and} \quad T_{an} = W^2(\sigma_{p_1} \nu_1 + \sigma_{p_2} \nu_2)n \tag{16.13}
\]

To clarify equilibrium we return to the system of three surfaces. Let \( \zeta^{(i)} \) denote the variation (vector) of surface \( S^{(i)} \). The variation is the sum of three versions of (9). Let us make the convention that \( n^1 \) and \( n^2 \) point upward and \( n^3 \) points downward. Then we find

\[
\sum_{i=1}^{2} \left( \int_{\Omega^+} \zeta^{(i)} \cdot n^{(i)} \ \text{div} \ F^{(i)} dx \right) + \int_{\Omega^-} \zeta^{(3)} \cdot n^{(3)} \ \text{div} \ F^{(3)} \ W^{(3)} dx \\
+ \sum_{i=1}^{3} \left( \int_{\Gamma'} \zeta \cdot T^{(i)}(\nu^{(i)}) \ ds \right) = 0 \tag{16.14}
\]

subject to \( \zeta^{(1)} = \zeta^{(2)} = \zeta^{(3)} \) on \( \Gamma' \)

Thus, in addition to the equilibrium of equations given by (16.7), we obtain the natural boundary condition, which is the Herring Condition,

\[
T^{(1)}(\nu^{(1)}) + T^{(2)}(\nu^{(2)}) + T^{(3)}(\nu^{(3)}) = 0 \tag{16.15}
\]

For example, in the case where \( \sigma \) is independent of \( n \), we immediately verify that

\[
n^{(1)} + n^{(2)} + n^{(3)} = 0,
\]

namely, that the surfaces meet at angles of \( 2\pi/3 \).

We shall discuss simulating a large network of evolving surfaces subject to (16.8) with (16.15). An important property of this system is that, in the absence of critical events, it is dissipative. To check this, observe that the total energy of the network is given by

\[
E(t) = \sum_{S} \int_{\Omega} \sigma_S(\nu) dS \tag{16.16}
\]

and computing \( dE/dt \) corresponds to setting \( \zeta = -\partial u/\partial t \) in the first variation. After substituting (16.8), we obtain that

\[
\frac{dE}{dt} = -\sum_{S} \int_{\Omega} \frac{1}{\mu_S} (v_n)^2 dS + \sum_{\Gamma} \int_{\Gamma} v \cdot \sum_{\Gamma} T ds \tag{16.17}
\]

Thus, when (16.15) holds, the integrals over the triple lines vanish and

\[
\frac{dE}{dt} = -\sum_{S} \int_{\Omega} \frac{1}{\mu_S} (v_n)^2 dS \leq 0
\]
16.2 Discretization

We proceed by presenting the discretization of the equations above using a dissipation principle. Given a triangulation $T = \beta$ of a surface $S$ consisting of triangles $\beta$, the discrete energy of $S$ is

$$E = \sum_{\beta \in T} \sigma(n_\beta)A_\beta$$

where $n_\beta$ is the normal to $\beta$ and $A_\beta$ is the area of $\beta$.

The energy rate of change resulting from a motion of the vertices of the triangles $\beta \in T$ is

$$\frac{dE}{dt} = \sum_{\beta \in T} \left[ \nabla_n \sigma(n_\beta) \cdot \frac{d\beta}{dt} A_\beta + \sigma(n_\beta) \frac{dA_\beta}{dt} \right]$$

(16.18)

Let $x_i, x_j$ and $x_k$ be three nodes of $\beta$, and $W = (x_j - x_i) \times (x_k - x_i)$. Then we may write

$$n_\beta = \frac{W}{|W|} \quad \text{and} \quad A_\beta = |W|$$

$$\frac{d\beta}{dt}|W| + n_\beta \frac{d|W|}{dt} = \frac{dW}{dt}$$

(16.19)

Combining (16.18) and (16.19) yields

$$\frac{dE}{dt} = \sum_{\beta \in T} \left[ \nabla_n \sigma(n_\beta) \cdot \left( \frac{dW}{dt} - n_\beta \frac{d|W|}{dt} \right) + \sigma(n_\beta) \frac{W}{|W|} \cdot \frac{dW}{dt} \right]$$

$$= \sum_{\beta \in T} \left[ \left( \nabla_n \sigma(n_\beta) + \sigma(n_\beta) n_\beta \right) \cdot \frac{dW}{dt} \right]$$

(16.20)

We compute

$$\frac{dW}{dt} = \left( \frac{dx_i}{dt} - \frac{dx_j}{dt} \right) \times (x_k - x_i) + (x_j - x_i) \times \left( \frac{dx_k}{dt} - \frac{dx_j}{dt} \right)$$

and let $H = \nabla_n \sigma + \sigma n_\beta$, then

$$\frac{dE}{dt} = \sum_{\beta \in T} \left[ H \cdot \left( \left( \frac{dx_i}{dt} - \frac{dx_j}{dt} \right) \times (x_k - x_i) \right) - H \cdot \left( \left( \frac{dx_k}{dt} - \frac{dx_j}{dt} \right) \times (x_j - x_i) \right) \right]$$

$$= \sum_{\beta \in T} \left[ - \left( \frac{dx_j}{dt} - \frac{dx_i}{dt} \right) \cdot (H \times (x_k - x_i)) + \left( \frac{dx_k}{dt} - \frac{dx_j}{dt} \right) \cdot (H \times (x_j - x_i)) \right]$$

$$= \sum_{\beta \in T} \left[ H \times (x_k - x_j) \cdot \frac{dx_k}{dt} + H \times (x_j - x_i) \cdot \frac{dx_j}{dt} + H \times (x_j - x_i) \cdot \frac{dx_k}{dt} \right]$$
Introducing notation $t_{\beta l} = x_k - x_j$, this is reduced to

$$
\frac{dE}{dt} = \sum_{\beta \in T} \sum_{x_l \in \beta} \left< H \times t_{\beta l}, \frac{dx_l}{dt} \right>
$$

$$
= \sum_{\beta \in T} \sum_{x_l \in \beta} \left< H \times (n_{\beta} \times m_{\beta l}) |t_{\beta l}|, \frac{dx_l}{dt} \right>
$$

$$
= \sum_{\beta \in T} \sum_{x_l \in \beta} \left< (H \cdot m_{\beta l}) n_{\beta} |t_{\beta l}| - (H \cdot n_{\beta}) m_{\beta l} |t_{\beta l}|, \frac{dx_l}{dt} \right>
$$

We obtain

$$
\frac{dE}{dt} = \sum_{x_l \in \text{Nodes}} \sum_{\beta_l} \left< \left( \nabla_n \sigma(n_{\beta}) \cdot m_{\beta l} \right) |t_{\beta l}| - \sigma(n_{\beta}) m_{\beta l} |t_{\beta l}|, \frac{dx_l}{dt} \right> (16.21)
$$

where $\beta_l$ is the set of triangles in $T$ involving the node $x_l$.

This gives the discrete form of the evolution equations,

$$
v_n = - \sum_{\beta_l} \left[ (\nabla_n \sigma(n_{\beta}) \cdot m_{\beta l}) |t_{\beta l}| - \sigma(n_{\beta}) m_{\beta l} |t_{\beta l}| \right] (16.22)
$$

which clearly are dissipative.

To see the relation between these expression and the continuous version given in the previous section, we use the tangential divergence theorem, and the tangential Green’s theorems,

$$\int_{\Omega} \text{div}_\Omega u \, d\Omega = \int_{\Omega} u \cdot n \, d\Omega + \int_{\partial\Omega} u \cdot m \, d\Gamma$$

$$\int_{\Omega} \nabla_\Omega \varphi \, d\Omega = \int_{\Omega} \varphi n \, d\Omega + \int_{\partial\Omega} \varphi m \, d\Gamma$$

where $u, n, m$ are vectors and $\varphi$ is a scalar. This gives,

$$v_n = - \text{div}(\nabla_n \sigma(n) + \sigma n)$$

which agrees with formula (16.8).

### 16.3 Numerical Implementation

We describe here the details of our numerical scheme. The basic objects involved in three dimensional simulation are grain boundaries, triple lines and grains, (in two dimension we use triple junctions, and grain boundaries). Grain boundaries and triple lines are discretized and evolved in our simulation. Unlike several other methods, in this method we do not discretize grains; they are determined by the triple lines and grain boundaries. Our discretization of surfaces and triple lines uses second order finite difference approximations, with either explicit or implicit time stepping. The time evolution of nodal points is done in two steps. First, grain boundaries are moved by Mullins Equation, and then triple lines are moved to enforce the Herring Condition. As grain boundaries and triple lines move, critical events must be considered to reflect actual changes in the physical topology, such as collapse and disappearance of individual grains, as well as to avoid element collapse in the simulation.
Grain boundaries are defined by a collection of triangles consisting of three nodal points and evolution is done by moving the nodal points according to (16.22). In the case of isotropy, $\sigma = 1$, it reduces to

$$ v_n = \sum_{\beta_l} R_{\beta_l} $$

(16.23)

where $\beta_l$ has three nodal points $x_i, x_j, x_k$ and

$$ R_{\beta_l} = \frac{1}{4A_\beta} \left[ -|x_k - x_j|^2 x_l + (x_k - x_j) \cdot (x_k - x_l) x_j + (x_j - x_k) \cdot (x_j - x_l) x_k \right] $$

With an explicit time scheme, this gives the discrete evolution equation

$$ x_l(t + \tau) = x_l(t) + \tau \sum_{\beta_l} R_{\beta_l}(t) $$

(16.24)

where $\tau$ denotes the time step.

Now, we consider an implicit time scheme to accelerate the computation. Let $N$ denote the number of nodal points of the grain boundary $S$ and

$$ x(t) = [x_1(t), x_2(t), \cdots, x_N(t)]^T $$

Define an operator

$$ Lx(t) = \left[ \sum_{\beta_1} R_{\beta_1}(t), \sum_{\beta_2} R_{\beta_2}(t), \cdots, \sum_{\beta_N} R_{\beta_N}(t) \right]^T $$

Then an implicit method requires solving the system of equations

$$ \left( I - \frac{\tau}{2} L \right) x(t + \tau) = \left( I + \frac{\tau}{2} L \right) x(t) $$

(16.25)

and it is approximated by applying Gauss-Seidel iterations.

Triple lines are defined by a set of nodal points and the Herring Condition is imposed on these points. The Herring Condition at a nodal point $x_i$ is given by

$$ \sum_{\beta_l} R_{\beta_l} = 0 $$

(16.26)

where $\beta_l$ is the set of triangles in $T$ involving the node $x_i$. And it is also approximated by Gauss-Seidel iterations.

### 16.3.0.1 Critical Events

During evolution, three types of critical events occur in three dimensions: loss of grains, loss of facets and loss of triple lines. The critical events are detected by monitoring the size and the rate of changes of topological components (i.e. grains, facets and triple lines). Components that are shrinking fast compared to their size trigger critical events, unless they are growing. This scheme follows Kuprat (Kuprat, 2000).
**Loss of grains:** In this event a neighboring grain, denoted by $w_g$, will absorb the small target grain $t_g$. The choice of $w_g$ is made as follows. We first pick the smallest facet of the target grain $t_g$. Then consider all the facets of the grain $t_g$ that have a common triple line with that small facet. We pick from these facets the one with the largest area. It has two grains on its two sides. One of them is our target grain. The other is the neighboring grain $w_g$ that will absorb.

**Loss of facet:** Implementation of this event involves two processes, 1) creation of a new volume, 2) loss of grain event. Creation of a new grain is done by growing the target facet, denoted by $t_s$, in the normal direction, into a new grain $n_g$. The new grain is built using a partial volume of two grains meeting at the target facet $t_s$. A loss of grain event is then performed using all neighboring grains except the two which contributed to growing the new grain.

**Loss of triple line:** Basically, the same idea as in loss of facet applies to the loss of triple line. We build a new volume $n_g$ around the target triple line $t_l$ using a partial volume from the three grains meeting at that line. We then perform a loss of grain event using all neighboring grains except those three that contributed to growing the new grain.

There are challenges to implement the critical events. First, the general algorithms for critical events fail when a two boundary facet is involved in the event. And this case should
16.4 Numerical Results

We illustrate our method using some numerical results in two and, briefly, in three dimensions. For simplicity and ease of comparison, we confine ourselves to the case where the energy density $\sigma = \text{const.}$ and the mobility $\mu = 1$. In two dimensions, we may typically begin from a 25,000 to 50,000 Voronoi-type initial configuration. We are able to verify these diagnostics: the average area grows linearly with time, the Mullins-von Neumann $n - 6$-rule holds for individual grains (not undergoing a critical event), and a ‘round’ grain shrinks inversely to its radius at the proper rate (to second order in spatial discretization.) For the general large scale simulation, we may plot relative area histograms, as in Figure 16.4. below. These show a transient phase and then remarkable self-similarity. More details are available in Kinderlehrer et al. (submitted).

Other statistics may be collected as well, for example, the fraction of grains with a given number of sides, illustrated in Figure 2 at a time step when the relative area histogram is stationary.

An important feature of the our technique is the ability to simulate with anisotropic energy densities $\sigma = \sigma(n, \alpha)$, that is, energies which depend both on the normal to the curve $n$ and the lattice misorientation $\alpha$. Lack of space prevents us from discussing this, but a glimpse of the new and interesting phenomena we are able to investigate is given in Ta’asan et al. (2003).

In three dimensions we are presently implementing large scale simulations. On average, grains with a large number of facets, say greater than 14, grow and grains with few facets shrink. An example of a critical event is portrayed below.

Figure 16.3: 1. the short horizontal triple line $tl$ is subject to removal; 2. a new volume $ng$ is built; 3. $ng$ is merged into $wg$ due to the loss of grain; 4. curvature driven motion have straightened the boundaries
16.5 Conclusion

In this paper we have presented a consistent variational approach to the mesoscale simulation of grain growth in three dimensions. We have presented results for a ‘bamboo structure’, a system with constant energy and mobility. The method accurately computes curvature driven growth with the Herring Condition imposed at triple junctions. The issues we had to resolve to accomplish this include the correct discretization scheme and the treatment of critical events.
Figure 16.6: Grain 73 has grown while grain 57 has shrunk. The figure shows the loss of grain 57 into grain 73. Faces are rendered as polyhedra but are, in fact, curvilinear.

An important feature of the mesoscale approach is that it respects the thermodynamic formulation given by Mullins and Herring. The reduced dimensionality of its data structure permits simulation of large scale systems, indeed, systems sufficiently large that we may inquire about their statistics. We found that the statistics produced by our simulation are very robust.

Future work can exploit this method to investigate interface dominated properties, for example, questions about anisotropy, texture, and abnormal grain growth, as mentioned in the Introduction. As a first project, we are exploring the role of anisotropy.

References


M. Upmanyu, R.W. Smith, D.J. Srolovitz (1998), Atomic Simulation of Curvature Driven Grain Boundary Migra-


17 Dislocation Dynamics Simulations of Particle Strengthening

Volker Mohles

Dislocation glide in a ductile matrix with particles of a secondary chemical phase is computer simulated. From these simulations the critical resolved shear stress (CRSS) is derived. The simulations are based on the local equilibrium of resolved stresses along the dislocations in one glide plane. Within this continuum-elastic approach, most accurate models are used. The elastic interaction of dislocations with each other and with themselves is fully allowed for (Brown 1964, Bacon 1967), i.e. no line tension is assumed. The spherical particles have a radius distribution and a spatial arrangement very close to that of a real Ostwald-ripened crystal. The dislocations glide in one plane (no cross-slip, no climb), but the interaction between particles and dislocations is modeled in three dimensions. Examples are given for dispersion strengthening, order strengthening and lattice mismatch strengthening. For some mechanisms, analytical expressions for the CRSS have been derived from the simulations. They contain no unknown parameters. The simulation results of lattice mismatch strengthening are compared to measurements on a corresponding real crystal. Further chances of dislocation dynamics simulations for particle strengthening are outlined.

17.1 Introduction

Most structural materials derive their strength from particles of secondary chemical phases which impede dislocation glide. The particles increase the material strength throughout the deformation process as long as the material remains ductile. Especially, the particles make a strong contribution, \( \tau_p \) (particle stress), to the critical resolved shear stress (CRSS) of single crystals. The amount of \( \tau_p \) has been widely investigated; a detailed overview of the basic processes and more refined approaches has been given by Nembach (1996). A short draft of the developments is given here in order to point out the deficiencies of the hitherto results from literature and the value and the chances of the present computer simulations.

In the first models of particle strengthening, the particles were considered as point obstacles in the glide plane, and a dislocation was modeled as a string under tension. For this tension a dislocation line energy \( E \), which is a force, was assumed. Each obstacle was able to impose a certain maximum force \( F_{\text{max}} \) on a dislocation which impedes dislocation glide. For weak obstacles \( (F_{\text{max}} \ll 2E) \), the Friedel (1956) model yields the particle stress \( \tau_p \) analytically as a function of \( E \), \( F_{\text{max}} \), and of the number of obstacles per unit area. Strong obstacles \( (F_{\text{max}} \gtrsim 2E) \) are circumvented by dislocations; for this case Orowan (1948) has derived an
analytical expression for $\tau_p$. Later refinements of these basic theories may be divided into three categories:

(i) Improvements of the dislocation model: De Wit and Koehler (1959) have drawn a distinction between line energy and line tension for screw and edge dislocations. Brown (1964) and Bacon (1967) have introduced the concept of the elastic dislocation self-interaction, which fully replaces the line tension approximation. Later this concept has been extended from isotropic to anisotropic elasticity (Barnett et al. 1972, Scattergood and Bacon 1975), and effects of the dislocation core (Prinz et al. 1978) and dislocation dissociation (Bacon 1978, Duesbery et al. 1992) have been considered. Moreover the concept has been extended to three dimensions (see Chapter 8 by H. M. Zbib). Schwarz and Labusch (1978) have assigned a mass (per unit length) to a dislocation and hence investigated inertial effects by computer simulations.

(ii) Improvements of the obstacle models: To calculate $\tau_p$ for real obstacles, the basic approach is to derive the maximum obstacle force $F_{\text{max}}$ from realistic obstacle models. This single parameter is then utilized by applying geometrical theories like that of Friedel (1956) or derivations thereof (e.g. Nembach 1996). More accurate approaches allow for more obstacle parameters. The effects of the finite obstacle extension along the dislocation line has been considered by Ham (1968). Mott and Nabarro (1948) have stressed the importance to account for the finite obstacle extension normal to the dislocation line especially for weak obstacles. Schwarz and Labusch (1978) have generalized this approach in computer simulations by using obstacles which are described by a force profile (a function) normal to the dislocation. Simulations with obstacles extended in all directions have been presented by Rönnapagel and co-workers (Fuchs and Rönnapagel 1993, Pretorius and Rönnapagel 1994).

(iii) Improvements of the obstacle arrangements: The effects of a random obstacle arrangement in the glide plane instead of a regular one have been worked out in the pioneering computer simulations of Foreman and Makin (1966, 1967). Aspects of a real three-dimensional particle arrangement have been discussed by Nembach (1996): even with all particles being equally large, the obstacle sizes which are effective in the glide plane have a certain distribution because the particles intersect the glide plane at various heights. Moreover, the obstacle strengths have a distribution, too. This can make it quite difficult and laborious to calculate the CRSS, as can be seen from publications on lattice mismatch strengthening (Gerold and Haberkorn 1966, Brown and Ham 1971).

Most of the quoted improvements to the basic models of particle strengthening have been considered and tested singularly but not in combination, so that uncertainties remain. Currently predictions of the strengthening contributions of particles are no more accurate than by a factor two. This conclusion must be drawn from Nembach’s (1996) comparisons of theories and computer simulations with experimental data. However with the computational power available today it is possible to combine all the refined models quoted above. This provides a valuable means to improve the understanding of strengthening mechanisms and to make accurate predictions. Like in all works quoted above thermally activated processes like cross-slip and climbing, which require a three dimensional dislocation model, are disregarded. Usually
these mechanisms don’t govern the CRSS. Examples are given subsequently for dispersion strengthening, order strengthening and lattice mismatch strengthening.

17.2 Simulation Method

The following description of the simulation method should, in principle, enable the readers to implement their own simulations; however more details have been described elsewhere (Mohles, 2001a).

17.2.1 Basis of the Method

In the present simulations a dislocation is considered as an abstract flexible line in one glide plane; the actual shearing process is disregarded. This line moves according to the resolved stresses \( \tau \) which derive from the external load, from the particles and from dislocations, including itself. A dislocation is represented by a polygon. The absolute value and the sign of the velocity normal to the dislocation line, \( v^\perp \), is calculated for each segment of this polygon.

\[
v^\perp = \frac{b}{B} \left( \tau_{\text{ext}} + \tau_{\text{disloc}} + \sum_k \tau_{\text{obst},k} \right)
\]  

(17.1a)

The external stress \( \tau_{\text{ext}} \) is the same for all segments of a dislocation. As detailed later on, the CRSS \( \tau_p \) is derived from \( \tau_{\text{ext}} \). The stress \( \tau_{\text{disloc}} \) describes the elastic interaction of a dislocation with other dislocations as well as with itself. Hence this term includes the resistance of a dislocation against bending. \( \tau_{\text{disloc}} \) is detailed in Section 17.2.3. The obstacle stress is comprised of the contributions \( \tau_{\text{obst},k} \) of all particles \( k \) considered; details are given in Sections 17.3 and 17.4. The fraction \( (b/B) \) is the coefficient of a viscous drag stress \( \tau_{\text{drag}} \), with \( (b/B)\tau_{\text{drag}} = -v^\perp \). With this latter equation, it gets obvious that Equation (17.1a) is an equilibrium of resolved stresses; the sum of all stresses vanishes. Other stress contributions can be included in Equation (17.1a), for instance in order to consider inertial effects (Schwarz and Labusch 1978), thermal activation (Labusch 1989) or a stacking fault (Bacon 1978, Duesbery et al. 1992). \( b \) denotes the magnitude of the Burgers vector, and \( B \) is the viscous drag coefficient as defined in literature (e.g. Schwarz and Labusch 1978). At present, the choice of \( B \) is arbitrary because it only affects the time scale. This time scale is meaningless because only static solutions will be looked for at present. Equation (17.1a) defines the sign and the magnitude of the velocity for each location \( x \) on the dislocation. The direction of the motion is normal to the local line vector \( s \). This determines the local velocity vector

\[
\dot{x} = v^\perp(x) s^\perp(x)
\]

(17.1b)

Here \( s^\perp \) is the unit vector normal to \( s \). Together, Equations (17.1a) and (17.1b) form a partial differential equation. It is solved numerically using a Runge-Kutta-type integrator. For this, a discrete representation for the dislocation is needed.
17.2 Simulation Method

17.2.2 Dislocation Segmentation

In the present simulations a dislocation has two representations. The basic representation is a set of \( N \) vectors \( x_i \) in the glide plane, \( 1 \leq i \leq N \), which point from the origin to points on the dislocation. Adjacent indices \( i \) and \( i + 1 \) denote neighboring points. These points are called points of integration (PI) in the following. The set of PIs is the basic representation because Equations (17.1a) and (17.1b) are evaluated for all these locations \( x_i \) on the dislocation. The mean distance between the PIs is denoted as \( \bar{s} \). The secondary representation is that of a polygon consisting of \( 2N \) short straight dislocation segments. It is needed to calculate \( \tau_{\text{disloc}} \) (Section 17.2.3). The polygon is derived from the PIs as follows:

(i) A circle through the points \( x_{i-1}, x_i, \) and \( x_{i+1} \) is constructed for each PI \( i \) as indicated in Figure 17.1(a). The local line direction \( s_i^0 = s_i / |s_i| \) and the local radius of curvature, \( R_{\text{curv},i} \), are derived from this circle. The direction \( s_i^{\perp 0} \) normal to \( s_i^0 \) is defined counterclockwise in the plane (Figure 1(b)). \( s_i^{\perp 0} \) is used in Equation (17.1b) for \( s_i^{\perp 0}(x_i) \); \( R_{\text{curv},i} \) is needed in Section 17.2.3.

(ii) A straight dislocation segment with the direction \( s_i^0 \) is assigned to PI \( i \) so that this PI lies in the center of the segment. This is called a main segment. In Figure 17.1(b) the main segments are drawn as thick lines. The length of the main segments is chosen to be \( s_i / 2 \), where \( s_i = \frac{1}{2} |x_i - x_{i-1}| + \frac{1}{2} |x_{i+1} - x_i| \) is the mean distance to the left and the right neighbor.

(iii) The end points of the main segments are connected by intermediate segments so that a closed polygon results. The intermediate segments have no PI in their centers; in Figure 17.1(b) they are drawn as thin lines for distinction.

With this segmentation there are twice as many segments as PIs; to each PI belong one main segment and two half intermediate segments on the left and the right. This segmentation may seem quite laborious, but it avoids possible errors (see Section 17.2.3) and numerical instabilities (Duesbery et al. 1992, Mohles 2001a) involved with more simple segmentations like that of Bacon (1967). In the following, the index \( i \) is always used for PIs \( (1 \leq i \leq N) \); the index \( j \) is used for segments \( (1 \leq j \leq 2N) \). As indicated in Figure 17.1(b), the starting...
and ending points of segment \( j \) are denoted by \( x_a^j \) and \( x_b^j \), respectively. It is emphasized that the mean length of segments, main and intermediate, is only \( s/2 \).

### 17.2.3 Dislocation Self-Interaction

The dislocation stress \( \tau_{\text{disloc}} \) in Equation (17.1a) is calculated from the stress tensor \( \sigma_{\text{disloc}} \) of a general closed dislocation loop in an elastically isotropic material. No distinction needs to be made whether the source of \( \sigma_{\text{disloc}} \) is another dislocation or the same one that senses the resolved stress \( \tau_{\text{disloc}} \) (elastic self-interaction). If the stress sensing PI \( i \) is located in the origin, i.e. \( x_i = (0, 0, 0) \), then the components \( \sigma_{\text{disloc},\alpha\beta} \) of \( \sigma_{\text{disloc}} \) are (e.g. Hirth and Lothe 1982 or Chapter 8):

\[
\sigma_{\alpha\beta,\text{disloc}} = -\frac{G}{8\pi} \oint_{\text{loop}} b_m \varepsilon_{nm\alpha} \frac{\partial}{\partial x_n} \nabla^2 |x| dx_{\beta} - \frac{G}{8\pi} \oint_{\text{loop}} b_m \varepsilon_{nm\beta} \frac{\partial}{\partial x_n} \nabla^2 |x| dx_{\alpha} 
\]

(17.2)

The indices \( \alpha, \beta, k, m \) and \( n \) stand for the directions \( x_1, x_2 \) and \( x_3 \). Einstein’s summation convention is applied. \( \delta_{mn} \) is the Kronecker symbol and \( \varepsilon_{kmn} = \omega_k \cdot (\omega_m \times \omega_n) \) is the Einstein permutation operator. \( G \) and \( \nu \) are the shear modulus and Poisson’s ratio, respectively. \( b_m \) denotes the Burgers vector components of this dislocation loop. To derive the resolved stress \( \tau_{\text{disloc}} \), the Peach-Koehler (1950) formula is applied to \( \sigma_{\text{disloc}} \).

\[
\tau = b^0 \cdot \sigma \cdot n^0 
\]

(17.3)

Here \( n^0 \) is the unit vector normal to the glide plane and \( b^0 \) is the direction of the Burgers vector of the stress receiving dislocation. This Burgers vector may differ from that of the stress source used in Equation (17.2) for instance when Shockley partial dislocations are simulated.

For the present simulations a dislocation is decomposed into short straight segments (Section 17.2.2). The self-stress \( \tau_{\text{disloc},i} \) exerted on PI \( i \) in Equation (17.1a) is calculated as the sum of the contributions \( \tau_{\text{seg},j \rightarrow i} \) of all segments \( j \) (main and intermediate) of all dislocation polygons. Only those segments within a certain distance \( R_{\text{cut-off}} \) (see below) from \( x_i = (0, 0, 0) \) are taken into account.

\[
\tau_{\text{disloc},i} = \sum_j \tau_{\text{seg},j \rightarrow i} = \sum_j \tau_{\text{seg}}(x^a_j, x^b_j - x_i, b_i^0, b_j) 
\]

(17.4)

where \( j \) complies with

\[
|x_j - x_i| < R_{\text{cut-off}} 
\]

\( x_j = \frac{1}{2}(x^a_j + x^b_j) \) is the center of segment \( j \) spanning from \( x^a_j \) to \( x^b_j \) (Figure 17.1). The function \( \tau_{\text{seg}}(\Delta x^a_j, \Delta x^b_j, b_i^0, b_j) \) is calculated from Equations (17.2) and (17.3). The integration in Equation (17.2) is performed analytically for a general straight dislocation segment. When
the glide plane is chosen to be $\mathbf{n}^0 = (0, 0, 1)$ and all segments and their Burgers vectors are within this plane the integration yields

$$
\tau_{\text{seg}, j \rightarrow i} = \frac{G}{4\pi} \frac{1}{D_{ij}} \left( (b_j \cdot h_{ij})(b_i^0 \cdot s_j^0) + \frac{(b_i^0 \cdot h_{ij}^+)(b_j \cdot s_j^+)}{1 - \nu} \right)
$$

(17.5a)

with

$$
s_j = x_j^b - x_j^a
$$

(17.5b)

$$
h_{ij} = (x_i^b - x_i^0) - (x_j^b - x_j^0)
$$

(17.5c)

$$
D_{ij} = (x_j^b - x_i) \cdot (x_j^a - x_i)^\perp
$$

(17.5d)

Again, the vectors $s_j^\perp$, $h_{ij}^\perp$ and $(x_j^b - x_i)^\perp$ are normal to and of the same length as their counterparts $s_j$, $h_{ij}$ and $(x_j^b - x_i)$. The evaluation of Equation (17.5a–d) is much simpler and faster than the stress formula for a general straight segment in three dimensions (Chapter 8). However $\tau_{\text{seg}, j \rightarrow i}$ is not well-defined because the integrand in Equation (17.2) is not unequivocal. For example the first authors who applied the self-interaction concept in computer simulations (Brown 1964, Bacon 1967), have used a different approach. Their stress formula for a single straight dislocation segment differs from Equation (17.5a–d), but when summed up according to Equation (17.4) for a closed dislocation, the resulting stress $\tau_{\text{disloc}, i}$ equals the present result of Equations (17.4) and (17.5a–d) (Mohles 2001a). For the present computer simulations, $\tau_{\text{seg}, j \rightarrow i}$ of Equation (17.5a–d) is preferred for its faster numerical evaluation.

If $x_i$ lies on the tangent of segment $j$ (but not on the segment itself), the denominator $D_{ij}$ in Equation (17.5a) vanishes. In this case the non-divergent approximation $\tau_{\text{tangent}, j \rightarrow i}$ (Equation (17.5e)) is used instead of $\tau_{\text{seg}, j \rightarrow i}$ (Equation (17.5a)).

$$
\tau_{\text{tangent}, j \rightarrow i} = \frac{G}{4\pi} \frac{\nu}{1 - \nu} \left( b_i^0 \cdot s_j^0 \right) \left( x_j^b - x_i \right)^{-1} \left( x_j^a - x_i \right)^{-1}
$$

(17.5e)

On the dislocation itself $\tau_{\text{disloc}}$ diverges; the linear elastic continuum model is not applicable in the dislocation core. In order to deal with the local interaction of a dislocation with itself, in the calculation of $\tau_{\text{disloc}}$ the dislocation core is “cut out”. To do this, two ways have been established. Instead of calculating $\tau_{\text{disloc}}$ on the dislocation, Brown’s (1964) approach was to calculate the stresses in front of and behind the dislocation in equal distances $R_{\text{core}}$; then the mean value of these stresses was taken as $\tau_{\text{disloc}}$. The energy of the dislocation core can be attributed to $R_{\text{core}}$. Bacon (1967), followed by others, has used a different but equivalent method to deal with the stress divergence: For each PI $i$, a piece of dislocation around $x_i$ of the length $L_{\text{core}}$ is cut out of the calculation. The physical meaning of the cut-out length $L_{\text{core}}$ is equivalent to that of $R_{\text{core}}$. But with a piece of dislocation being cut out, the integral of Equation (17.2) is no longer closed. In general this may lead to errors of any magnitude because the integrand of Equation (17.2) and hence $\tau_{\text{disloc}}$ is not unequivocal (Chapter 8). However with the segmentation procedure of Section 17.2.2, possible errors are fully avoided because, by definition, each PI lies in the center of a (main) segment where the self-stress of this segment unequivocally vanishes for symmetry reasons. Hence, this segment may be seen as being cut out, but still the integral of Equation (17.2) is closed (Mohles 2001a).
A straight dislocation segment does not interact with itself, but still its length \( s_j \) defines the distance to the neighbor segments and thus affects the self-interaction of a dislocation. If only Equations (17.4) and (17.5a–d) were used to calculate \( \tau_{\text{disloc}} \), then the local segment length \( s_j \) would be closely related to \( L_{\text{core}} \) and hence to the dislocation core energy. But while \( L_{\text{core}} \) has a physically fixed meaning, the lengths \( s_j \) vary for two reasons: firstly, the individual lengths vary during a simulation run, and secondly, the mean segment length (or, the density of PIs) must be chosen depending on simulation parameters (Section 17.2.4). To allow for an independent choice of \( L_{\text{core}} \) and \( s_j \) an additional stress term \( \tau_{\text{arc}} \) is needed for the segment \( j \) the center of which is \( \Pi_i \). The corresponding PIs and segments are denoted by \( j \leftrightarrow i \). In Figure 17.1b for instance, \( \Pi_i \) corresponds with segment \( j+3 \).

\[
\tau_{\text{arc}}(j \leftrightarrow i) = K \frac{G}{b} \frac{b}{R_{\text{curv},j}} \ln \frac{s_j}{L_{\text{core}}}
\] (17.6)

with

\[
K = \frac{1 + \nu}{1 - \nu} \left( b \cdot s \right)^2 + \frac{1 - 2\nu}{1 - \nu} \left( b^* \cdot s^* \right)^2
\]

The local curvature radius \( R_{\text{curv},j} \) is derived from the circle construction in Figure 17.1a. \( \tau_{\text{arc}} \) replaces \( \tau_{\text{seg},j \rightarrow i} \) in Equation (17.4) if \( j \leftrightarrow i \). This has been shown to be fully consistent with the present segmentation procedure (Mohles 2001a). It is also equivalent to Bacon’s (1967) approach to describe the local self-interaction by a short dislocation arc (therefore the name \( \tau_{\text{arc}} \)). However, the meaning of the length of this arc is slightly different from that of \( L_{\text{core}} \) due to the different segmentation procedures. For the present simulations a reasonable (Hirth and Lothe 1982) constant value \( L_{\text{core}} = 2b \) is used.

In Equation (17.4), all segments \( j \) at distances larger than \( R_{\text{cut-off}} \) from \( x_i \) are disregarded because their stress contributions are low. \( R_{\text{cut-off}} \) is chosen so large that the maximum possible collective stress of segments at this distance is lower than a certain stress value \( \delta \tau \). The choice of \( \delta \tau \) is detailed in Section 17.2.4. As an approximation for the maximum collective stress, the stress of a straight dislocation of infinite length, \( \tau_{\text{straight}} \), is used (e.g. Hirth and Lothe 1982). \( \tau_{\text{straight}} \) is a function of the shortest distance to this straight dislocation, for which \( R_{\text{cut-off}} \) is inserted.

\[
\delta \tau \geq \tau_{\text{straight}}(R_{\text{cut-off}}) = (1 - \nu)^{-1/2} \frac{G}{2\pi} \frac{b}{R_{\text{cut-off}}}
\] (17.7)

With the factor \((1-\nu)^{-1/2}\), an average of edge and screw dislocation character is used for this approximation. When resolved to \( R_{\text{cut-off}} \) Equation (17.7) yields an evidential reasonable cut-off distance to be used in Equation (17.4) as a function of \( \delta \tau \) (Section 17.2.4).

17.2.4 Simulation Procedure and Accuracy

As a configuration to start a simulation with, usually one (or more) straight dislocation near the bottom of a rectangular field of obstacles is used. The generation of such fields is detailed in Section 17.3. The dislocation glide is simulated until the first dislocation touches the upper side of the field. On the left and the right side, periodic boundary conditions are imposed: the
17.3 Particle Arrangement

arrangement of the obstacles (particles) is periodic, and also the dislocation’s self-interaction is periodically continued as if the simulated area was wrapped around a cylinder and the dislocations moved in the direction of the cylinder axis.

A simulation run is started with a low external stress $\tau_{\text{ext}}$. This drives the dislocation forwards against the obstacles; the dislocation bows out between them, as Equation (17.1a) is solved numerically, until the self-stress $\tau_{\text{disloc}}$ compensates for $\tau_{\text{ext}}$ in every PI. Hence a static equilibrium of resolved stresses is found (here $\tau_{\text{obs}} = 0$ is assumed between the obstacles). Then $\tau_{\text{ext}}$ is increased by a small step $\delta \tau$, which is chosen at about 3% of the expected (estimated) CRSS. Then the temporal integration is continued until the next static equilibrium is found, and so forth. The dislocation overcomes the obstacles by “shearing” or circumventing them and eventually reaches the upper side of the obstacle field. Then $\tau_{\text{ext}}$ is so high that the dislocation would glide continuously through this field. The CRSS, or $\tau_p$, is taken to equal the last value of $\tau_{\text{ext}}$ which yielded a stable equilibrium configuration, i.e. before the dislocation contacted the upper field boundary.

$$\tau_p = \text{Max} [\tau_{\text{ext}}] - \delta \tau.$$  \hspace{1cm} (17.8)

Due to this procedure, the accuracy of $\tau_{\text{ext}}$ and hence $\tau_p$ is limited by the stress step $\delta \tau$. Therefore none of the other stress contributions of the stress equilibrium, Equation (17.1a), needs to be more accurate than $\delta \tau$. This fact is used to avoid a waste of computation time. One example is the viscous drag stress $\tau_{\text{drag}}$, described in Section 17.2.1, which vanishes in a static equilibrium: the condition to recognize a static equilibrium, $|v| = 0$ in every PI (which would never be found numerically), is relaxed to $|v| < (b/B) \delta \tau$. Likewise, the accuracy limit $\delta \tau$ for $\tau_{\text{disloc}}$ defines the outer cut-off radius $R_{\text{cut-off}}$ via Equation (17.7). In principle, the maximum error of the simulations is several times $\delta \tau$ because several contributions in the stress equilibrium involve this error. But it has been verified that the typical error of $\tau_p$ for a given obstacle field is actually $\pm 6\%$ of $\tau_p$ for obstacle fields containing 500 to 1000 obstacles (Mohles and Fruhstorfer 2002).

Another simulation parameter that affects the simulation accuracy is the density of PIs along the dislocation (or, the segment lengths). About six PIs are needed over the length of one obstacle diameter in order to sample the stress of this obstacle with sufficient accuracy (Mohles 2001a). But as the dislocations move and bow out, the local density of PIs changes. Therefore in regular intervals PIs are inserted or deleted as required in order to keep the local distances $s_i$ between PIs close to the predefined mean value $\bar{s}$.

17.3 Particle Arrangement

The stress $\tau_{\text{obs},k}$ of each obstacle $k$ in Equation (17.1a) is a function of the location $x_k$ in space relative to PI $i$ for which $\tau_{\text{obs},k}$ is calculated: $\tau_{\text{obs},k}(x_i) = \tau_{\text{obs}}(x_k - x_i)$, where the function $\tau_{\text{obs}}(\Delta x)$ describes the type of interaction between particles and dislocations. This includes the particle shape. The functions $\tau_{\text{obs}}(\Delta x)$ for the strengthening mechanisms used for simulations so far are given in the respective sections of the simulation results (17.4.1 to 17.4.3). The generation of particle arrangements in space (sets of $x_k$) is described subsequently.
Since real particles do not intersect, the particle arrangement in space depends on the particle shape and on the distribution of sizes and possibly orientations. At present the particles are chosen to be spherical. Their radii $r_k$ are randomly picked from a distribution $g(r_k/\bar{r})$, where $\bar{r}$ denotes the mean particle radius. An example is the function $g_{WLS}(r/\bar{r})$ that has been derived analytically by Wagner (1961), Lifshitz and Slyozov (1961) (WLS distribution) for Ostwald ripened crystals with low particle volume fractions $c$. For $r/\bar{r} < 1.5$,

$$g_{WLS}(r/\bar{r}) = \frac{4}{9} (r/\bar{r})^2 \left( \frac{3}{3 + r/\bar{r}} \right)^{7/3} \left( \frac{1.5}{1.5 - r/\bar{r}} \right)^{11/3} \exp \left( \frac{r/\bar{r}}{1.5} - 1.5 \right),$$  

(17.9)

and for $r/\bar{r} \geq 1.5$, $g_{WLS} = 0$. Experimental works have shown that radius distributions of real Ostwald-ripened specimens can differ from this function quite strongly (e.g. Wagner and Kampmann 1991); but there is also good experimental verification of the function $g_{WLS}(r/\bar{r})$ (Fruhstorfer et al. 2002). However, virtually any distribution function $g(r/\bar{r})$ may be used for the following procedure to generate arrangements of spherical particles. This procedure is an improved version of that of Rönnpagel and coworkers (Fuchs and Rönnpagel 1993, Pretorius and Rönnpagel 1994). The idea is that every particle has depleted its surrounding of the solute needed for its growth. Therefore no other particle can exist in this surrounding, called sphere of interest in the following. The procedure consists of five steps.

(i) Spheres with radii $R_k$ picked randomly from $g_{WLS}(R_k/\bar{R})$ are densely packed in a cuboid volume. Periodic boundary conditions are applied: if part of a sphere penetrates a side of the cuboid, this part comes back in on the opposite side. Such a periodic close packing can be achieved in several ways. A simple and fast one has been described by Mohles and Fruhstorfer (2002). The volume fraction $c_0$ of the spheres amounts to about 0.59, i.e. the packing is fairly dense. These spheres are meant to represent the spheres of interest from which the particles get their material.

(ii) The aspired particle volume fraction $c$ is chosen. For this, one particle with radius $r_k = (c/c_0)^{1/3} R_k$ is placed into the center of each sphere of interest; hence the particle radii are also WLS-distributed. With respect to the nearest neighbor spacings, the resulting particle arrangement is rather uniform.

(iii) Each particle $k$ is shifted out of the center of its sphere of interest by a 3D random vector $u_k = \xi R_k I_k$. Here $I_k$ is a random vector each component of which is in the range $[-0.5 \ldots 0.5]$, equally distributed. $\xi$ is a parameter which allows to choose the degree of randomness. If particle $k$ is intersecting another one after shifting, $I_k$ is rejected and a different random vector is used instead. The periodic boundary conditions are maintained.

(iv) The mean particle radius $\bar{r}$ is adjusted by multiplying all particle positions and radii and the size of the cuboid by a common factor.

(v) For the simulations a glide plane is chosen parallel to one side of the cuboid. This ensures that periodic boundary conditions are maintained in the plane. The glide plane may as well be chosen before the scaling procedure in step (iv).
This procedure to generate particle arrangements may appear artificial, but an approach like this is necessary because other methods, like Monte-Carlo-simulations of particle growth (Binkele and Schmauder 2003) yield much too small obstacle arrays. To avoid effects of the boundary conditions and strong statistical scatter in \( \tau_p \), a number of 500 obstacles or more must be effective in the glide plane. A three-dimensional particle array must contain even more particles. With the method described above such an arrangement is generated quickly. It has been proved (Mohles and Fruhstorfer 2002) that most realistic arrangements are attained with the present method when a randomness parameter \( \xi = 1.0 \) is used in step (iii). For this proof the distributions of nearest neighbor spacings in real and computer generated arrangements have been compared quantitatively. Particle arrays with particle shapes other than spheres (e.g. cuboids) will require a more refined characterization.

17.4 Strengthening Mechanisms

In the present section three types of interaction between particles and dislocations are described. Typical resulting dislocation configurations are shown, and some quantitative results are given.

17.4.1 Dispersion Strengthening

In dispersion strengthened materials dislocation glide is impeded by incoherent particles. Dislocations cannot penetrate these particles because the glide system in the ductile matrix is not continued inside the particles. An example for this is a copper matrix containing amorphous \( \text{SiO}_2 \)-precipitates grown by internal oxidation. In the simulations such detail is unimportant; incoherent particles are modeled by a virtually infinite negative stress inside (outside the stress vanishes). To avoid numerical problems, a certain finite stress \( \tau_{\text{inc}} \) (incoherent) is used which always suffices to keep the dislocation out (Baither et al. 2001). Moreover, on the particle surface the stress is smoothed to avoid infinite stress gradients. The smoothing is done by a soft step function with which the function \( \tau_{\text{obs}}(\Delta x) \) rises almost linearly from 5\% to 95\% of \( \tau_{\text{inc}} \) over the mean distance \( \bar{s} \) between the PIs (Mohles 2001b).

In Figure 17.2, two consecutive equilibrium configurations of one dislocation are plotted as continuous lines. The dark gray discs represent the particle intersections with the glide plane. The areas shaded in light and darker gray have been swept out by the dislocation. The darker shade of gray indicates the area of the second before last equilibrium position, i.e. at \( \tau_{\text{ext}} = \tau_p - \delta \tau \). The light gray area has been swept out after \( \tau_{\text{ext}} \) has been increased to \( \tau_p \); the respective dislocation configuration is the last stable one. As a starting configuration a pure edge dislocation (dashed line) has been used; however the equilibrium configurations show all dislocation characters. As to be expected (de Wit and Koehler 1959) the edge parts of the dislocation bow out more strongly than the screw parts. Except for the few particles below the starting configuration, all particles behind the main dislocation are surrounded by an Orowan loop or an Orowan island. The latter means that several particles are enclosed in a common loop inside of which not only the particles but also the matrix has not been sheared. Simulated dislocation configurations for different volume fractions \( c \) and degrees of
randomness of the particle arrangement (parameter $\xi$ in Section 17.3) have been published by Mohles and Fruhstorfer (2002).

![Figure 17.2: Two consecutive equilibrium dislocation configurations in a simulated crystal strengthened by incoherent particles. The configurations partly overlap. Areas shaded in light and darker gray have been swept out by the dislocation. Mean radius $\bar{r} = 100\text{b}$, volume fraction $c = 0.1$.](image)

To derive quantitative results, simulations have been performed in wide ranges of mean particle radii $\bar{r}$ and volume fractions $c$; edge and screw dislocation characters have been used as starting configurations (Mohles 2001b). The randomness parameter $\xi$ has been varied as well (Mohles and Fruhstorfer 2002). The tabulated data has been compared with Nembach’s (1996) analytical formula (N stands for Nembach, disp stands for dispersion strengthening) which in turn had been based on works of Orowan (1948), Bacon et al. (1973) and Hirsch and Humphreys (1969). $\tau_{\text{p,disp}}^{N}\bar{r},c$ regards the elastic dislocation self-interaction, the appropriate averaging procedure for screw and edge dislocations and a random obstacle arrangement.
Moreover it takes a real particle radius distribution \( g(r/\bar{r}) \) into account.

\[
\tau_{p,\text{disp}}^N = 0.9 \frac{G}{4\pi \sqrt{1 - \nu}} \frac{2b}{\omega L} \frac{\ln(2\omega_H r/b)}{\ln(\omega_L r/b)}^{3/2} \left\{ \ln(2\omega_H r/b) \right\}^{1/2}
\]

with \( \omega_L = \sqrt{\pi \omega_H / c - 2\omega_H} \) and \( \omega_H^{-1} = \omega_L^{-1} + (2\omega_L)^{-1} \) \hspace{1cm} (17.10)

The constants \( \omega_L \) and \( \omega_H \) depend on \( g(r/\bar{r}) \); they establish the statistical relation between \( \bar{r} \) and the mean radius \( (= \omega_L \bar{r}) \) and the mean area \( (= \omega_H \pi (\bar{r})^2) \) of the particle intersections with the glide plane. For the distribution \( g_{\text{WLS}}(r/\bar{r}) \), \( \omega_L = 0.82 \) and \( \omega_H = 0.75 \) (Nembach 1996).

The simulated data approved the function \( \tau_{p,\text{disp}}^N(\bar{r}, c) \) pretty well, especially so with respect to the dependence on \( \bar{r} \). However, the simulated dependence on \( c \) came out to be slightly stronger than predicted. One reason is that all Orowan loops are stress sources which attract the main dislocation and hence hold it back. This raises \( \tau_p \), especially so when there are many loops, i.e. for large volume fractions \( c \). This synergetic effect has not been considered by \( \tau_{p,\text{disp}}^N(\bar{r}, c) \). But probably there is another, purely geometrical cause for the systematic deviations: Zhu and Starke (1999) have also found an additional \( c \)-dependence of \( \tau_{p,\text{disp}} \). In their simulations these authors used the line tension model, which does not allow for the attracting stress of the Orowan loops.

Anyway, the present simulations are more reliable than Equation (17.10) because they allow for possible synergetic effects automatically. Therefore a slightly altered version of Equation (17.10) has been suggested for \( \tau_{p,\text{disp}} \) (Mohles 2001b):

\[
\tau_{p,\text{disp}} = 0.89 \frac{G}{4\pi \sqrt{1 - \nu}} \frac{2b}{\omega L} \frac{\ln(4\omega_H r/b)}{\ln(\omega_L r/b)} \hspace{1cm} (17.11)
\]

\( \tau_{p,\text{disp}}(\bar{r}, c) \) differs from \( \tau_{p,\text{disp}}^N(\bar{r}, c) \) mostly by larger exponents of the rightmost fraction. This formulation has been found by trial and error. It has no physical justification, but it accounts for the additional \( c \)-dependence found very well.

### 17.4.2 Order Strengthening

An order strengthened material consists of a disordered matrix containing long-range ordered coherent precipitates (short-range ordering effects are disregarded here). When a perfect matrix dislocation (D1) cuts through such a particle it destroys the order inside and creates an antiphase boundary (APB) with the energy \( \gamma_{\text{APB}} \) per unit area. This energy expense holds the dislocation back. If a second dislocation (D2) follows and restores the order this dislocation will be driven forwards. Together, D1 and D2 form a superdislocation. Regardless of the type of order the obstacle stress \( \tau_{\text{obst},k} \) of particle \( k \) is mainly determined by \( \gamma_{\text{APB}} \):

\[
\tau_{\text{obst},k} = \begin{cases} 
-\gamma_{\text{APB}}/b_1 & \text{inside particle } k, \text{ D1} \\
+\gamma_{\text{APB}}/b_2 & \text{inside particle } k, \text{ D2} \\
0 & \text{outside} 
\end{cases} \hspace{1cm} (17.12)
\]
Dislocation Dynamics Simulations of Particle Strengthening

Here \( b_1 \) and \( b_2 \) are the lengths of the Burgers vectors of D1 and D2, respectively. If D1 and D2 are actually matrix dislocations, \( b_1 \) and \( b_2 \) are equal. But they are not equal if the matrix dislocations dissociate into partial and superpartial dislocations inside the particles, like in the Condat-Décamps-model (1987). However in the following the case \( b_1 = b_2 \) is considered. An example for such a system is the commercial Nickel-base superalloy Nimonic PE16. It consists of an f.c.c.-matrix (\( \gamma \)-phase) with a Burgers vector of the type \( a_0/2\langle 110 \rangle \) \((a_0 = \) lattice constant) and coherent L1\(_2\) long-range ordered spherical precipitates (\( \gamma' \)-phase, Burgers vector \( a_0(110) \)). However, for the computer simulations such detail is not required. Like for incoherent particles (Section 17.4.1), \( \tau_{\text{obst},k} \) of Equation (17.12) is smoothed on the particle surface over the distance \( \bar{s} \) to avoid infinite stress gradients.

In Figure 17.3(a) and (b), equilibrium configurations of four matrix dislocations (D1, D2, D3, D4), or two superdislocations, in a Nimonic PE16 crystal are plotted. In Equation (17.12) and in the following only D1 and D2 are mentioned; however the same statements hold for D3 and D4, respectively. The crystal is near the peak-aged aging state: \( \tau_p \) is near its maximum with respect to the mean particle radius \( \bar{r} \). The leading dislocation of each pair (D1) bows out strongly between the particles while the trailing one (D2) is rather straight. This is in line with observations by transmission electron microscopy in real PE16 crystals with similar parameters \( \bar{r} \) and \( c \) (e.g. Nembach et al. 1985, \( \bar{r} = 32b, c = 0.089 \)). D1 is pushed forward by \( \tau_{\text{ext}} \) and by the stress caused by D2, whereas for D2, \( \tau_{\text{ext}} \) and the stress caused by D1 mostly compensate for each other.

It appears that D1 determines the overall, larger-scale configuration of a pair because it touches many more particles than D2. The main effect of D2 is that it drives D1 against the obstacles in addition to \( \tau_{\text{ext}} \). This lowers the stress needed to overcome them, \( \tau_p \), by the factor two. Brown and Ham (1971) have pointed out that \( \tau_p \) is lowered further because D2 is pushed forwards not only by \( \tau_{\text{ext}} \) but also by \( \tau_{\text{obst}} \) inside the obstacles. Two such cases can be seen in the inset of Figure 17.3(b). But simulations of the present kind (Mohles 2003) have yielded that the latter effect is unimportant, in agreement with the assumption of Haasen and Labusch (1979).

But D2 has another implication: when it touches the same particle that holds back D1, it will be driven forwards and thereby drive D1 out of the particle. Then this particle is overcome. The importance of this process, called strong pair coupling (SPC), has already been pointed out by Hüther and Reppich (1978). The present simulations show that SPC only occurs in screw configurations. In Figure 17.3(a), showing dislocations of predominant screw character, there are about nine cases where D2 almost touches the same particle as D1; two are magnified in the insets. The edge dislocations in Figure 17.3(b) show only one such case; it is magnified in the inset. But in this case, the local dislocation line vectors have changed so strongly that D1 and D2 have mixed rather than edge character. Screw dislocations are stiffer than edge dislocations (de Wit and Koehler 1959). Therefore they are much more likely than edge dislocations to penetrate a particle; only then D2 has a chance to follow.

The impact of SPC on \( \tau_p \) is quite significant. In the example of Figure 17.3 (\( \bar{r} = 36b, c = 0.1, \gamma_{\text{APB}}/(Gb) = 0.015 \)), \( \tau_p/G \) for edge and screw dislocations equals 0.0021 and 0.0016, respectively. The difference is caused by SPC alone; this has been proved on the basis of more quantitative data (Mohles 2003). It is emphasized that SPC may become quite important if the particles are non-spherical: if large particles have sharp edges, SPC is likely to occur so that they are sheared in spite of their size; otherwise large particles are circumvented.
17.4 Strengthening Mechanisms

Figure 17.3: Equilibrium configurations of four matrix dislocations in an order strengthened crystal (Nimonic PE16) near the peak-aged aging state. \( r = 36b, c = 0.1, \gamma_{APB}/(Gb) = 0.015, \tau_{ext} = \tau_p. \) (a) screw dislocations, \( \tau_p = 0.0016G; \) (b) edge dislocations, \( \tau_p = 0.0021G. \)

Therefore the curvature radius of the particle edges will be an important parameter to be considered when such materials are modeled.

### 17.4.3 Lattice Mismatch Strengthening

Lattice mismatch strengthened materials contain over- or undersized particles which are coherently embedded in the matrix. The size mismatch causes a coherence stress \( \sigma_{coh}. \) The tensor \( \sigma_{coh} \) of misfitting ellipsoids in an elastically isotropic continuum has been calculated by Eshelby (1956). By applying Equation (17.3) (Peach and Koehler 1950), the corresponding
obstacle stress $\tau_{\text{obst}, k}$ of particle $k$ is derived from $\sigma_{\text{coh}}$. For a spherical particle $k$ with radius $r_k$ located at $x_k = (x_k, y_k, z_k)$, $\tau_{\text{obst}, k}$ at $x$ is

$$\tau_{\text{obst}, k} = \begin{cases} 6\varepsilon G r_k^3 (\Delta x_k \cdot b^0) |\Delta x_k|^{-5} & \text{if } |\Delta x_k| > r_k \\ 0 & \text{if } |\Delta x_k| < r_k \end{cases}$$  \hspace{1cm} (17.13)

where $\Delta x_k = x_k - x$ and $\varepsilon = (a_p - a_0)/a_0$ is the constrained lattice mismatch parameter ($a_0$ and $a_p$ are the relaxed matrix and particle lattice constants, respectively). Inside the particles ($|\Delta x_k| < r_k$) there is only hydrostatic stress; hence $\tau_{\text{obst}}$ vanishes there (Equation (17.3)). Outside $\tau_{\text{obst}}$ has a minimum and a maximum.

The gray background of Figure 17.4 is a grayscale plot of the obstacle stress of all particles (the sum in Equation (17.1a)). Three particles are marked by the letters A, B and C. Each particle shows an area of positive (darker gray) and negative (lighter gray) obstacle stress. The loops are stabilized by particles that either do (e.g. loop X in Figure 17.4) or don’t (e.g. loop Y) intersect the glide plane. The respective numbers of both of these cases are roughly the same. This demonstrates that for simulations of lattice mismatch strengthening, it is important to use a real three-dimensional particle arrangement. In the cases of dispersion and order strengthening (Sections 17.4.1 and 17.4.2) this is less important because in these cases the obstacle size distribution can be treated analytically (Nembach 1996). Dislocation configurations for other parameters $\vec{r}$, $c$ and $\varepsilon$ have been published elsewhere (Mohles 2001c, 2001d, Mohles and Nembach 2001, Mohles 2002).

For the CRSSs of peak-aged and overaged mismatch strengthened crystals, the following analytic expressions $\tau_{p,\text{mis,peak}}(\vec{r}, c)$ and $\tau_{p,\text{mis,over}}(\vec{r}, c, \varepsilon)$, respectively, have been derived from simulations so far (Mohles 2001c, Mohles and Nembach 2001). They are quoted here to allow for a comparison with measurements on real specimens.

$$\tau_{p,\text{mis,peak}} = 1.29\varepsilon G \sqrt{\frac{\pi a_p}{\bar{\lambda}_l}}$$  \hspace{1cm} (17.14)

$$\tau_{p,\text{mis,over}} = 1.83(1 - 1.5c)\varepsilon^{0.5}(\bar{r}/b)^{0.4} r_p^{\text{disp}}(\vec{r}, c)$$  \hspace{1cm} (17.15)

In the case of Equation (17.14), only the numerical factor 1.29 for the general theory of peak-aged crystals (e.g. Nembach 1996) has been derived. Equation (17.15) has been based on Nembach’s expression $\tau_{p,\text{disp}}^N$ for dispersion strengthening (Equation (17.10)). Equation (17.15) differs from $\tau_{p,\text{disp}}^N$ mainly by the factor $\varepsilon^{0.5}(\bar{r}/b)^{0.4}$. Like in the case of dispersion strengthening this factor has been deduced by trial and error. Still there is a rather simple explanation for it (Mohles 2001c): The number of strong obstacles, e.g. those which hold Orowan loops in Figure 17.4, scales with the maximum force of all obstacles, which is proportional to $\varepsilon$ and $\bar{r}$. $\tau_{p}$, in turn, is basically proportional to the square root of the number of obstacles; this would explain a factor $\varepsilon^{0.5}(\bar{r}/b)^{0.5}$. For underaged crystals no analytic expression is known yet because the results are quite complicated (see below).
Figure 17.4: Two consecutive equilibrium configurations in a slightly overaged lattice mismatch strengthened crystal. The gray background is a grayscale plot of the coherence stress as seen by the dislocation. Dark gray: $\tau_{obst} > 0$; light gray: $\tau_{obst} < 0$. Particles A/B are located under / above the glide plane; particle C intersects the glide plane. $\bar{r} = 100b$, $c = 0.02$, $\varepsilon = 0.02$. 

In Figure 17.5 the measured CRSS of copper single crystals strengthened by cobalt-rich particles, (Martin 1979, Büttner et al. 1987) is plotted as a function of $(\bar{r}/b)^{1/2}$. The strengthening mechanism and the particle distribution of this alloy are known to be the same as in the present simulations; hence the simulation results can be compared with these measurements directly. The accuracy of the measurements roughly equals the size of their symbols in Figure 17.5. The measured CRSSs contain a small contribution (at most 5%) of solid solution strengthening, but this is disregarded here. Variations of the deformation temperature are also disregarded. The functions $\tau_{p,\text{mis,peak}}(c, \varepsilon)$ and $\tau_{p,\text{mis,over}}(\bar{r}, c, \varepsilon)$ with $c = 0.02$ (as measured) and $\varepsilon = 0.015$ (Büttner et al. 1987) are plotted for comparison. The material constants $G = 42.1$ GPa, $\nu = 0.43$ and $b = 0.256$ nm are used. To some degree this choice for $G$ and $\nu$
accounts for the elastic anisotropy of copper (Bacon et al. 1973).

\[ \tau_p / \text{MPa} = \frac{1}{(\bar{r} / b)^{1/2}} \]

Figure 17.5: CRSS of a Cu-matrix strengthened by Co-particles, \(\epsilon = 0.020 \pm 0.001\). Measured: ■ Martin (1979), • Büttner et al. (1987). Simulation results with material parameters taken from literature and \(\epsilon = 0.02\): —— Equation (17.15), —— Equation (17.14), —— Equation (17.11). Single simulations of a pair of Shockley partial dislocations: + edge, ○ screw character.

In the range \(6 \leq (\bar{r} / b)^{1/2} \leq 10\) the agreement between simulated and experimental results is downright excellent. It is emphasized that no parameter at all has been adjusted. For \((\bar{r} / b)^{1/2} > 10\) the measured CRSS systematically deviates from the prediction of Equation (17.14). There are several possible explanations for this:

(i) The particles cause stresses not only in the primary glide plane; this may support cross-slip. In the present simulations this has not been allowed for.

(ii) Prismatic dislocation loops may form around large particles during growth or deformation which compensate for the elastic coherence stress. The absolute constrained mismatch \(\Delta d\) of a particle with radius \(r\) over its diameter is \(\Delta d = 2\epsilon r\). For \(r = 100b\), \(\Delta d\) equals \(3b\) so that a single prismatic dislocation loop would already cause a significant stress reduction. This would also lower \(\tau_p\).

(iii) The Co-particles may change their lattice from f.c.c. (as enforced by the Cu-matrix) to h.c.p. (the normal lattice of Co at room temperature) when they are large. Then the precipitates become semi coherent or incoherent. The simulation result of incoherent particles (Equation (17.11)) with \(\epsilon = 0.02\) has been included in Figure 17.5. The agreement for \((\bar{r} / b)^{1/2} > 11.5\) suggests that the large particles may indeed be incoherent.
In the underaged aging state of the crystal, $(\bar{r}/b)^{1/2} < 5$ in Figure 17.5, a linear increase of $\tau_p$ with $(\bar{r}/b)^{1/2}$ is expected from Friedel’s theory (1956). In this theory, point obstacles of equal strength and the line tension approach have been used. Computer simulations of Foreman and Makin (1966, 1967) have confirmed $\tau_p \sim (\bar{r}/b)^{1/2}$ for weak obstacles; the deviation for strong obstacles is not important here. The CRSSs measured by Martin (1979) and others (see Nembach 1996), indeed, yield $\tau_p \sim (\bar{r}/b)^{1/2}$ (Figure 17.5). This agreement looks compelling; however the present simulations indicate that the agreement is rather accidental. When the same suppositions (point obstacles, line tension) are used in the present computer code, the results of Foreman and Makin (1966, 1967) are recovered accurately. This has been verified thoroughly. But with the models described in Sections 17.2 and 17.3 the simulations involve much more detail about lattice mismatch strengthening (distributions of obstacle sizes and strengths, dislocation self-interaction, dislocation dissociation) and hence yield more complicated results:

Neither edge nor screw dislocations show the proportionality $\tau_p \sim (\bar{r}/b)^{1/2}$. For the CRSS of the edge dislocation, $\tau_p \sim (\bar{r}/b)^{1/3}$ is a slightly better description. This proportionality has been deduced by Mott and Nabarro (1948) and Schwarz and Labusch (1978) for extended obstacles (in contrast to point obstacles). But a good description of the results of Figure 17.5 and other simulated data (Mohles 2001d, 2002) is only found with $\tau_p \sim (\bar{r}/b)^{1/8}$. Even more surprising, at first sight, is the CRSS of screw dislocations: from $(\bar{r}/b)^{1/2} = 2$ to $(\bar{r}/b)^{1/2} = 3$, screw dislocations show a decrease of $\tau_p$ in Figure 17.5. It has been deduced (Mohles 2001d) that this behavior emerges from the dislocation dissociation: with $(\bar{r}/b)^{1/2} \approx 2$, the particle size is close to the mean dissociation width which equals about $5b$ for a screw dislocation with the material constants quoted above.

While the simulation results for underaged crystals are ambiguous, the measured CRSS is always in between. This supports the view that the unique CRSS of a real crystal is given by a mean value of the results for edge and screw dislocations. First simulations of expanding dislocation loops (Mohles and Nembach 1999) support this view, too. For overaged and peak-aged crystals, the geometric mean is to be used (Hirsch and Humphreys 1969, Mohles and Nembach 2001). When the geometric mean of the simulation results in Figure 17.5 is compared with the measured CRSS, a maximum deviation of 20% is found. In contrast, Nembach’s (1996) comparisons have shown that the predictions of hitherto theories overestimate the CRSS by a factor two to three.
17.5 Summary and Outlook

Computer simulations of dislocation glide are useful to improve the understanding of strengthening mechanisms and to make accurate predictions. Examples have been given for dispersion strengthening, order strengthening and lattice mismatch strengthening. In some cases it is possible to derive analytical expressions (e.g. Equations (17.11), (17.14), (17.15)) for the CRSS. For this it is useful to have a theoretically based expression, like $\tau_{N,\text{disp}}(\bar{r}, c)$ of Equation (17.10) (Nembach 1996): only rather simple alterations of this expression were necessary to find an accurate description of the simulation results for dispersion strengthened and for overaged lattice mismatch strengthened materials. In other cases, like for underaged lattice mismatched strengthened materials, the simulation results are quite complicated because they involve many details of the respective strengthening mechanism: the mean particle radius $\bar{r}$, the volume fraction $c$, the mismatch parameter $\varepsilon$ and the stacking fault energy must be considered as parameters. Moreover the distributions of particle sizes and strengths and the Poisson ratio $\nu$ or even elastic anisotropy will have some impact. This makes it quite difficult to find a general analytical description for the simulated data. Still, for a material with concrete parameters the present simulations produce useful predictions for the CRSS. This has been demonstrated for a Cu-rich CuCo alloy, for which the strengthening mechanism (lattice mismatch) and the parameters are well-known. Conversely, if measurements and simulations strongly contradict each other then a different or additional mechanism must be effective in the real specimen. An example is the strongly overaged state of the CuCo-alloy in Section 17.4.3.

So far only spherical particles have been considered to keep the number of parameters low and to treat the main parameters first. The simulations can also handle other particle shapes; however this will require a new approach to generate an appropriate particle array, and this in turn requires an accurate characterization of the geometry. For cuboid particles like in many Ni-based $\gamma/\gamma'$ superalloys, for instance, the channel widths between particles and their distribution are important parameters. But also the curvature radius of the edges can become important (Section 17.4.2). Moreover, other strengthening mechanisms (e.g. modulus mismatch, stacking fault energy mismatch) and also their mutual superpositions can be considered. Hence the strengthening effect of particles in real materials can be investigated in great detail.

Acknowledgements

For fruitful discussions and continued support I would like to thank Prof. Dr. E. Nembach. Financial support by the Deutsche Forschungsgemeinschaft is gratefully acknowledged.

References

References


18 Discrete Dislocation Dynamics Simulation of Thin Film Plasticity

Burghard von Blanckenhagen and Peter Gumbsch

Here it is illustrated how discrete dislocation dynamics (DDD) simulations can be used to investigate various aspects of the plastic deformation of thin films. The description of polycrystalline fcc metal films which can be found as interconnects in integrated circuits is the main subject. In the first section, the topic of thin film deformation is introduced via experimental results and basic thin film models. Furthermore, it is outlined why thin films are well suited for DDD simulations. The second section considers aspects of the DDD simulation specific to thin films. In the last section of this chapter results are presented and the thin film specific deformation mechanisms are discussed.

18.1 Thin Film Plasticity

The increasing miniaturization in micro-electronics and micro-mechanics has led to a drastic reduction of the dimensions of the components in these devices. Interconnects in integrated circuits or tiny “structural parts” in micro-electro-mechanical (MEMS) devices made from fcc metals (e.g., Cu, Al, Ag or Au) often have thicknesses of the order 1 µm or well below. Due to the confinement in (at least) one dimension, the mechanical properties of these films differ dramatically from their bulk counterparts. Understanding the mechanical properties of thin metal films is therefore crucial in predicting the reliability of such small scale devices.

Measuring the mechanical properties of thin films requires dedicated experimental techniques like micro-tensile tests (e.g., Hommel et al., 1999), wafer curvature experiments (Dorner et al., 1986; Venkatraman and Bravman, 1992), bending experiments (Nix, 1989), bulge test (Kalkman et al., 1999) and nano-indentation (Nix, 1989). These experiments have consistently shown that the flow stress of thin films is much higher than for the corresponding bulk material and increases with decreasing film thickness irrespective of the metal or alloy under investigation. (Pb (Murakami, 1979), Al (Venkatraman and Bravman, 1992), Al-0.5%Cu (Venkatraman et al., 1990), Cu (Keller et al., 1998), Ag (Kobrinsky and Thompson, 1998)).

To comprehend the plastic deformation of the polycrystalline thin metal films used in micro-electronics and micro-mechanics, it is necessary to study the effect of the confined geometry on dislocation multiplication and motion. Grain size and film thickness of these films are usually of the order of µm and below and therefore smaller than the characteristic length scale of dislocation networks observed in bulk material after deformation. Indeed, dislocation cell structures have not been observed in thin films (Keller et al., 1996). Apparently, the collective behavior of many interacting dislocations, which would lead to the formation
of cell structures, is impeded by grain boundaries and interfaces. Therefore, the number of dislocations, which have to be considered in order to understand thin film plasticity, is rather small.

A model often used to describe thin film deformation considers the motion of dislocations in a thin channel (Freund, 1987) and (Nix, 1989) which is based on Matthews et al. (1970) model of misfit dislocations in epitaxial semiconductor films. Since the films are usually deposited on a substrate, it is assumed that dislocations moving through the film (threading dislocations) cannot leave at the substrate–film interface. As a consequence, interface dislocations are deposited by the threading dislocations (see Figure 18.1).

![Figure 18.1: Dislocation moving in a thin film. The interfaces to the top and bottom are impenetrable. Interface dislocations are deposited while the leading part of the dislocation (threading segment) advances.](image)

Balancing the energy necessary to produce the interface dislocations with the work done by moving the threading dislocation in the applied stress field, a thickness dependent flow stress is obtained. Neglecting a weak logarithmic dependence the flow stress is proportional to \(1/h\). The same reasoning for dislocations deposited at the grain boundaries leads to an additional critical stress which is proportional to the inverse grain size (Thompson, 1993).

These models qualitatively explain the dependence of the flow stress on film thickness, but the predicted stresses are roughly four times smaller than the experimentally measured ones (Murakami, 1979; Venkatraman and Bravman, 1992). This is displayed in Figure 18.2, where the stresses from wafer curvature experiments at room temperature (Vinci et al., 1995; Keller et al., 1998; Weiss, 2000) and the stresses at 0.5% plastic strain from micro-tensile tests (Hommel and Kraft, 2001) are shown together with the predictions from the Nix–Freund model (see von Blanckenhagen et al., 2003). The measured stresses in wafer-curvature experiments and micro-tensile tests are similar and much larger than the theoretical predictions. Since corrections to the Nix–Freund model due to different elastic constants or anisotropy are
18.2 Simulation of Dislocations in Thin Films

In order to identify a model system which captures the essential features of a thin film and can be used to study the plastic deformation, the typical microstructure of the films has to be analyzed. In many cases thin films are deposited on a substrate (e.g., silicon). Sometimes an oxidation layer forms on the top of the films or a passivation layer (cap layer) is deposited. The grain size of polycrystalline films after annealing is usually of the same order as the film thickness (Mullins, 1958). The grains are columnar with boundaries perpendicular to the interfaces. Most of the films have a pronounced \(\langle 111\rangle\) texture (e.g., Venkatraman et al., 1991).

To investigate dislocations in configurations similar to the Nix–Freund model, the simulation of the deformation of a single grain is an appropriate starting point. Since only a few expected to give only small contributions, other hardening mechanisms have to be identified which explain the high flow stresses in Figure 18.2 Candidates are the interaction of moving dislocations with each other and with other immobile interface dislocations or the nucleation and multiplication of the dislocations themselves. While analytical treatment is only possible for special cases, these scenarios can easily be probed with dislocation dynamics simulations.

**Figure 18.2:** Flow stresses of thin copper films on silicon substrate at room temperature determined in wafer curvature experiments (Vinci et al., 1995; Keller et al., 1998; Weiss, 2000) and stresses at 0.5% plastic strain of Cu films on polyimide substrate determined in micro-tensile tests (Hommel and Kraft, 2001) together with the prediction from the Nix–Freund model for a capped Cu film on a substrate. The minimum stress for a two-fold activation of a source (source model, see Section 18.3.2) is also shown. The stresses are plotted versus the reciprocal value of the smaller of film thickness or grain size.
dislocations are investigated, the correct treatment of the dislocation self-interaction (which causes the dislocation line-tension) is essential. Excellent agreement with analytical solutions can be obtained by approximating the dislocation line between the node of stress calculation and its neighboring nodes by a circular arc and calculating the self-stress of the arc (Bacon et al., 1973; Schwarz, 1999) using the averaging procedure proposed by Brown (1967). The results are insensitive to the number of nodes used to represent the dislocations. If the distance between the nodes becomes very small (on the order of $1b$) an averaging of the stress due to the next nearest neighbors similar to the Brown procedure is required (Schwarz, 1999). In contrast to methods which do not replace the segments adjacent to the point of self-stress calculation with a smooth line-segment (e. g. (Zbib et al., 1998; Gomez-Garcia et al., 1999)) an adjustment of the discretization parameters to the length scale of the problem under investigation is not required here.

An adequate treatment of the formation of dislocation locks and intersection events is necessary to allow for the short range interactions between dislocations in a thin film (Devincre et al., 2001; Rhee et al., 1998; Wickham et al., 1999; Weygand et al., 2002; von Blanckenhagen et al., 2002). Comparisons of DDD simulations and atomistic simulations of the formation and destruction of dislocation locks have shown good agreement, indicating that a treatment based on continuum theory is appropriate (Bulatov et al., 1998; Rodney and Phillips, 1999; Shenoy et al., 2000).

18.2.1 Boundary Conditions

An important aspect in the simulation of thin film deformation is the interaction of dislocations with interfaces and grain boundaries. The Nix–Freund model considers dislocations which are deposited at the interfaces by the moving threading segment. This situation can easily be implemented in a DDD simulation by making the interfaces impenetrable and allowing for dislocation motion only within the film (Figure 18.1). If the elastic constants of the substrate and cap layer are different from the film, the dislocations will experience image forces. Using an expression for the critical stress derived by Nix (1989) the effect of the image forces on the flow stress can be estimated. An increase of only 10% arises if the appropriate elastic constants for a silicon substrate and a Si$_3$Ni$_4$ passivation layer on a copper film are used. If substrate and passivation layer are stiffer than the film, the interface dislocations tend to stay away from the interfaces. However, the stand-off distance is small compared to the film thickness (Embury and Hirth, 1994). In the limit of infinitely stiff substrate and passivation layer, the critical stress is only 20% larger than for the same elastic constants (Nix, 1998). The estimated relative increase of the flow stress is independent of the film thickness.

If passivation and substrate are elastically more compliant, the stress field of the interface dislocations will be weaker. In the extreme case of a free standing film, dislocations can leave the film at the free surfaces. Two aspects are important when considering free surfaces in a DDD simulation: (1) The image forces of the free surfaces have to be taken into account and (2) the termination of dislocations at the interface has to be implemented.

Dislocations in a semi-infinite crystal with one free surface can be simulated by employing the principle of superposition of the elastic fields: The stress field is calculated for the infinite body and point forces are introduced at the free surface in order to compensate for the tractions on the surface and to yield the correct boundary conditions. The stress field in the body
due to the point forces can be calculated analytically and contributes to the forces exerted on the dislocations (Fivel et al., 1996). Free standing films with two free surfaces can also be simulated using this method, if point forces are applied on both surfaces and the interaction between the surfaces is taken into account (Fivel and Canova, 1999; Hartmaier et al., 1999). More complex boundary conditions can be realized if the stresses which are required to fulfill the boundary conditions are calculated using the finite element method. Displacement boundary conditions and interfaces to materials with different elastic constants can then be realized (Fivel et al., 1997; Lemarchand et al., 2001; Weygand et al., 2002).

Dislocations which touch the surface leave the crystal. In the simulation they have to be terminated at the surface. The terminating node has to be moved artificially. Schwarz (1999) proposed a method where the orientation of the segment which ends at the surface is chosen in order minimize the forces acting on it. A screw and an edge segment would have a right angle to the boundary since the dislocation is pulled to the surface. A segment with mixed character would have a certain inclination due to the tendency to align in screw direction (Lothe, 1970). Comparative simulations with the point force method have shown, that in cases where the line-tension dominates (e.g., activation of a Frank–Read source) the error due to the neglect of the image forces is smaller than the inherent error of a dislocation simulation, if this termination is used (Liu et al., 2001).

Considering the enormous additional computational cost of an explicit treatment of the image force problem (increase in computational time by about two orders of magnitude for coupled free surfaces (von Blanckenhagen, 2002)) and the relatively small corrections compared with the difference between the Nix–Freund model and the experimental data (Figure 18.2), it seems to be justified to use approximative schemes like impenetrable interfaces or terminated dislocations when trying to identify the underlying mechanisms of thin film deformation. The crucial and still open question is whether the interfaces should be treated as impenetrable obstacles for dislocations, where interface dislocations form, or whether dislocations should disappear at the interfaces. Transmission electron microscopy (TEM) studies of dislocations in thin metal films do not show consistent results. The formation of interface dislocations has been observed in some cases (Jawarani et al., 1997; Keller–Flaig et al., 1999). The disappearance of these interface dislocations after a short time of radiation with the electron beam was also often reported (Kuan and Murakami, 1982; Venkatraman et al., 1991; Müllner and Arzt, 1998). In contrast, in situ TEM studies of cross-sectional specimen did not show interface dislocations. The interface rather seemed to attract the dislocations almost like a free surface (Legros et al., 2001; Dehm and Arzt, 2000). The formation of interface dislocations also depends on the nature of the interface and the materials on both sides.

A DDD simulation can not be used to investigate the fundamental interaction of a dislocation with an interface, but the consequences of certain models for the dislocation–interface interaction with respect to thin film plasticity can be investigated. In particular simulations can be performed for impenetrable interfaces or free surfaces as the two extremes and then be compared with the experimental data.
18.3 Thin Film Deformation, Models and Simulation

18.3.1 Mobility Controlled Deformation

The task of the simulation is to identify a mechanism which can explain the high flow stress and its dependence on film thickness of thin films. The starting point is to investigate additional contributions to the Nix–Freund model. Freund (1990) and Nix (1998) analytically studied the interaction of moving dislocations with interface dislocations on crossing slip planes (Figure 18.3). The analytical estimates considered a simplified geometry of straight dislocations and predicted a strength enhancement of $\sim 2.6\gamma_0$ where $\gamma_0$ is the critical stress of the threading dislocation alone (solid line in Figure 18.2). In contrast, DDD simulations have shown that the stress increase for the overcoming of interface dislocations only amounts to 30% (Schwarz and Tersoff, 1996; Gomez-Garcia et al., 1999; Pant et al., 2001). The analytical estimates overestimated the stress increase considerably since they did not account for the flexibility of the dislocation which significantly helps to surmount the obstacles and reduces the critical stress tremendously. Hence, interface dislocations on intersecting slip planes, which act as obstacles for moving dislocations can not explain the flow stresses of thin films observed experimentally (Figure 18.2).

If only one slip system is activated in the film, interface dislocations on slip planes parallel to the plane of the moving dislocation will form and their stress fields will influence the advancement of the moving dislocation (Figure 18.4). Depending on the character and distance of the deposited interface dislocations, the motion of a threading dislocation can be hindered or aided (Anderson and Kreidler, jr., 1998). In (111) oriented films, the stress field is always repulsive for screw dislocations and also repulsive for edge dislocations if the distance between the interface dislocation and the moving dislocation is smaller than half the film thickness (von Blanckenhagen, 2002). The stress needed to move a dislocation in the field of interface dislocations will increase considerably if many interface dislocations exist to
both sides of the threading dislocation (Weihnacht and Brückner, 2001). However, in order to reach the experimentally measured high flow stresses, the distance between the parallel dislocations has to be much smaller than the film thickness. This leads to a flow stress which only depends on the distance of the parallel dislocations and does not depend on the film thickness. This is in contradiction to experiments (von Blanckenhagen, 2002).

The interaction of two threading dislocations on different parallel glide planes was investigated with a DDD simulation by Pant et al. (2001). If the distance between the threading dislocations is small, they can form dipoles and immobilize. However, to sustain the high stresses in the films their distance has to be approximately 20nm. The chance of such a close encounter in a film where grain size and film thickness are of the order of 1 µm and a typical dislocation density corresponds to a maximum of 100 dislocations per grain is far too small to regard it as the controlling mechanism of thin film plasticity. (von Blanckenhagen, 2002).

The models and simulations mentioned above consider preexisting dislocations in the film and assume that their motion is the most difficult step in the deformation process. The discussion and the cited literature has shown that this assumption does not provide a coherent picture of thin film deformation. Consequently, the nucleation and multiplication of dislocations from dislocation sources is considered next.

18.3.2 Source Controlled Deformation

As mentioned in Section 18.1 dislocation networks and dense dislocation entanglements have not been observed in thin films. Consequently dislocation sources or multiplication sites inside the grain (Frank–Read sources) must be rare and their number is expected to decreases with decreasing grain dimensions. Eventually, dislocation sources have to be multiply operated to
achieve the imposed plastic strain. If this is the case, the geometrical constrains of the thin film have a profound influence on the plastic behavior as will be outlined in the following.

Figure 18.5: Critical dislocation configurations in the activation of dislocation sources of size $s$ in a quadratical grain. (a) activation of the Frank–Read source, (b) the passage between the pinning points and the interfaces is as difficult as the activation of the source if $s = h/3$, (c) passage is the most difficult step if $s > h/3$, (d) after multiple operation of the source the critical configuration is at the passage between the pinning point and the deposited dislocations (von Blanckenhagen et al., 2003). In order to find a lower bound of the source activation stress a source with edge character (horizontal Burgers vector) was chosen for the simulation.

Results of the simulation of the activation of a Frank–Read source inside a grain with impenetrable interfaces and boundaries are displayed in Figures 18.5 and 18.6 (von Blanckenhagen et al., 2003). The starting configuration consists of a pinned dislocation segment in the center of the grain (Figure 18.5(a)). The side arms of the source, which connect the segment with the boundaries, are neglected since their stress contribution is small (Foreman, 1967). The applied stress is gradually increased and the dislocations move accordingly. The source activation stress, defined as the stress necessary to produce a dislocation loop and to transfer the pinned segment back to its initial configuration, is plotted in Figure 18.6 versus the size of the source. Figure 18.6 reveals that sources with an effective size of approximately one third of the film thickness are activated most easily and that a minimum stress is required that is at least three times larger than predicted by the Nix–Freund model. As long as the grain size is of the same order as the film thickness this result also applies for free standing films where only the grain boundaries act as obstacles for the dislocations (von Blanckenhagen et al., 2001).

It is instructive to compare this result with experimental ones. This is done in Figure 18.2 where the stress to activate a source of size $s_{\text{eff}} = d/4$ (the most efficient size to produce two dislocation loops) is plotted together with experimental data. The stresses predicted by
Figure 18.6: Source activation stress in units of the stress to move the dislocation through the film versus size of the source normalized by the film thickness, for the first to the sixth activation. The letters refer to the configurations shown in Figure 18.5 (von Blanckenhagen et al., 2003).

this source activation model match the experimental data very well both in magnitude and dependence on film thickness.

Transmission electron microscopy studies further support the source model: Dislocation loops were observed to emerge in the grain interior and to move outwards during the deformation of copper films on silicon substrate in cross section (Dehm and Arzt, 2000; Dehm et al., 2001). Furthermore, the sudden movement of dislocations over long distances was observed (Dehm, 2002), which is an indication for strong obstacles and well compatible with the source model.

The symmetric dislocation configuration in Figure 18.5 was chosen to make the simulation and interpretation easier and to establish a lower bound for the flow stress. More realistic arrangements can be realized by randomly distributing randomly sized sources on all slip systems. These three dimensional DDD simulations also consider interactions between dislocations from different sources. It has been shown that the source activation and not the breaking of dislocation locks is the controlling step as long as sources are multiply activated (von Blanckenhagen et al., 2002). The simulation can be used to calculated stress strain curves or hardening rates and their dependence on grain dimensions and orientation. A free parameter is the number of dislocation sources which are placed into the grain at the beginning of the simulation. An initial dislocation density of \( \rho_{\text{ini}} = 0.46 \cdot 10^{14} \) which is comparable to experimentally observed dislocation densities (Keller–Flag et al., 1999; Kobrinsky and Thompson, 2000) (and corresponds to 40 dislocation sources for the case shown in Figure 18.7), was chosen for the calculations which are presented in the following. Since simulations with different
Figure 18.7: The starting configuration consists of 40 randomly distributed dislocation sources of random sizes with slip systems which are randomly selected from the primary fcc slip systems \(\langle 110\rangle\{111\}\) (left figure). The sources are placed in a columnar grain with an edge length of 512 nm in a free standing film. The grain is viewed along the film normal which is the \(\langle 111\rangle\) crystallographic axis. The right figure shows the dislocation configuration after 0.58\% plastic strain.

random dislocation-starting-configurations will result in slightly different stress-strain curves, the stress-strain curves are averaged over several runs with the same dislocation density, in order to minimize the statistical error.

As an example of a three dimensional simulation the dislocation structure of a columnar grain in a free standing film (where the dislocations can leave the film at the surfaces) after 0.58\% plastic strain is shown in Figure 18.7. The initial dislocation configuration is shown in the left part of the figure. Different sources of different slip systems have been active and produced pile-ups at the grain boundaries. The dislocation density at the grain boundaries is higher than in the grain interior after the film has been deformed. Similar dislocation distributions have been found in TEM studies of polycrystalline thin copper films (Hommel and Kraft, 2001).

Figure 18.8 displays a simulated stress-strain curve together with a curve gained in a micro-tensile experiment for a copper film deposited on a polyimide substrate (Brederek and Kraft, 2001), (for experimental details see Hommel et al., 1999). The initial plastic deformation is well described, but the simulation does not capture the plastic deformation at large plastic strains. Whereas in the micro-tensile test the work hardening decreases at plastic strains larger than 0.3\% and shows a stress plateau for plastic strains exceeding 0.5\%, the hardening rate deduced from the simulation stays approximately constant (Figure 18.8), since no relaxation mechanisms are implemented in the simulation. The simulated stresses therefore have to be considered as upper bounds. During unloading the experiment and the simulation show an initial elastic region followed by plastic deformation which already starts while the sample is still under tension. This pronounced Bauschinger type behavior, which is often observed in thin films (Kretschmann et al., 1997; Baker et al., 1998; Shen et al., 1998), is naturally explained by the source model: When the stress is reduced, the dislocations in the pile-ups
move backwards due to their mutual repulsion and annihilate. However, this backwards motion and annihilation does not start immediately at the beginning of the unloading, since the stress to activate the sources is higher than the stress required to push the dislocations against the boundaries. Therefore a (small) elastic region can be observed in the unloading part of the stress strain curves.

To estimate the effect of the stress-contributions from dislocations in neighboring grains simulations have been done where several grains are simulated. Figure 18.10 compares the stress-strain curve of an isolated grain with that of an assembly of columnar grains in a free standing film, where the cubic central grain is surrounded by eight grains (the nearest neighbors to the sides and the next nearest neighbors at the corners) of the same dimensions. In each grain 40 random sized dislocations were randomly placed. The plastic strain was calculated from the dislocation motion in the central grain. The hardening rate of the grain with neighbors is smaller than that of the isolated grain. Here, the stress fields from dislocations piling up at a grain boundaries from both sides can compensate each other. Furthermore, the stress from dislocations in neighboring grains in addition to the applied stress aids to activate dislocation sources as can be seen by a comparison of Figure 18.7 with 18.9. Nevertheless, the compensation is not perfect and the stress continues to increase with increasing plastic strain. In particular, a comparison of similar simulations of grains in a capped film with impenetrable interfaces did not show significant differences between the curves of an isolated grain and a conglomeration of grains (von Blanckenhagen, 2002).
Figure 18.9: Dislocation configuration of a grain surrounded by eight grains after 0.58% plastic strain. The initial dislocation configuration in the central grain was the same as in Figure 18.7, the dislocation sources in the neighboring grains are randomly placed.

Figure 18.10: Stress vs. plastic strain for a grain surrounded by eight grains of the same dimensions ($d = h = 512$ nm, see Figure 18.9). All grains have a $\langle 111 \rangle$ orientation, the rotation around the normal axis is chosen randomly. The plastic strain was calculated only from the central grain. The stress was applied homogeneously to all grains. The reference curve was calculated for a single grain with the same source configuration (see Figure 18.7).
To conclude, even if the elastic interaction between different grains is taken into account, the stress plateau observed in micro-tensile tests at large strains (Figure 18.8) cannot be reproduced. This discrepancy between experiment and simulation indicates where further development of the DDD simulation is required. In order to simulate thin film deformation more realistically, the condition of impenetrable grain boundaries has to be modified. The absorption of dislocations at the head of a pile-up by the grain boundary or their transmission to the neighboring grain in analogy to the Hall–Petch model has to be included in order study the plastic interaction in addition to the elastic interactions between different grains. Another point concerns the implementation of temperature dependent effects: Relaxation experiments of copper films have shown that at room temperature roughly 80% of the stress is athermal (Weihnacht and Brückner, 2001), nevertheless mechanisms like cross slip, dislocation climb and thermally activated dislocation glide have to be implemented to study the deformation at higher temperatures. Further work is also required to investigate the formation of dislocation sources. The latter point and the interaction of dislocations with grain boundaries or interfaces cannot be solved within a DDD simulation, here atomistic simulations are needed. Using the results of atomistic calculations as parameters in the DDD simulations allows to connect the microscopic with the macroscopic scale.

To summarize, DDD simulations are a useful tool to investigate dislocation dynamics in thin films and help to understand the deformation mechanism. The simulation of stress–strain curves, which could be used as an input to macroscopic simulations, is possible for the initial deformation, but further work is still required to reach larger strains and more complex loading histories.

Acknowledgments

This work was financially supported by the Deutsche Forschungsgemeinschaft (Gu 367/18).

References


J. Lothe. The image force on dislocations at free surfaces — comments on the concept of line tension. In J. A. Simmons, R. de Wit, and R. Bullough, editors, *Fundamental Aspects of Dislocation Theory*, volume 317, pages
References


19 Discrete Dislocation Dynamics Simulation of Crack-Tip Plasticity

Alexander Hartmaier and Peter Gumbsch

The fracture toughness of transition metals or semiconductor crystals strongly depends on loading rate and temperature in such a way that the fracture toughness increases with temperature or with decreasing loading rate. For a given loading rate a more or less well-defined temperature exists where the material behavior undergoes a transition from brittle to ductile. In our investigations we conducted two-dimensional discrete dislocation dynamics simulations to study semi-brittle fracture, where a plastic region around the crack tip starts to develop, but cannot shield the crack completely, such that the material still fails by fracture. The simulations show that in the semi-brittle regime crack-tip plasticity and fracture toughness are thermally activated with a constant activation energy. By means of a theoretical analysis of the numerical data an Arrhenius-like relation is derived for loading rate and temperature at points of constant fracture toughness. This scaling relation is also in complete agreement with experimental data of three-point bending tests of tungsten single crystals. Thus, the proposed scaling relation can be used to predict fracture toughnesses in a wide range of temperatures and loading rates, based on only a small number of experiments.

19.1 Introduction

The investigation of the failure of materials is one of the main tasks of materials science, because the failure properties of a material determine the limits of its technological applicability. General failure properties of a material, as for example its yield or rupture strength, as well as the material’s reaction to pre-existing flaws are important engineering parameters. Fracture mechanics studies and describes the strength of pre-cracked materials which can be quite different from the strength of the homogeneous material, due to the stress concentration at pre-cracks or flaws. Generally, failure of materials is classified into brittle and ductile failure. Materials with a face centered cubic structure usually show ductile behavior at all temperatures and deformation rates. Materials with stronger interatomic forces like body centered cubic metals (Gumbsch et al., 1998; Roberts et al., 1994), intermetallic phases (Bergmann and Vehoff, 1995; Ebrahimi and Shrivastava, 1998), semi-conductors (St.John, 1975; Brede and Haansen, 1988; Samuels and Roberts, 1989; Warren, 1989; George and Michot, 1993; Serbena and Roberts, 1994) and also ionic crystals (Narita et al., 2002) often show a transition between brittleness at low temperatures (or high deformation rates) and ductility at high temperatures (or low deformation rates). Materials with a high dislocation density before testing usually exhibit a more or less pronounced semi-brittle regime, where the fracture toughness
rises steadily. The testing of virtually dislocation free specimens like silicon or alumina single crystals reveals an exceptionally sharp brittle-to-ductile transition. Despite this peculiarity silicon has often been chosen as model material for studies of fracture and the brittle-to-ductile transition, because large, dislocation-free single crystals are easily available.

In the following, a model system for crack-tip plasticity of engineering materials is developed that assumes a high density of dislocation sources close to the crack tip, and thus easy dislocation nucleation. The behavior of this model system under constant loading rates is investigated and general findings of the numerical simulation are put in a theoretical framework. It is shown that treating crack-tip plasticity as a thermally activated process yields general relations for the dependence of fracture toughness on temperature and loading rate. A critical comparison of this work with the more classical treatments of fracture mechanics and crack-tip plasticity is given in the discussion.

19.2 Model

In this work an isotropic elastic medium with a semi-infinite crack subject to pure mode I loading with the stress intensity factor (sif) \( K_a \) is considered. The load is applied with a constant rate \( \dot{K} \). The temperature \( T \) is assumed to be constant during the loading process. For simplicity the slip systems are reduced to pure slip of edge dislocations on the slip plane with the highest resolved shear stress, which is inclined to the crack plane by an angle of \( \phi = 70.5^\circ \).

A sketch of the geometry of the system under consideration is provided in Figure 19.1.

![Figure 19.1: Sketch of the model system of a crack loaded in mode I with sif \( K_a \). Dislocations nucleate at the source position and glide on a single slip plane.](image)

Classical treatments of crack-tip plasticity (Rice, 1968; Rice and Thomson, 1974; Weertman et al., 1983; Thomson, 1986; Brede and Haasen, 1988) have been restricted to the case where the dislocation population is always in equilibrium, which inherently assumes static or quasi-static loading conditions. Closed form solutions of the balancing equations can be derived (Weertman et al., 1983; Thomson, 1986) for this case. Under static loading conditions the ductility or brittleness of a material is given by its possibility to emit dislocations below the failure stress (Rice and Thomson, 1974). It is evident that such a static model yields either completely brittle or completely ductile material behavior.
Table 19.1: Young’s modulus $E$, the Poisson ratio $\nu$, and the lattice parameter $a$ employed here are adapted to tungsten (Pearson, 1956). The norm of the Burgers vector is calculated as $b = \sqrt{3}/2a$, representing a $1/2\{111\}$ lattice vector.

<table>
<thead>
<tr>
<th>$E$ (GPa)</th>
<th>$\nu$</th>
<th>$a$ (nm)</th>
<th>$b$ (nm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>393.9</td>
<td>0.29</td>
<td>0.3159</td>
<td>0.274</td>
</tr>
</tbody>
</table>

The systems considered in this work are not necessarily in equilibrium. This allows for a more general description of material behavior. The plastic relaxation is assumed to follow the Orowan law

$$\dot{\varepsilon}(\tau, T) = b\rho v_d(\tau, T),$$

where the plastic strain rate $\dot{\varepsilon}$ is a function of the Burgers vector $b$, the density of mobile dislocations $\rho$ and the dislocation velocity $v_d$. If the strain rate in the plastic zone is equal to or larger than the applied loading rate, the load on the crack-tip does not reach the criterion for crack advance. In this case of ductile behavior the material shows yielding rather than fracture. If, in contrast, the failure criterion is met at a certain stage of loading after some plastic relaxation has occurred, the material fails by fracture. To discriminate such behavior from completely brittle cleavage we speak of semi-brittle fracture. This semi-brittle fracture regime is characterized by an increase in fracture toughness with temperature.

The elastic shielding of the crack tip by the dislocation population, and thus the reduction of the local load on the crack tip by plastic relaxation, are quantified by defining a local sif as (Thomson, 1986)

$$k_t = K_a \frac{E}{4(1 - \nu^2)} \sum_{j=1}^{N_d} \frac{b_j}{\sqrt{2\pi r_j}} \left[ \frac{1}{2}\sin\left(\frac{\varphi_j}{2}\right) + \sin\left(\frac{3\varphi_j}{2}\right) + \frac{1}{2}\sin\left(\frac{5\varphi_j}{2}\right) \right],$$

where $b_j$ and $(r_j, \varphi_j)$ are Burgers vector and position of dislocation $j$ in cylindrical coordinates. It is seen that the shielding of the dislocations is inversely proportional to their distance to the crack tip, and that it accumulates over the total number of dislocations $N_d$. $E$ and $\nu$ are Young’s modulus and the Poisson ratio, respectively. In the numerical examples given here all material parameters are adapted to tungsten (see Table 19.1).

Brittle crack growth occurs once a critical value $k_c$ for the local sif is reached. In the case of a sharp crack tip this critical sif can be expressed in the form of the Griffith criterion (see for example (Rice, 1968))

$$k^2_{c,s} = \frac{2E}{(1 - \nu^2)} S,$$

where $2S$ is the surface energy of the two fracture surfaces. In the case of a blunted crack tip the failure criterion is assumed to be reached, when the tensile stress in front of the crack tip reaches the cohesive strength $R_c$ of the material. Thus, denoting the blunting radius of the crack tip by $r_t$,

$$k^2_{c,b} = 2\pi r_t R_c^2$$

(19.4)
is obtained as the failure criterion for the blunted crack. In this work we define failure of the material as the instance when the failure criterion for the local sif is met, and refer to the applied sif at this stage as the fracture toughness $K_c$. For simplicity, we assume a sharp crack tip throughout and employ the critical value of $k_c = 2 \text{MPa} \sqrt{\text{m}}$, corresponding to the low temperature fracture toughness of tungsten (Gumbsch, 1995; Riedle et al., 1996).

Considering the rate of plastic deformation, it is seen from the Orowan law in Equation (19.1) that either the dislocation nucleation rate (affecting $\rho$) or the dislocation velocity $v_d$ are the rate limiting factors. Analytical analysis (Schöck and Püschl, 1991; Rice and Beltz, 1994) and numerical simulation (Xu et al., 1997) have shown that one can define a critical local sif for dislocation nucleation at crack tips. However, the rate of dislocation nucleation and its dependence on sif and temperature has not yet been investigated systematically. The critical sif values for homogeneous dislocation nucleation at a straight crack tip are found to be very high in body centered cubic materials, such that homogeneous nucleation can be excluded for normal loading conditions. Consequently, experimental (Roberts et al., 1994; Hsia et al., 2001) and theoretical (Zhou and Thomson, 1991; Xu et al., 1997) work has shown that dislocation nucleation occurs preferentially at imperfections in the crack front. In this work a new dislocation is introduced whenever the resolved shear stress at the source position (see Figure 19.1) tends to repel the dislocation from the crack-tip. The source position is located at a distance $r = 30 b$ from the crack tip. For this distance the first dislocation nucleation occurs at 20% of the critical local sif. This procedure causes dislocation nucleation to occur easily, i.e. without energetical barrier, and homogeneously along the crack front, which refers to a case where there is a high density of imperfections at the crack front acting as dislocation sources.

Under this condition one finds that the nucleation rate is limited by the speed at which previously nucleated dislocations leave the crack-tip region (Thomson, 1986; Brede and Haasen, 1988). In face centered cubic metals dislocation velocity is usually regarded to be proportional to the resolved shear stress and widely independent of temperature (see for example (Jassby and Vreeland, Jr., 1970)). In contrast, in body centered cubic materials and also in semi-conductors the velocity of both, screw (Seeger and Šesták, 1978; Brunner, 2000) and non-screw (Schadler, 1964; Prekel et al., 1968) dislocations, must be considered to be thermally activated. In the literature, different dislocation velocity laws have been derived for silicon (Brede and Haasen, 1988; Hirsch et al., 1989), tungsten (Brunner, 2000; Hartmaier, 2000), and NiAl (Ochmann and Vehoff, 1995).

In most of this work the empirical formulation

$$v_d(\tau, T) = v_0 \left( \frac{\tau}{\tau_0} \right)^m \exp \left( -\frac{Q}{k_B T} \right) \tag{19.5}$$

is used, with the pre-exponential factor $v_0$, the normalization stress $\tau_0$, the stress exponent $m$, and the activation energy $Q$ ($k_B$ is the Boltzmann constant). The mobility of non-screw dislocations in tungsten has to be described by a temperature dependent stress exponent of the form $m(T) = a/T + b$ (Schadler, 1964). A more rigorous thermodynamical treatment of dislocation mobility yields (Kocks et al., 1975)

$$v_d(\tau, T) = v_0 \exp \left( -\frac{Q \left[ 1 - (\tau/\tau_0)^m \right]^n}{k_B T} \right) \tag{19.6}$$
Table 19.2: Material parameters describing dislocation mobility of several types of dislocations. For non-screw types Equation (19.5) is employed, while for screw dislocations Equation (19.6) is used. The parameters for non-screw (Schadler, 1964) and screw (Brunner, 2000; Hartmaier, 2000) dislocations in a ⟨111⟩{110} slip system in tungsten are given in the first two rows. The parameters in the last two rows are employed in a parameter study represented in Figures 19.5 and 19.6.

<table>
<thead>
<tr>
<th>type</th>
<th>$v_0$ (m/s)</th>
<th>$\tau_0$ (GPa)</th>
<th>$Q$ (eV)</th>
<th>$m$</th>
<th>$n$</th>
</tr>
</thead>
<tbody>
<tr>
<td>W non-screw</td>
<td>0.32 \cdot 10^{-8}</td>
<td>10^{-3}</td>
<td>0.32</td>
<td>2.81 + (592 K)/T</td>
<td>–</td>
</tr>
<tr>
<td>W screw</td>
<td>3000</td>
<td>1.3</td>
<td>1.8</td>
<td>1</td>
<td>2.8</td>
</tr>
<tr>
<td>A non-screw</td>
<td>0.32 \cdot 10^{-8}</td>
<td>10^{-3}</td>
<td>0.3</td>
<td>5, 10</td>
<td>–</td>
</tr>
<tr>
<td>B screw</td>
<td>3000</td>
<td>0.4, 0.8</td>
<td>0.3</td>
<td>2, 4</td>
<td>1</td>
</tr>
</tbody>
</table>

for the dislocation velocity, where a second stress exponent $n$ is introduced and $\tau_0$ is interpreted as the Peierls stress. This formulation is mainly used to describe screw dislocation mobility in body centered cubic metals. The material parameters employed here are given in Table 19.2. The dependence of crack-tip plasticity on the different formulations of dislocation velocity has been studied previously (Hartmaier and Gumbsch, 2002), where it has been observed that the scaling relation for loading rate and temperature strongly depends on the formulation of $v_d$. In this work the theoretical basis for this observation is given.

The physical quantities $K_a$, $\dot{K}$, and $T$ are the external variables determining the state of the model system as well as the state of experimental fracture samples. Together with the material parameters describing the elastic ($E$, $\nu$) and visco-plastic ($b$, $Q$, $\tau_0$, $m$, $n$, $v_0$) properties, as well as the failure criterion $k_c$, these external state variables determine the fracture toughness of the sample.

The model of crack-tip plasticity described above is expected to capture the most important features of fracture processes in real materials. The most severe simplifications are made by neglecting crack-tip blunting and by assuming homogeneous dislocation nucleation close to the crack tip on a single slip plane, thus restricting the model to the case of intrinsic plasticity. In the case of extrinsic plasticity the dislocations would be generated more homogeneously in the bulk of the material. However, dimensional analysis of energy dissipating processes at crack tips shows that the assumption of intrinsic plasticity is justified for most well-annealed metals (Kysar, 2003). The assumption of homogeneous dislocation nucleation has been investigated previously (Hartmaier and Gumbsch, 1999a,b), where it has been found to be justified at elevated temperatures. In the low temperature regime, in contrast, dislocation nucleation must be considered to occur only at nucleation sites distributed sparsely along the crack front. In the literature discrete dislocation dynamics (DDD) models can be found for intrinsic plasticity (Lin and Thomson, 1986; Hirsch et al., 1989; Brede, 1993; Pippan and Riemelmoser, 1995; Ochmann and Vehoff, 1995; Roberts, 1996; Riemelmoser et al., 2001; Hartmaier and Gumbsch, 1999a, 2002) as well as for extrinsic plasticity (Deshpande et al., 2001, 2003). The models focusing on intrinsic plasticity assume that dislocations nucleate at or close to the crack-tip on slip planes containing the crack front. Such models usually only consider slip either on one or on two slip planes. The shielding of the crack from the applied load is due to the crack-closing stresses of the dislocations, as given in Equation (19.2). The
more recent approaches considering extrinsic plasticity assume that dislocation sources are readily available in the material. Therefore many slip planes become activated and a plastic zone develops than can be compared with the predictions of continuum crystal plasticity (Deshpande et al., 2003). Due to the combination of DDD with a finite element model the blunting of the crack tip can be taken into account additionally to the purely elastic shielding. Finally the combination with a cohesive element technique allows for an explicit description of the fracture process.

The comparatively simple model of crack-tip plasticity employed in this work is certainly not well-suited to give an accurate description of the plastic zone and the fracture toughness under all conditions. However, it allows for a detailed analysis of the interaction of all defects in the crack tip region and, thus, it allows for an improved understanding and a general description of crack-tip plasticity and fracture toughness as a function of loading rate and temperature in the semi-brittle regime. A detailed description of the two-dimensional DDD scheme employed here can be found in (Hartmaier and Gumbsch, 1999a, 2002; Hartmaier, 2000). The DDD method in general is explained in part I, Chapter 8 of this book, an application to thin-film plasticity is given in Chapter 18.

A typical result of a simulation with the DDD model of crack-tip plasticity is represented in Figure 19.2, where the local sif is plotted over the applied load. It is seen that the local sif drops with each dislocation nucleation and increases again as the dislocation moves away, until the next dislocation is nucleated. The value of the local sif immediately before a nucleation event rises with the number of dislocations. As shown later the critical value of $k_t$ for dislocation nucleation is described by Equation (19.10). The simulation is stopped when failure criterion $k_t = k_c$ is reached. In the example shown in Figure 19.2 this yields the fracture toughness $K_c = 7.1 \text{ MPa}\sqrt{\text{m}}$. 

**Figure 19.2**: Local sif plotted over the applied load for $K = 1.0 \text{ MPa}\sqrt{\text{m}}/s$ and $T = 100 \text{ K}$. The data points up to $K_a = 1 \text{ MPa}\sqrt{\text{m}}$ are connected with lines as visual guides, for larger values of $K_a$ only distinct data points are given. The thick dashed line represents the best fit of Equation (19.10) with the constants $C = 1.54 \text{ (MPa)}^2/\text{m}$, $s = 1.45$, and $q = 0.556$. Tungsten non-screw type dislocation mobility is assumed for this simulation (see Table 19.2).
Another important quantity resulting from the numerical simulations is the dislocation distribution. Figure 19.3 shows the dislocation distribution at different stages of the loading. The highest dislocation density is always assumed close to the crack tip. The value of this maximum as well as the spatial expansion of the dislocation distribution increase with the applied load.

In the remainder of this work, the mutual dependencies of state variables and material parameters on transient microstructural quantities like local sif, number of dislocations, and dislocation distribution will be studied.

Figure 19.3: Dislocation distribution at different stages of the loading as given in the legend. The data refers to the same simulation as in Figure 19.2.

19.3 Crack-Tip Plasticity

The following analysis of crack-tip plasticity is based on the force balance between the different defects, i.e., the crack and the dislocations, in the region close to the crack tip. Following the treatise of Weertman et al. (1983), this force balance for an elastic medium containing a crack and a set of dislocations is written as

\[ G = g_t + \sum_{j=1}^{N_d} g_{d(j)}, \]

where \( g_t \) is the force on the crack tip, \( g_{d(j)} \) is the force on dislocation \( j \), and \( G \) is the total force on all defects. We only consider forces in the direction of crack advance, therefore all quantities here are given as real valued scalars. The total force \( G \) is identified with the energy release rate of the crack, which is for pure mode I loading

\[ G = \frac{K_a^2 (1 - \nu^2)}{E}. \]

In the same way, the force on the crack tip can be written as

\[ g_t = \frac{k_t^2 (1 - \nu^2)}{E}, \]

with the local sif \( k_t \) as defined in Equation (19.2).
19.3 Crack-Tip Plasticity

The validity of Equations (19.7) and (19.9) is strictly proven only for a mode III crack and a population of screw dislocations that is symmetrical with respect to the crack plane. However, numerical simulation shows that in our case, too, this force balance is fulfilled to a very good approximation. Figure 19.4 shows that the relative error is extremely small, and stays below $10^{-3}$ even when there are systematic deviations. At the beginning of the simulation, the systematic deviations are caused by numerical round off errors in the division with the small value $K_a^2$. The systematic deviations at later stages that lead to negative values of the relative error are due to disturbances of the system immediately after dislocation nucleation events.

![Figure 19.4: Relative error of the balance of forces on crack tip and dislocations as expressed in Equation (19.7). The sum of the forces on the dislocations is denoted by $S(g_d) = E/(1 - \nu^2) \sum g_d^{(j)}$. The numerical results are obtained under the same conditions as given in Figure 19.2.](image)

The sum over all dislocations in Equation (19.7) cannot be reduced to known quantities without further severe simplifications, but the analysis of the results of DDD simulations with a variety of parameters shows that this term can be expressed as product of powers of the local stress $k_t$ and the number of dislocations $N_d$ in the form

$$E \frac{1}{(1 - \nu^2)} \sum_{j=1}^{N_d} g_d^{(j)} = K_a^2 k_t^2 - k_t^2 = C \left( \frac{k_t}{k_c} \right)^s N_d^q,$$

(19.10)

with three constants $C, s, q$. In Figure 19.2 Equation (19.10) has been solved numerically with respect to $k_t$ as a function of $K_a$ and the result has been fitted to the relative maxima of $k_t$ immediately before a new dislocation is nucleated (the value of $N_d$ for a given $K_a$ is known from the simulation results). It is seen that the fit with the values $C = 1.54$ (MPa)$^2$/m, $s = 1.45$, and $q = 0.556$ gives a good representation of the critical $k_t$ for dislocation nucleation for values of $K_a$ larger than 2 MPa$\sqrt{\text{m}}$, while it underestimates the correct value for smaller $K_a$.

An unexpected and very general result is that these constants are only functions of the material parameters $E, \nu$, and $b$, but neither of the external state variables nor of the parameters or the formulation of the dislocation velocity law. The interpretation of this finding is that the force balance of the crack-tip region is only a function of the number of dislocations that itself, of course, depends on the state variables and the dislocation velocity law. On the microstructural level, the shielding of the crack tip depends only on the dislocation distribution, given by the
number and distances of the dislocations to the crack tip. Therefore, a physical motivation of the finding that the dislocation microstructure – in our confined geometry – depends only on the number of dislocations, but not independently on state variables or parameters of the dislocation velocity, is that the collective dynamics of the dislocation population depends only on the mutual elastic interaction of the defects, which is completely described by $E$, $\nu$, and $b$ (see for example (Lin and Thomson, 1986)). Therefore, the dislocation population may reach the same total number of dislocations at different applied loads, but size and structure of the dislocation distribution will non the less be similar.

The associated observation that the fracture toughness is only a function of the number of emitted dislocations has been reported earlier (Hartmaier and Gumbsch, 1999a, 2002). In the general formalism used here, the fracture toughness is derived from Equation (19.10) by inserting $k_t = k_c$. This yields

$$K_c = \sqrt{k_c^2 + C N_d^3},$$

(19.11)

which again is only a function of the number of dislocations. In Figure 19.5, where $K_c$ is plotted as a function of $N_d$ for a number of loading rates, temperatures, activation energies and stress exponents, it is verified that the fracture toughness does not depend on these quantities independently, but only through the number of dislocations.

![Figure 19.5: Fracture toughness as function of the total number of dislocations for various types of dislocations. The parameters for dislocation mobility refer to Table 19.2, parameters that are varied are given in the legend. If not stated otherwise, the loading rate is $\dot{K} = 1$ MPa/√m/s. For each set of parameters results for several temperatures are represented in the plot. Equation (19.11) with the same constants as in Figure 19.2 is plotted as theoretical curve.](image)

### 19.4 Scaling Relations

A further general result of the DDD simulations is that the number of dislocations $N_d$ is directly proportional to the spatial expansion of the dislocation distribution on the slip plane.
The number of dislocations can thus be expressed as the time integral of the velocity of the first nucleated dislocation

\[ N_d \propto \int_{t'=0}^{t} v_d(\tau^{(1)}, T) \, dt'. \]  

(19.12)

Together with Equation (19.5), and by making use of the constant loading rate in form of the substitution \( dK_a = \dot{K} \, dt \), this gives

\[ N_d = \frac{A'}{K} \exp \left( -\frac{Q}{k_B T} \int_{K'_a=0}^{K_a} \left( \frac{\tau^{(1)}}{\tau_0} \right)^m dK'_a, \right) \]  

(19.13)

with the proportionality constant \( A' \). The separability of the functions containing stress and temperature in the velocity law has been used for the derivation of Equation (19.13). For the velocity law in Equation (19.6) or a temperature dependent stress exponent, stress and temperature are not separable functions. However, the exponential term can always be extracted, but a temperature dependent term under the integral remains. In each case, the integral cannot be evaluated further at this stage, because the dependence of \( \tau^{(1)} \) on \( K_a \) is unknown. In the further, the integral is written as an unknown function \( F(K_a) \). Inserting Equation (19.13) at \( K_a = K_c \) into Equation (19.11) and collecting all terms containing \( K_c \) on the left hand side of the equation yields

\[ \left( \frac{K^2 - k_c^2}{F(K_c)} \right)^{1/q} = \frac{A}{K} \exp \left( -\frac{Q}{k_B T} \right) \]  

(19.14)

with the new constant \( A = C^{1/q} A' \). This implicit expression of the fracture toughness \( K_c \) as a function of the state variables \( \dot{K} \) and \( T \) is at the same time an Arrhenius-like relation for points of constant fracture toughness. Such an Arrhenius-like relation has been postulated already in earlier work (Hartmaier and Gumbsch, 2002) to explain the observed scaling behavior of experimental and numerical results. All combinations \((\dot{K}_i, Q_i/T_i), i \in [1, 2, 3, \ldots] \), of loading rates and temperatures that lead to the same fracture toughness – and consequently to the same total number of dislocations – must fulfill the relation

\[ \frac{\dot{K}_1}{\dot{K}_2} = \exp \left[ -\frac{1}{k_B} \left( \frac{Q_1}{T_1} - \frac{Q_2}{T_2} \right) \right] \]  

(19.15)

or

\[ T_2 = \frac{Q_2}{Q_1} \left( \frac{k_B}{Q_1} \ln \frac{\dot{K}_1}{\dot{K}_2} + \frac{1}{T_1} \right)^{-1}. \]  

(19.16)

Equation (19.16) can be used to re-scale the temperature axis of results obtained for different loading rates and activation energies, thus projecting these results on a single master curve. Such scaling is demonstrated in Figure 19.6 for the velocity law of Equation (19.5). Results obtained for different parameters re-scale exactly on the respective master curves.
If only data from the same material are re-scaled, where $Q_1 = Q_2 = Q_{\text{scal}}$, the scaling procedure defines the apparent activation energy for crack-tip plasticity $Q_{\text{scal}}$. The derivation above has shown that this apparent activation energy for crack-tip plasticity is identical with the activation energy for dislocation motion if stress and temperature contribute to the dislocation velocity as separable functions. It has been shown (Hartmaier and Gumbsch, 2002) that this scaling relation also stays valid to a very good approximation even for velocity laws, where this condition is not met. In this case, however, the slight temperature dependence of the function $F(K_c)$ in Equation (19.14) causes the apparent activation energy for crack-tip plasticity $Q_{\text{scal}}$ used in the scaling relation to be somewhat smaller than the activation energy for dislocation motion. The reduction amounts to $Q_{\text{scal}} = 0.12$ eV for tungsten non-screw dislocations (for which $Q = 0.32$ eV) and to $Q_{\text{scal}} = 1.4$ eV for screw dislocations ($Q = 1.8$ eV). The validity of this scaling relation is also demonstrated in Figure 19.7, where experimental data of three-point bending tests on tungsten single crystals (Gumbsch et al., 1998) are plotted. For this experimental data the apparent activation energy for crack-tip plasticity $Q_{\text{scal}} = 0.19$ eV, obtained by the scaling procedure, is found to be equal to the activation energy for the brittle-to-ductile transition as derived from the brittle-to-ductile transition temperatures at different loading rates.

![Figure 19.6](image1)

Figure 19.6: Fracture toughness-temperature plot with re-scaled temperature axis for type A dislocations (Table 19.2). Results are shown for $Q_1 = 0.3$ eV (open symbols) and $Q_1 = 1.8$ eV (filled symbols) and for $K_1 = 0.1$ MPa$\sqrt{\text{m/s}}$ (circles), $K_1 = 1.0$ MPa$\sqrt{\text{m/s}}$ (triangles), and $K_1 = 10.0$ MPa$\sqrt{\text{m/s}}$ (squares). The constant stress exponents $m$ are given in the legend. The basis values for scaling according to Equation (19.16) are $Q_2 = 0.3$ eV and $K_2 = 0.1$ MPa$\sqrt{\text{m/s}}$.

![Figure 19.7](image2)

Figure 19.7: Experimental data gathered on tungsten single crystals in 3-point-bending tests with constant loading rate and temperature (Gumbsch et al., 1998). Data are provided in a fracture toughness-temperature plot with re-scaled temperature axis. The basis loading rate is $K_2 = 0.1$ MPa$\sqrt{\text{m/s}}$, all loading rates given in the legend are in units of 1 MPa$\sqrt{\text{m/s}}$. The apparent activation energy of crack-tip plasticity has been determined as $Q_{\text{scal}} = 0.19$ eV.
19.5 Discussion

The observation that the total force on the dislocation population on a single slip plane can be expressed as a product of powers of the number of dislocations and of the local $\text{sif}$ has been motivated by a physical reasoning, but certainly has to be subject to further investigations. A comparison with experimental data would be of special interest and is possible via Equation (19.11) which states a relation between the fracture toughness and the total number of dislocations in form of a power law. The empirical constants of this power law are only functions of the elastic properties of a material and its Burgers vector, but not of external state variables or the dislocation velocity law.

A relation between loading rate and the brittle-to-ductile transition temperature has been derived by Brede and Haasen (1988, Equation (18) therein) based on a similar reasoning as the formalism presented here, however, taking into account dislocation distributions for the static case. This relation is similar to Equation (19.14) which is the basis of the scaling relation for points of equal fracture toughness. Brede and Haasen assumed that the dislocation nucleation rate is proportional to the dislocation velocity at the brittle-to-ductile temperature and at a very low stress level (close to the critical stress to move a dislocation). The mathematical formalism of the present work is much more general and based on the result of DDD simulations that the number of dislocations is proportional to the spatial expansion of the inverse pile-up in front of the crack. It follows immediately that the dislocation nucleation rate is proportional to the velocity of the leading dislocation in the inverse pile-up, at whichever temperature and stress level. It is worthwhile noting that this finding is also valid if an athermal friction stress acts on the dislocations, for example caused by phonon drag or by obstacles. Such a constant back-stress on all dislocations impedes their motion, but does not alter their mutual interaction and their collective dynamics. Therefore, the scaling behavior of simulation results with a constant athermal stress is identical as in the cases without friction stress studied above. Since the velocity of the leading dislocation is smaller in the presence of a friction stress, the fracture toughness for otherwise identical internal and external parameters is also smaller. However, the fracture toughness for same numbers of nucleated dislocations is higher than in the absence of a lattice friction, because the dislocations remain closer to the crack tip. If the dislocation motion is completely blocked, as for example at grain boundaries, it must be assumed that the scaling relation is no longer valid and crack-tip plasticity becomes athermal. However, since the spatial expansion of the dislocation distribution is on the micrometer scale, this affects only small grained materials.

The proportionality between the dislocation nucleation rate and dislocation velocity is non-trivial even under the prerequisite that dislocation nucleation occurs easily and homogeneously along the crack front. If dislocation nucleation itself is a thermally activated process, it is possible – depending on the activation energy – that dislocation nucleation is the rate limiting process. In this case, the apparent activation energy for crack-tip plasticity can be assumed to be related to the activation energy of dislocation nucleation, rather than to the activation energy for dislocation motion. There is, however, experimental evidence that in tungsten single crystals dislocation mobility is in fact the rate limiting process, at intermediate and elevated temperatures (Gumbsch et al., 1998).

A practical use of the derived scaling relation is that for a given material the fracture toughness in the semi-brittle regime can be predicted for different temperatures and loading rates.
This requires the knowledge of a fracture toughness-temperature-curve for a single loading rate and the apparent activation energy for crack-tip plasticity, which can be derived from single experiments for at least two different loading rates. Such a model-based extrapolation of experimental data can save quite a large number of experiments and, moreover, the fracture toughness can be predicted for temperatures or loading rates that are not easy to establish in experiments.

It is noted that the model of crack-tip plasticity presented here does not show a transition to purely ductile behavior. This must consequently be attributed to the simplifications of this model. The most severe simplifications in this respect are the neglect of dislocation multiplication and of the blunting of the crack tip. Especially in simulations, where more than thousand dislocations are nucleated, the blunting of the crack-tip should be taken into account and the failure criterion according to Equation (19.4) should be employed. This would give a critical local sif rising with the square root of the number of dislocations. If the local sif then increases with a power of $N_d$ close to or smaller than $1/2$, there are conditions where the critical value for crack advance is practically never met, and the material consequently behaves in a completely ductile manner.

19.6 Conclusions

Two-dimensional discrete dislocation dynamics simulations have been performed to study crack-tip plasticity. The theoretical analysis of the numerical results reveals that crack-tip plasticity and fracture toughness in the semi-brittle regime are thermally activated with a constant activation energy. This activation energy for crack-tip plasticity is identical with or – depending on dislocation velocity – at least closely related to the activation energy for dislocation motion. An analytical scaling relation for loading rate and temperature at points of constant fracture toughness could be derived based on the treatment of crack-tip plasticity as a thermally activated process. This scaling relation also describes experimental results very well, and thus can be employed for a model-based extrapolation of fracture toughnesses of engineering materials.

References


Recent computational advances have permitted mesoscale simulations of systems containing on the order of $10^6$ dislocations. However, while typical dislocation densities in metals range between $10^6$–$10^{11}/\text{cm}^2$, depending on the degree of deformation, it is important to note that many gross features of the macroscopic deformation can be accurately described with constitutive relations that depend on relatively few variables. For example, the plastic response of an fcc single crystal in single-slip conditions is well-described by a single parameter that describes the slope of the stress-strain curve in each stage of hardening, independent of the specifics the many dislocations that make up the microstructure. (Weertman and Weertman (1983); Kocks et al. (1975); Kocks and Mecking (2003)) The large-scale reduction in the number of degrees of freedom required to characterize plastic deformation implies that a homogenization, or coarse graining, of variables is appropriate over some range of length scales. Here, we describe recent work in which temporal and spatial coarse-graining strategies are identified that link the mesoscale with the continuum. Our goal is to show how seemingly unrelated problems can be viewed as part of a unified picture of coarse-grained dislocation behavior.

20.1 Introduction

A successful homogenized theory of dislocation behavior requires the identification of the fundamental coarse-grained variables from among the myriad of dislocation coordinates as well as the time scale for (often overdamped) defect motion. Unfortunately, there has been, to date, little effort to devise workable coarse-graining strategies that properly reflect the long-ranged nature of dislocation-dislocation interactions as well as the short-range structure. If the ultimate goal of a dislocation theory is to predict materials response, then when discussing the coarse-graining (i.e., averaging) of dislocation variables, it is essential to consider what is being left out of the final description. Given clear regimes of behavior, as witnessed by fcc materials, it is likely that alternate coarse-graining approaches may be needed in different regimes.

This chapter is organized as follows. We first briefly review some aspects of recent modeling of the properties of discrete dislocations. We then discuss general issues in coarse graining, followed by specific applications to energetics and dynamics of dislocation structures. Finally, we comment on prospects for the future.
20.2 Dynamics of Discrete Dislocations

Here we briefly review two common methods for simulating microstructures consisting of discrete dislocations. The goal is to discuss assumptions and approximations in the methods, not to give a comprehensive review. More details are given elsewhere in this volume in Chapters 8, 18, and 19.

20.2.1 Dislocation Dynamics Methods

In the Dislocation Dynamics (DD) method, dislocations are represented as discrete objects, the interactions between them are calculated, and their motion in response to those interactions is determined. The first simulation of note represented dislocation lines in a plane, ignored the interactions, and used a line-tension model and interaction with solutes to examine hardening. (Foreman and Makin, 1967) More recently, with the availability of good computational resources, discrete dislocation simulations have become more common.

The key to dislocation dynamics simulations is the representation of the dislocations. Systems of straight, parallel dislocations can be viewed as points in a plane perpendicular to the line direction of the dislocations, with interactions determined from the results for straight dislocations. (Hirth and Lothe, 1967) The first simulations were of such two-dimensional systems. (Lepinoux and Kubin, 1987; Gulluoglu et al., 1989; Amodeo and Ghoniem, 1990) Two-dimensional simulations continue to this day, despite the extreme approximations involved in assuming that all dislocations are straight and parallel.

The extension to three-dimensional dislocation dynamics simulations occurred in the early 1990s. (Kubin, 1993) For three-dimensional simulations, a major issue has been the description used to identify the location of the dislocation line. Kubin and coworkers (Devincre and Kubin, 1994) approximate a dislocation loop as connected straight segments, each with pure edge or screw character. Zbib’s group developed an approach based on the use of connected straight segments with mixed edge and screw character. (Zbib et al., 1998) Ghoniem and coworkers recently introduced a parametric approach in which the dislocations are represented as splines between nodes. (Ghoniem and Sun, 1999) In each of these approaches, interactions between dislocations (usually determined assuming isotropic linear elasticity) are evaluated using line integrals over the dislocations loops, using analytic expressions for straight segments and numerical integrations when employing parametric dislocations. Other ingredients of this type of modeling, regardless of the representation of the dislocations, are source terms, models for dislocation reactions and annihilation, etc. Once the forces on a dislocation are determined, the equations of motion are solved. Typically, one assumes some sort of dissipative motion, very often in the over-damped limit that results in a linear force-velocity relation. Recently, a stochastic discrete dislocations dynamics approach was introduced based on random fluctuations in the stress. (Hiratani and Zbib, 2002)

20.2.2 Phase-Field Methods

Another approach employed to describe the properties of discrete dislocations is the phase-field method, in which the dislocation lines are not modeled directly, but arise from a change in a field, e.g., the displacement field. The energetics are not described as line integrals as in the
20.3 Static Coarse-Grained Properties

DD simulations, but generally involve determination of the dislocation density tensor (by summing over discrete dislocations) and the use of the formalism from continuous dislocations, as described in Section 20.3. The key point in any phase-field method is the identification of spatially-dependent order parameters. One then uses a phenomenological equation of motion for the order parameters to find the time dependent changes in the order parameters. One choice for the dynamics of dislocations is described in detail in Section 20.4, Discontinuities in the order parameters indicate, in this case, the location of the dislocations.

There are a few methods that have been proposed for the phase-field description of discrete dislocations. In the work of Khachaturyan and coworkers, the order parameter is the number of dislocations along a given slip direction with a given Burgers vector. (Wang et al. (2001a,b,c); Jin and Khachaturyan (2001); Wang et al. (2002a,b)) Others have taken a similar approach, with some differences in implementation. (Chen, 2002; Rodney et al., 2003) A recent formulation uses a criteria based on discontinuities in the displacement field. (Koslowski et al., 2002) In the latter work, ground-state distributions of dislocations in the presence of distributed particles in single crystals were determined as a function of strain.

The advantage of the phase-field method is that it can easily be put on a grid and solved with standard numerical techniques. The dislocations are tracked by looking for discontinuities in the order parameters and thus come naturally from the formalism, without any of the complexities needed in the dislocation dynamics methods. There are a number of issues that remain in the phase-field approach. These include the phenomenological nature of the dynamics and uncertainties in how to include details of dislocation sources, dislocation reactions, etc.

20.3 Static Coarse-Grained Properties

The first requirement of a successful theory of dislocation behavior is an identification of the fundamental coarse-grained variables from among the myriad of dislocation coordinates. The challenge to those engaged in dislocation dynamics modeling at the mesoscale is, then, to devise both spatial and temporal homogenization schemes that lead to accurate, yet simplified descriptions of defect energetics and dynamics based on a small number of relevant macrovariables. (LeSar and Rickman, 2003)

Several factors complicate the task of developing such schemes. First, dislocation interactions are long-ranged, so local defect structural details depend, in general, on dislocation configurations some distance away. One implication of this constraint is that strictly local coarse-grained descriptions of dislocation ensembles, based for example on the dislocation density, may be inadequate, necessitating the introduction of density gradient terms in a coarse model. Second, the energy landscape of a system of dislocations is very complex, with a multitude of metastable states separated by barriers. (Wales, 2001) Since the energy scale associated with dislocation interactions is typically larger than the thermal energy (except at very high temperatures), the ground state, i.e., the state with lowest-energy, is not necessarily accessible to the system. The corresponding dislocation structures are inherently complex owing to the anisotropy of dislocation-dislocation interactions, the crystallographic constraints on slip motion, and the kinetic barriers to relaxation. Thus, understanding dislocation structures requires a characterization of the low-lying metastable states. Third, a proper incorporation of non-
conservative modes of motion (e.g., climb) complicates the description even further via the introduction of additional degrees of freedom, namely the coordinates of point defects that facilitate this motion.

Given these difficulties, several interrelated strategies have been proposed for constructing a dislocation dynamics model based on a reduced variable set. Perhaps the most appealing of these strategies, are those based on phenomenological equations of motion that follow from a suitably defined free energy, (Nelson and Toner, 1981; Rickman and Viñals, 1997) which is generally taken to be a functional of the local density of dislocations or, in some cases, the second-rank density tensor that carries both line direction and Burgers vector information for a group of underlying dislocations that are contained in some averaging volume. In Section 20.3.1 these approaches are briefly reviewed.

Another approach to the problem is to strive from the outset to obtain a self-consistent coarse-grained description that follows directly from mesoscale dynamics. One prescription for accomplishing this aim is to somehow integrate out the short-length scales in the system to obtain a Hamiltonian that depends on a reduced variable set. In this context self-consistency would mean that the resulting coarse Hamiltonian faithfully reproduces the energetics and important, large-scale structural features found in the mesoscale model. While this avenue for linking the length scales presents numerous technical difficulties, including the proper definition of an averaging volume and the identification of a restricted variable set, it has the virtue of directly connecting the meso- and macroscales and thereby incorporating short-range interactions and dynamical constraints. In Section 20.3.2.3, we review ongoing work in this area and, in so doing, suggest some promising methods for systematically obtaining a reduced variable set.

20.3.1 Continuous Dislocation Theory

The construction of a coarse-grained model for a collection of dislocations is motivated by earlier unified hydrodynamic treatments of condensed systems. (Martin et al., 1972; Forster, 1975; Marchetti and Saunders, 2002) While dislocations are not strictly equilibrium defects, it is nevertheless useful to apply this formalism to the long-lived metastable state represented, for example, by a heavily-worked metal. From this perspective, the relevant macrovariables are those that decay slowly on the time scale over which the system is observed, namely variables that satisfy some conservation law or those that are associated with a broken symmetry. In the case of a crystalline solid, the displacement field, \( u(\vec{r}) \), relates atomic positions to lattice sites and is an additional slow variable (absent in a simple fluid system) connected with the broken translational and rotational symmetries that give rise to the lattice. Indeed, It has been argued (Martin et al., 1972) that the existence of point defects (e.g., vacancies) follows from the independence of the atomic density and the displacement field at the microscopic scale.

Upon introducing dislocations into the picture, the displacement field becomes multivalued and one can employ the decomposition \( u(\vec{r}) = u_1(\vec{r}) + u_{\text{def}}(\vec{r}) \), where \( u_{\text{def}}(\vec{r}) \) is the singular part of the field associated with a line defect and \( u_1(\vec{r}) \) is the remaining, phonon contribution. (Nelson and Toner, 1981) Furthermore, this singular field can be related to the dislocation density tensor field components (Mura, 1987), \( \rho_{ij}(\vec{r}) \). This field, in turn, satisfies an equation of continuity that embodies the conservation of Burgers vector. (Landau and Lifshitz, 1986) Thus, it is convenient to identify the dislocation density as an additional “slow”
variable needed to characterize a crystalline solid with some dislocation content. (Rickman and Viñals, 1997) As will be seen below, if the dislocation density is regarded as a coarse-grained field, then additional information will be required in the set of macrovariables to characterize properly substructural information.

We start with ideas from continuous dislocation theory. Formulated in the 1950s and 60s by many researchers (including (Kröner, 1958, 2001; Eshelby, 1956a; de Wit, 1960; Mura, 1963; Kosevich, 1979, 1965)), the theory of continuous dislocations is based on a continuous dislocation density tensor. A coarse-graining volume $\Omega$ is introduced, over which we average the dislocation density. The dislocation density is a tensor field with components $\rho_{kl}(\vec{r})$, where $k$ indicates the component of the line direction, $i$ the component of the Burgers vector and $\vec{r}$ a point in space. While an elegant and powerful approach, continuous dislocation theory has not proved to be particularly useful for practical applications. One of the deficiencies in continuous theory is the question of what the dislocation density means, i.e., over what volume are dislocations averaged and what dislocation substructures may be missed in the averaging. This point will be discussed in detail below.

More recent applications of continuous dislocation theory as a basis for models of materials deformation have been presented. Nelson and Toner showed how to define a free energy of a system of dislocations and then use that free energy to discuss a dislocation loop mechanism for transitions in liquid crystals. (Nelson and Toner, 1981) More recently, Shizawa and Zbib used the dislocation density tensor as the basis of a thermodynamic theory of gradient elastoplasticity. (Shizawa and Zbib, 1999) While the main focus of Shizawa and Zbib was on developing a gradient theory, Acharya recently presented work that has the ambitious goal of developing a a continuum theory of elastoplasticity based on the ideas of continuously distributed dislocations. (Acharya, 2001) While the formalism was outlined, no applications were made to predict microstructure or response. Rickman and Viñals (Rickman and Viñals, 1997), starting with the free-energy formalism from Nelson and Toner (Nelson and Toner, 1981), derived a dynamical equation for the dislocation density.

Kosevich (Kosevich, 1979) gave an expression for the interaction energy of a systems of continuous dislocations (assuming isotropic linear elasticity) in terms of Kröner’s incompatibility tensor (which is related to the dislocation density tensor). (Kröner, 1958) The interaction energy was re-expressed by Nelson and Toner (Nelson and Toner, 1981) in terms of the Fourier transform

$$E = \frac{1}{2} \int \frac{d^3q}{(2\pi)^3} K_{ijkl}(\vec{q}) \rho_{ij}(\vec{q}) \rho_{kl}(\vec{q}),$$

(20.1)

where $\rho_{ij}(\vec{q})$ is the transform of the dislocation density tensor, $\mu$ is the shear modulus and $\nu$ is Poisson’s ratio. Repeated indices are summed. The kernel (without core energy contributions) is

$$K_{ijkl} = \frac{\mu}{q^2} \left[ Q_{ik}Q_{jl} + C_{il}C_{kj} + \frac{2\nu}{1-\nu} C_{ij}C_{kl} \right].$$

(20.2)

In Equation (20.2), the projection operators $Q_{ij} = \delta_{ij} - q_iq_j/q^2$, $C_{ij} = \epsilon_{ijk} g_k/q$, and $\epsilon_{ijk}$ are the components of the Levi-Civita tensor density. The energetics of the disordered core regions near each line can be incorporated, at least approximately, by the inclusion of a
phenomenological energy penalty term in the kernel above. Because the dislocation density tensor is solenoidal, we can rewrite the above kernel more in a more convenient form as

\[ K_{ijkl} = \frac{\mu}{q^2} \left[ \delta_{ik} Q_{jl} + C_{il} C_{kj} + \frac{2\nu}{1 - \nu} C_{ij} C_{kl} \right]. \tag{20.3} \]

It is sometimes convenient to work in with the dislocation tensor directly (as opposed to the transform). The interaction energy is this case is (Rickman and LeSar, 2001)

\[ E_I = \frac{\mu}{16\pi} \int \int \epsilon_{ipd} \epsilon_{jmn} R_{mp} (\vec{r}, \vec{r}') \times \left\{ \rho_{ij}(\vec{r}) \rho_{mn}(\vec{r}') + \delta_{ij} \rho_{kl}(\vec{r}) \rho_{kn}(\vec{r}') + \frac{2\nu}{1 - \nu} \rho_{il}(\vec{r}) \rho_{jn}(\vec{r}') \right\} d\vec{r} d\vec{r}' \tag{20.4} \]

where the integrals are over the entire system, \( \rho_{ij} \) are the components of the dislocation density tensor, and \( R_{mk} \) indicates the derivative \( \partial^2 |\vec{r} - \vec{r}'|/\partial x_m \partial x_k \).

We note that in most applications of this theory, such as the phase-field methods mentioned above, the dislocation density is constructed from discrete dislocations, as opposed to being considered as a microvariable of the system. This interpretation is certainly legitimate, as insertion of a density made up of discrete dislocations (i.e., the density is a sum of delta functions indicating the position of the dislocation line) into Equation (20.1) or Equation (20.4) yields the standard interaction terms between dislocations. (Rickman and LeSar, 2001)

Consider now the interpretation of a continuous dislocation density, i.e., when it is applied not for sums of discrete dislocations. In the development of the theory, it is generally assumed that the averaging volume \( \Omega \) is large relative to the dislocation spacing yet small relative to the system size. (Nelson and Toner, 1981) However, it is not clear exactly what this averaging prescription means, nor is it obvious at what scales a continuum theory should hold. If one takes the above assumption that a continuum theory holds for length scales much greater than the typical dislocation spacing, then it restricts applicability of the method to scales much greater than the dislocation structures known to be important for materials response. Kröner, one of the founders of the theory, recently stated that “The greatest shortcoming is that the dislocation density tensor...measures the average dislocation density only and therefore, regards the internal mechanical state utmost incompletely.” (Kröner, 2001) Kosevich is one of the few who discuss the interpretation of various limits to the density. (Kosevich, 1965) The issue thus becomes whether one can use continuum theory to describe actual dislocation microstructures. Certainly, without corrections to the theory, any structure at scales less than the coarse-graining size are ignored.

There are a few ways that one could proceed to improve on the continuum dislocation theory. One option would be to start at continuum level (e.g., Equation (20.4)) and add corrections to the density. The most obvious approach would be to include gradient corrections to include non-local effects, an approach that would be similar to that employed in many local density functional methods, e.g., Thomas-Fermi theory in electronic structure. We choose to take an alternative avenue, in which we start not from continuous dislocations, but from distributions of discrete dislocations. We express the energy in the form of Equation (20.4) and then seek corrections to that energy for systems with finite averaging volumes. Here we summarize results from such an approach, with a more complete derivation given elsewhere. (LeSar and Rickman, 2002, 2003,a)
20.3 Static Coarse-Grained Properties

20.3.2 Extensions to the Continuous Theory

20.3.2.1 Multipole Expansion

Limitations to and improvements of the continuous theory can be better understood by starting from a discrete picture. Consider a small region of space with volume $\Omega$ containing $n$ distinct dislocation loops, not necessarily entirely contained within $\Omega$. We can define a set of moment densities of the distribution of loops in $\Omega$ as (LeSar and Rickman, 2002, 2003a)

$$\rho_{lj}^{(\Omega)} = \frac{1}{\Omega} \sum_{q=1}^{n} b_j^{(q)} \oint_{C^{(q)}, \Omega} dl_{l}^{(q)},$$

(20.5a)

$$\rho_{lj\alpha}^{(\Omega)} = \frac{1}{\Omega} \sum_{q=1}^{n} b_j^{(q)} \oint_{C^{(q)}, \Omega} r_{\alpha}^{(q)} dl_{l}^{(q)},$$

(20.5b)

where $\vec{b}$ is the Burgers vector and the notation $(C^{(q)}, \Omega)$ indicates that we integrate over those parts of dislocation line $q$ that lie within the volume $\Omega$. Here $\rho_{lj}^{(\Omega)}$ is the dislocation density tensor and $\rho_{lj\alpha}^{(\Omega)}$ is the dislocation dipole moment tensor for volume $\Omega$. Higher order moments are given elsewhere (LeSar and Rickman, 2003).

Consider next two regions in space denoted by $A$ and $B$. We can write the interaction energy between the dislocations in the two regions as sums of pair interactions, which can in turn be expressed as line integrals over the dislocation loops (Ghoniem and Sun, 1999; de Wit, 1960) If we take the volumes to be well separated, then the interaction energy can be written as a multipole expansion (LeSar and Rickman, 2002). Then, upon truncating the expansion at zeroth order (i.e., the “charge-charge” term), letting the averaging volume $\Omega \rightarrow dx dy dz$, summing over all regions, we find that we recover the Kosevich form for continuous dislocations in Equation (20.4) and the dislocation density tensor approaches asymptotically the continuous result (LeSar and Rickman, 2003).

Corrections to the Kosevich form associated with a finite averaging volume can be obtained by including higher-order moments in the expansion. For example, the first-order term (“charge-dipole”) has the form (LeSar and Rickman, 2003)

$$E_I^{(cd)} = \frac{\mu}{16\pi} \int \epsilon_{ipt} \epsilon_{jmn} R_{m\mu\alpha\alpha}$$

$$\times \left\{ \rho_{jl}(\vec{r})\rho_{lna}(\vec{r'}) + \delta_{ij}\rho_{kl}(\vec{r})\rho_{kn\alpha}(\vec{r'}) + \frac{2\nu}{1-\nu}\rho_{il}(\vec{r})\rho_{jna}(\vec{r'}) \right\}$$

$$- \left\{ \rho_{jla}(\vec{r})\rho_{ln}(\vec{r'}) + \delta_{ij}\rho_{kla}(\vec{r})\rho_{kn}(\vec{r'}) + \frac{2\nu}{1-\nu}\rho_{ita}(\vec{r})\rho_{jln}(\vec{r'}) \right\}$$

$$d\vec{r}d\vec{r'},$$

(20.6)

where $R_{m\mu\alpha\alpha}$ is the next higher-order derivative of $R_{mp},$ (LeSar and Rickman, 2002) Higher-order terms (e.g., “dipole-dipole”, “charge-quadrupole”) have been presented elsewhere (LeSar and Rickman, 2003a) The expression in Equation (20.6) (and higher-order
terms) can be used as a basis for a continuous dislocation theory with local structure by including the dipole (and higher) dislocation moment tensors as descriptors.

Equations (20.4) and (20.6) can be used as a basis for describing the coarse-grained energy of a system of dislocations. Consider a system described by the coarse-grained continuous dislocation density tensor, \( \rho_{ij}(\vec{r}) \). Now consider overlaying a mesh of the three-dimensional volume. For convenience assume a uniform cubic mesh with \( M = N \times N \times N \) cells, each with volume \( v_s \). The integrals over all space can be rewritten as integrals over the cells. The energy of the system (for now we take only the term in Equation (20.4)) can then be written as sums over the cells

\[
E_I = \sum_{s=1}^{M} \sum_{t=1}^{M} \phi_{st} + \sum_{s=1}^{M} \phi_{ss},
\]

(20.7)

where the prime indicates that the \( s = t \) terms are not included and the terms \( \phi_{st} \) are between cells \( s \) and \( t \). \( \phi_{ss} \) is the cell energy, which we discuss below. If we now assume that we can take the dislocation density tensors to be constant over each cell, the interaction terms take the form

\[
\phi_{st} = \frac{\mu}{16\pi} \epsilon_{ijl} \epsilon_{jmn} \left\{ \rho_{il}^{(s)}(\vec{r}) \rho_{in}^{(t)} + \delta_{ij} \rho_{kl}^{(s)}(\vec{r}) \rho_{kn}^{(t)} + \frac{2\nu}{1-2\nu} \rho_{il}^{(s)}(\vec{r}) \rho_{jn}^{(t)} \right\} F_{stmp},
\]

(20.8)

where all distance information between cells is contained in the function

\[
F_{stmp} = \int_{v_s} \int_{v_t} R_{mp}(\vec{r}_s, \vec{r}_t) d\vec{r}_s d\vec{r}_t.
\]

(20.9)

Given an evaluation of the functions \( F_{stmp} \), the minimum energy dislocation configuration of the system (given the approximation of the density being locally constant over the cells) can be found by a minimization of Equation (20.7) with respect to the parameters \( \{ \rho_{kl}^{(s)} \} \). The global minimum of Equation (20.7) has the solution \( \{ \rho_{kl}^{(s)} \} \to 0 \), which is not very interesting.

We can, however, introduce constraints on the system that will lead to non-zero dislocation densities.

To proceed, we first split the components of the dislocation density tensor, which can be positive or negative, into parts that represent opposite directions of the Burgers vector or line direction, i.e., \( \{ \rho_{kl}^{(s)} \} = \{ \alpha_{kl}^{(s)+} - \alpha_{kl}^{(s)-} \} \). The \( \{ \alpha_{kl}^{(s)+}, \alpha_{kl}^{(s)-} \} \) have the advantage of being positive. The net scalar dislocation density can then be written as the volume-weighted average of the \( \{ \alpha_{kl}^{(s)+}, \alpha_{kl}^{(s)-} \} \) over the system, i.e.,

\[
\rho = \frac{\sum_{s=1}^{M} v_s}{V} \sum_{k=1}^{3} \sum_{l=1}^{3} \left( \alpha_{kl}^{(s)+} + \alpha_{kl}^{(s)-} \right).
\]

(20.10)

For meshes with uniform volumes, \( v_s/V \to M \).

We could then, for example, consider two constraints: 1) constant scalar dislocation densities and 2) zero Burgers vector. The first constraint is simply that \( \rho \) as defined in Equation (20.10) is constant. The second constraint requires that

\[
\sum_{s=1}^{M} \alpha_{kl}^{(s)+} = \sum_{s=1}^{M} \alpha_{kl}^{(s)-}.
\]

(20.11)
20.3 Static Coarse-Grained Properties

We could also add phenomenological source terms, annihilation, etc.

The coarse-grained approach just described brings to light a number of issues. First, no dislocation structure is captured for distances smaller than the cell size – the densities are average properties. Sub-cell structure can be represented by including higher-order multipole moments, as in Equation (20.6), and so one can imagine building a hierarchy of multipoles, increasing the sub-cell description. An important consideration, however, is the cell energy \( \phi_{ss} \), which represents a sum of the dislocation self energies and the interactions between dislocations within a cell. Since dislocations form low-energy dipolar structures, the cell energies can be very large. Numerical studies based on coarse graining structures from discrete dislocation simulations indicate that the cell energies can be much larger than the net interaction energy. Accurate determinations of the cell energies may indeed depend on building into the theory descriptions of the dislocation microstructures. One approach to developing such descriptions is through the use of correlation functions, which are described next.

20.3.2.2 Correlation functions

We have seen that is not possible to develop a meaningful coarse-graining method without incorporating, at least to some extent, the local, sub-scale structures, a conclusion reached some time ago.(Wilkens (1967); Wilkens (1969); Kocks and Wilkens (1967)) A standard approach in statistical mechanics to incorporate the effects of local structures is the use of correlation functions or, equivalently, structure factors. There have been relatively few papers that calculated such functions for dislocation systems.

The first calculation of correlation functions for 2D dislocation systems was in the very early work of Gulluoglu et al.(Gulluoglu et al., 1989) They considered systems of straight, parallel edge dislocations in the case of a single slip system, i.e., along the \( x \) axis. Starting from random dislocation positions, they relaxed the system, finding low-energy metastable states. They then calculated pair correlation functions weighted by the relative sign of the interacting dislocation pairs, finding a strong correlation of like-signed dislocations along the \( y \) axis and of opposite-signed dislocations along the \( 45^\circ \) axes. More recently, a detailed study was made of the structure of a series of two-dimensional simulations with varying slip systems and boundary conditions.(Wang et al., 1998) Once again, the authors considered systems of parallel straight edge dislocations, but with a variety of slip systems (single slip along \( \hat{x} \), double slip along \( \hat{x} \) and \( \hat{y} \), triple slip along \( \hat{x} , \hat{x} + \sqrt{3}/2\hat{y} \), and \( \hat{x} + \sqrt{3}/2\hat{y} \), and random slip directions). Based on a fast-multipole method,(Wang and LeSar, 1995) they studied the effects of differences between full periodic boundary conditions and simulations based on truncated interactions. In addition to defining a series of angularly-weighted pair distribution functions, they also examined clustering of the dislocations. Argaman and coworkers have recently examined the formation of cellular structures for two-dimensional edge simulations, looking at larger-scale correlations.(Argaman et al., 2001a,b) Groma developed a theoretical formalism for calculating one-, two-, three-, etc. particle distribution functions,(Groma, 1997) which formed the basis for calculations for (once again) a system of parallel straight edge dislocations.(Zaiser et al., 2001a) All the calculations mentioned so far have focused on two-dimensional simulations. To date, no similar calculations have been performed for three-dimensional dislocation structures, though a formalism has been developed for this purpose.(El-Azab, 2000)
20.3.2.3 Self-Consistent Coarse Graining Strategies

To connect directly mesoscopic dislocation dynamical models with their macroscopic counterparts, it is imperative to construct a functional of the dislocation density that captures the energetics of the system across length scales and that incorporates the relevant length- and time-scale information needed for a characterization of dislocation microstructures formed via polygonization. This can, perhaps, best be accomplished by self-consistent, numerical homogenization procedures in which the Hamiltonian of the system is obtained at different levels of a coarse-grained hierarchy. Such strategies has been employed in other contexts to create the homogenization necessary to reduce the set of variables at each level of description. Similar methods are just now being brought to bear on the dislocation problem, and so we will merely outline here the salient features of some promising homogenization schemes.

Broadly speaking, there are two complementary avenues to pursue in a numerical coarse-graining strategy in which length scales shorter than some prescribed cutoff are eliminated from the problem. The first, real-space, approach involves subdividing the system into \( n \) equally-sized blocks and then, after integrating out information on scales less than the block size, rewriting the Hamiltonian in terms of block variables. This strategy has been employed in Monte Carlo coarse graining (MCCG) studies of spins systems.(Binder, 1981; Kaski et al., 1984; Gracheva et al., 2000) In these studies, histograms summarizing the probability \( P \) of observing block spin variables \( \{S_i\} \), \( i = 1, 2, \ldots, n \) (often the mean value of a property \( s \) of the objects contained) are tabulated from the results of a MC simulation. The desired coarse-grained free energy of the block variables \( F(\{S_i\}) \) is then obtained by noting that \( P(\{S_i\}) \propto \exp(-F(\{S_i\})/k_B T) \), where \( k_B \) is Boltzmann’s constant and \( T \) is temperature, and assuming some functional form for \( F \). Alternatively, one can adopt the second approach of eliminating short length scales by carrying out the homogenization in reciprocal space. Indeed, this strategy has also been employed in studies (of the critical behavior) of Ising systems.(Swendsen, 1981) We note that cell distribution functions for coarse-grained simulations have been studied, at least to some degree.(Kaski et al., 1984)

The energy scale associated with dislocation interactions is typically much greater than \( k_B T \), so the finite-temperature sampling that forms the basis of the MCCG technique is not well suited for application to systems of dislocations. Nevertheless, it will be advantageous to consider similar coarse-graining strategies and to identify the analogues of the block spin variables. The real-space approach again focuses on subdividing the system into coarse-grained blocks and then calculating the energies associated with these blocks. The long-ranged nature of the interaction dictates, however, that the associated Hamiltonian comprise two separate pieces, one due to interactions between blocks and the other due to interactions between dislocations within a block.(Cannas, 1995) The block-block interactions embody multipolar contributions to the total energy (c.f., Section 20.3.2.1), while the intra-block interactions are somewhat more local. It is important to realize, however, that intra-block contributions to the energy can be significant even for relatively small block sizes, and so the associated energetics must be contained in the Hamiltonian. The challenges here are then to relate the Hamiltonian directly to observed dislocation microstructures, and therefore to characterize the intra-block energetics in terms of macrovariables involving the dislocation density.

Consider now a reciprocal-space approach, currently under development, that allows for compact expression of dislocation multipole moments and the Green function that mediates...
20.4 Dynamic Coarse-Grained Properties

For the purposes of illustration, we imagine that a collection of $N$ straight, edge dislocations, each with $\ell = \hat{z}$ and $\vec{b} = \pm \hat{x}$, is located in real space on the sites of a spatially periodic square lattice. To simplify the situation a bit more, it is assumed that these objects can glide in discrete steps along prescribed slip planes and that, at the low temperatures being considered here, climb is effectively prohibited. Thus, apart from the slip constraints, this system is similar to a two-dimensional lattice gas with each constituent carrying a vector charge. One of a multiplicity of low-energy configurations for this system, often characterized by the presence of dislocation dipoles and walls, can be obtained via zero-temperature MC simulation that properly incorporates the imposed glide and climb constraints as well as the long-ranged nature of the dislocation interactions.

In reciprocal space the relationship between dislocation microstructure and energetics is best highlighted by first defining the structure function, $\tilde{s}(\vec{k})$ for an individual configuration as

$$\tilde{s}(\vec{k}) = \tilde{\rho}_{31}(\vec{k}) \tilde{\rho}_{31}(-\vec{k}),$$ (20.12)

where $\tilde{\rho}_{31}(\vec{k}) = \sum_{i=1}^{N} b_i \exp(i \vec{k} \cdot \vec{r})$, $b_i = \pm b$ and the tilde denotes a Fourier transform. The corresponding energy for this configuration is then

$$E = \sum_{\vec{k}} \tilde{G}(\vec{k}) \tilde{s}(\vec{k}),$$ (20.13)

where $\tilde{G}(\vec{k})$ is the interaction Green function for edge dislocations and the sum is over the wavevectors of the first Brillouin zone of the lattice. It should be noted that, by considering many low-energy configurations, one can define a structure factor $\langle \tilde{s}(\vec{k}) \rangle$ and corresponding average energy $\langle E \rangle$, where $\langle \rangle$ denote an average over the ensemble of configurations.

Our focus here is to develop accurate approximations to the structure factor near the center of the Brillouin zone. It is especially important to build into our description basic information as to angular-dependent spatial correlations among dislocations, as discussed in Section 20.3.2.2.

20.4 Dynamic Coarse-Grained Properties

A history of the theory of dislocation dynamics goes back to the early work of Frank (Frank, 1949), Eshelby (Eshelby, 1956b), Mura (Mura, 1963) and others who deduced the equations of motion for isolated edge and screw dislocations. These treatments dealt with inertial response of line defects in an isotropic medium of density $\rho_0$, having Lame coefficients $\mu$ and $\lambda$, as embodied in the Newtonian equation of motion

$$\rho_0 \frac{\partial^2 u}{\partial t^2} = \mu \frac{\partial^2 u_i}{\partial x_j \partial x_j} + (\mu + \lambda) \frac{\partial^2 u_j}{\partial x_j \partial x_i},$$ (20.14)

where $\vec{u}$ is the displacement field. Indeed, solutions to these equations have been found, for example, for single straight screw and edge dislocations moving in both bulk and semi-infinite media (Frank, 1949; Hirth and Lothe, 1967)
The formulation of equations of motion for an ensemble mutually interacting dislocations at finite temperature presents numerous challenges. As discussed above, it is useful to describe a collection of dislocations in terms of coarse-grained tensor density components \( \rho_{ij} \) and their spatial derivatives. From a hydrodynamic point of view, (Martin et al., 1972; Marchetti and Saunders, 2002) the dislocation density can be regarded as a “slow” macrovariable whose fluctuations decay slowly on the time scale of observation. From this perspective one naturally wishes to construct a time evolution equation for the dislocation density. On physical grounds it is expected that the resulting dislocation dynamics will be dissipative in nature, owing to both mutual dislocation interactions and interactions with other “fast” variables that constitute a thermal bath.

Having identified the relevant slow variables, the construction of a kinetic model is motivated by earlier work in the field of critical dynamics. (Gunton and Droz, 1977; Hohenberg and Halperin, 1977) This model then takes the form of a (possibly stochastic) set of differential equations that reflect any conservation laws that constrain the evolution of the variables. For example, in the absence of sources, the Burgers vector is conserved in a system with dislocations, the corresponding density satisfies such an equation of constraint. In one framework,(Hohenberg and Halperin, 1977) one can describe dislocation dynamics in terms of prototypical Models A through G for systems composed of conserved variables, non-conserved variables or both. In this context Models A and B, which describe the relaxational dynamics of a nonconserved field and the dynamics of a conserved field, respectively, are most relevant. The corresponding equations of motion can be summarized compactly for some prototypical scalar field \( \psi(\vec{r}) \) and free energy functional \( F[\psi(\vec{r})] \) as

\[
\frac{\partial \psi(\vec{r})}{\partial t} = -\Gamma \frac{\delta F[\psi]}{\delta \psi(\vec{r})} + \zeta(\vec{r}, t),
\]

where \( \zeta \) is a stochastic noise term that represents the thermal bath and where \( \Gamma \) is a kinetic coefficient that is a constant for Model A and equal to \( M \nabla^2 \) (related to an Onsager coefficient) for Model B. (Gunton and Droz, 1977)

Various workers have employed the formalism outlined above in dislocation dynamics simulations. For example, in early work in this area Holt (Holt, 1970) postulated a dissipative equation of motion for the scalar dislocation density, subject to the constraint of conservation of Burgers vector, with a driving force given by gradients in fluctuations in the dislocation interaction energy. Rickman and Viñals (Rickman and Viñals, 1997), motivated by an earlier statistical mechanical treatment of free dislocation loops (Nelson and Toner, 1981) and by hydrodynamic descriptions of condensed systems, (Martin et al., 1972) considered a dynamics akin to a noise-free Model B to track the time evolution of the dislocation density tensor in an elastically isotropic medium. Equations of motion for conserved dislocation densities have also been advanced for a viscoelastic medium containing unbound dislocations (Marchetti and Saunders, 2002) and in a continuum model of misfitting heteroepitaxial films. (Haataja et al., 2002) More recently, Khachaturyan and coworkers (Wang et al., 2001c; Jin and Khachaturyan, 2001) have used Model A (i.e., a time-dependent Ginzburg-Landau equation) to describe the relaxational dynamics of a nonconserved density in several phase-field simulations. It should be noted here that Hähner (Hähner, 1996) has criticized the conceptual foundation of this approach given the “non-equilibrium” nature of dislocation glide and given the presence of dislocation sources in most instances. While alternative models, such as those based on
reaction-diffusion phenomena (see below), have undoubtedly contributed as well to our understanding of line defect motion, we view the foregoing approach based on irreversible thermodynamics as an essential starting point for a theory of dislocation dynamics as the underlying conservation law can be modified to incorporate source terms and because non-linearities in the evolution of the dislocation density can be readily introduced into the formalism.

Indeed, the aforementioned reaction-diffusion models of dislocation dynamics are based on a modified equation of continuity that contains a source term reflecting possible dislocation reactions. By contrast with some irreversible thermodynamic formulations, however, these models distinguish among various dislocation “species” by tracking separately the evolution of mobile, immobile and dipolar (scalar) dislocation densities. (Hähner, 1996) This approach has been especially useful in the description of highly-defected systems, including those exhibiting dislocation pattern formation (Walgraef and Aifantis, 1985), and in cases where there are cyclic stresses (Zaiser and Hähner, 1994).

As indicated above, the long-ranged character of dislocation-dislocation interactions complicates modeling efforts, as evidenced by the difficulties encountered when attempting to link the mesoscale with the continuum. The construction of coarse-grained equations of motion is similarly problematic given the dissipative nature of the dynamics of a strongly interacting, defected system with many degrees of freedom. One avenue for incorporating long-ranged interactions into a dissipative dynamics is to employ the critical dynamical models highlighted earlier in this section as they follow from the construction of an entropy production and the identification of fluxes and conjugate thermodynamic forces. From this perspective the Peach-Koehler energies and corresponding forces comprise part of the driving force for dislocation motion, while other fast modes are subsumed into the thermal bath.

Another promising avenue for developing coarse-grained equations of motion also borrows from earlier work on renormalization and dynamic critical phenomena to deduce time evolution operators for homogenized variables (i.e., dislocation densities here). (Mazenko et al., 1980) This is perhaps one of the most conceptually appealing approaches for obtaining self-consistent dynamics from a master equation for individual dislocations. Its tractability in the current context is, however, still undetermined. Finally, we mention an interesting, recent description of the temporal evolution of a collection of dislocations, advanced by Hähner, (Hähner and Zaiser, 1997; Hähner, 2002) that regards the crystal as an effective medium wherein stresses and associated strain rates fluctuate self-consistently, leading to a dissipation of energy. This stochastic dynamical model is predicated on the assumption that the net effect of dislocation interactions with an evolving background configuration is describable in terms of a fluctuating, autocorrelated stress field. One interesting consequence of this model is the presence of fluctuation-induced patterning via the stabilization of non-uniform dislocation configurations.

20.5 Conclusions

There are numerous ongoing efforts to develop an appropriate coarse-graining procedure for dislocation structure and dynamics. Promising avenues of development have typically been based on continuous dislocation theory, offering various enhancements to that approach that include some description of short-range structure. Despite a certain degree of progress, real-
ization of a consistent coarse-graining approach remains elusive. One of the chief stumbling blocks is the metastable (i.e., kinetically driven) nature of the dislocation microstructures. Similar statements could be made about the development of a coarse-grained dynamics of dislocation microstructure. While progress has been made, including the development of a conserved dynamics (Model B) for continuous dislocations, the results have yet to be validated by comparison with experimental or simulation data.

Acknowledgements

The work of R. LeSar was performed under the auspices of the United States Department of Energy (US DOE) and was supported by the Office of Science/Office of Basic Energy Sciences/Division of Materials Science of the US DOE. J. M. Rickman would like to thank the National Science Foundation for its support under grant number DMR-9975384.

References


References


Mazenko, G. F., Nolan, M. J., and Valls, O. T., 1980. Real-space dynamic renormalization group. I. General formal-
References

47, 1159-1162.
and slip systems. Phil. Mag. A 78, 1195-1213.
theory of phase transformations to dislocation dynamics: model and three-dimensional simulations in a single
crystal. Phil. Mag. Lett. 81, 385-393.
1415.
Wilkins, M., 1969. Mean square stress ($\sigma^2$) for restrictedley random distributions of dislocations in a cylindrical
body. Acta Metall. 17, 1155-.
31, 1587-1592.
Zaiser, M., Carmen-Miguel, M., Groma, I., 2001a. Statistical dynamics of dislocation systems: The influence of
Sci. 40, 113-127.
21 Statistical Dislocation Modeling

Randan Sedláček

21.1 Introduction

Extensive numerical simulations, mostly with the aid of the Finite Element Method, have attained a high position in the contemporary engineering praxis. While the geometrical description of the simulated parts or structures is, thanks to the modern isoparametric finite elements, not far from being perfect, the simulation of the material response, especially when plastic deformation (including creep) is involved, is as a rule rather problematic.

Typically, purely empirical relations are fitted to experimentally measured deformation curves. In the course of a simulation, it is often necessary to employ the empirical relations whose fit parameters have no physical meaning outside the stress ($\sigma$) and temperature ($T$) range of experiments used for the fitting, since the type and number of experiments that can be performed is limited by costs and time. For example, cheaper creep experiments at relatively high stresses and temperatures are used for the fitting, whereas lower temperatures and stresses appear in the creep simulation. Such an extrapolation of empirical laws is highly hazardous or even simply wrong. Lacking any physical foundation, an empirical relation applied outside the region of fitting does not necessarily follow even the general qualitative tendencies, let alone quantitative predictions.

As a physically better justified alternative, there are generally applicable elastic-viscoplastic models of plastic deformation that reflect the state of the deforming material, e.g. (Chaboche, 1989; Lemaitre and Chaboche, 2000). Although equations are defined for the evolution of the internal state variables, the constants involved are used as fit parameters with only a restricted relation to the underlying physical mechanisms of the microstructure evolution. A serious practical problem in the application is the complexity of these general constitutive laws and the typically very large number of the fit parameters, which makes the fitting process itself rather difficult. A better understanding of the physical – especially of the microstructural and dislocation – mechanisms involved in the specific process to be simulated would enable to reduce the number of the internal variables and fit parameters for the particular application by choosing the relevant ones.

It is the purpose of this contribution to review a class of microstructure-based constitutive models of plastic deformation developed in the field of materials science to describe and explain the hardening behavior of crystalline materials. The models have been applied mostly to simulate simple uniaxial tests where the specimen geometry plays no significant role, typically at constant strain rate or constant stress, but their potential as constitutive laws for FE simulations begins to be exploited as well. These models account for the state of the deforming material in terms of evolution of a small number (typically one to three) of experimentally
accessible (measurable) internal variables with direct microstructural interpretation, the most prominent of them being the average dislocation density $\rho$. Since the approach to dislocation storage in this class of models is a statistical one, one speaks of statistical dislocation modeling. Admittedly, these microstructure-based models must be fitted to experimental data as well. However, the numerical values that the internal variables attain in the course of the simulations can be checked by physical reasoning and bounds for the fit constants can be set up on physical grounds. Moreover, as distinct from the purely empirical laws, the microstructure-based models have predictive power, which enables not only an extrapolation beyond the range of fitting, but also a prediction of the material behavior for a type of loading different from the one used for the fitting. For instance, in the one-parameter models, Section 21.2, the work-hardening behavior is described by the dependence of the flow stress $\sigma$ on the parameter $\rho$, the plastic strain rate $\dot{\epsilon}$, and temperature $T$, the so-called kinetic equation,

$$\sigma = \sigma(\rho, \dot{\epsilon}, T).$$  \hspace{1cm} (21.1)

Work hardening is assessed experimentally typically by tests at constant strain rate $\dot{\epsilon}$. To describe creep, that is plastic deformation at constant applied stress $\sigma$, the kinetic Equation (21.1) is considered in the inverted form,

$$\dot{\epsilon} = \dot{\epsilon}(\rho, \sigma, T).$$  \hspace{1cm} (21.2)

In both cases, the plastic strain $\epsilon$ is the natural structure evolution parameter, which results in the structure evolution equation for the dislocation density,

$$\frac{d\rho}{d\epsilon} = f(\rho, \dot{\epsilon}, T).$$  \hspace{1cm} (21.3)

Thus, the model is capable of predicting the creep behavior of a material from its work-hardening behavior (Estrin and Mecking, 1984).

This approach appears very useful in describing different modes of plastic deformation in a unified way, which is especially convenient for the simulation of plastic response under varying loading conditions, e.g. in FE simulations. In practice, however, the one-parameter model is often unable to describe the plastic behavior of materials with required accuracy and to account for experimentally observed features of deformation curves, especially those following abrupt changes in deformation conditions, and those related to transition between different hardening stages\(^1\). This is because the dislocation microstructure of a plastically deformed material is in general heterogeneous, typically consisting of dislocation cells or subgrains (Takeuchi and Argon, 1976; Bay et al., 1992; Hughes, 1993). The crucial impact of the microstructure is the fact that the cell walls are plastically harder than the cell interiors. Accordingly, even a pure crystalline material behaves in a way as a composite consisting of soft and hard ‘phases’ (Mughrabi, 1983), and the material behavior cannot be adequately described by the average dislocation density $\rho$ alone. Accordingly, a variety of microstructure-based models have been developed that make use of additional parameters, e.g. mean cell/subgrain size, dislocation densities in cell/subgrain walls and interiors, misorientation between cells/subgrains, etc. (Prinz and Argon, 1984; Nix et al., 1985; Zehetbauer, 1993; Argon and Haasen, 1993;\(^1\) For a review of hardening stages see Rollet and Kocks (1994).
Blum, 1993; Nes, 1998; Roters et al., 2000). In technical materials, even more parameters may be required, for example to account for particle strengthening. There is general agreement neither on the number of parameters to be used nor on the form of equations describing their evolution, nor even on the form of the kinetic equations. This is comprehensible, since the models aim at describing different materials, deformation modes, temperature ranges, etc.

In this chapter, we review a few constitutive laws of this type, starting with the one-parameter models, touching some multi-parameter ones, and finishing with the so-called composite model in the update version of Sedláček and Blum (2002). The latter is suggested to be able to serve as a unified kinematic framework for the microstructure-based multi-parameter models. Focusing on the general structure of the models, we refer to the original literature for the details of the materials-science foundations of the individual approaches.

Remark on Notation In the materials science approach, the normal stress applied to a specimen in a uniaxial test is denoted by $\sigma$. Ruling out any geometrical effects, the assumed uniaxiality allows for the additive decomposition of the material strain $\varepsilon$ that consists of an elastic and a plastic part, the former being expressed in terms of Hooke’s law,

$$\varepsilon = \frac{\sigma}{E} + \varepsilon,$$  \hspace{1cm} (21.4)

with $E$ being the Young modulus. In a polycrystal, the normal stress $\sigma$ is related to the average resolved shear stress $\tau$ responsible for the glide of dislocations by the Taylor factor $M \approx 3$, similarly as the plastic slip $\gamma$ that is due to the dislocation glide is related to the normal plastic strain $\varepsilon$,

$$\sigma = M \tau, \quad \gamma = M \varepsilon.$$  \hspace{1cm} (21.5)

In the following, the variables $\sigma$, $\varepsilon$, $\tau$, $\gamma$ will be used at need, assuming the validity of relation (21.5) implicitly.

In the continuum isotropic plasticity, $\sigma$ corresponds to the von Mises stress and $\varepsilon$ to the equivalent plastic strain,

$$\sigma = \sqrt{\frac{3}{2} T'_{ij} T'_{ij}}, \quad \varepsilon = \sqrt{\frac{2}{3} \varepsilon_{ij} \varepsilon_{ij}},$$  \hspace{1cm} (21.6)

where $T'$ is the deviatoric part of the Cauchy stress tensor, $\varepsilon$ is the plastic strain tensor, and the factor $\sqrt{3}$ relates the normal and shear quantities (Timoshenko and Goodier, 1951; Zienkiewicz and Taylor, 1991). The other possibility, not pursued here, would be the (anisotropic) crystal plasticity, where the crystal slip systems are explicitly accounted for (Asaro, 1983; MacLachlan et al., 2001). The additive decomposition (21.4) implies a geometrically linear (small strain) theory valid for infinitesimal stress and strain increments, which is sufficient for the individual time steps of a FE computation.
21.2 One-parameter Models

21.2.1 Phenomenological Model

In the one-parameter model of Kocks (1976), Mecking and Kocks (1981) and Estrin and Mecking (1984), the glide kinetics that is generally described by an Arrhenius equation is considered in the form of a power law as it is usual in the phenomenological modeling, e.g. (Asaro, 1983). For a prescribed strain rate we have

$$\sigma = \hat{\sigma} \left( \frac{\dot{\varepsilon}}{\dot{\varepsilon}_0} \right)^{1/m},$$  \hspace{1cm} (21.7)

for a prescribed stress we have

$$\dot{\varepsilon} = \dot{\varepsilon}_0 \left( \frac{\sigma}{\hat{\sigma}} \right)^m.$$  \hspace{1cm} (21.8)

Here $\dot{\varepsilon}_0$ is a referential plastic strain rate, $1/m$ is the strain-rate-hardening exponent. The value of $m$ ranges approximately between 200 and 500 (fcc metals at room temperature), the rate independent limit, i.e. $\sigma \equiv \hat{\sigma}$, corresponds to $m \to \infty$. In the rate-dependent formulation (21.7), there is no explicit yielding. In practice, however, $\dot{\varepsilon}$ is exceedingly small for $\sigma < \hat{\sigma}$, so that the hardness $\hat{\sigma}$ plays in fact the role of a yield stress. The hardness (athermal stress) $\hat{\sigma} = M \hat{\tau}$ due to short-range dislocation interactions (forest hardening) is expressed by the standard relation,

$$\hat{\tau} = \alpha G b \sqrt{\rho},$$  \hspace{1cm} (21.9)

with $\alpha \approx 0.3 \pm 0.1$ for fcc metals, $G$ being shear modulus and $b$ magnitude of Burgers vector. Equation (21.7) or (21.8) in combination with Equation (21.9) represent specific forms of the kinetic equations given generally by Equations (21.1) or (21.2). The structure evolution equation (21.3) takes into account the storage of dislocations that causes hardening and the dynamic recovery causing softening,

$$\frac{d\rho}{d\varepsilon} = k_1 \frac{1}{\Lambda} - k_2 \rho.$$  \hspace{1cm} (21.10)

The form of the production term (hardening) is motivated by the statistical approach to the athermal storage of dislocations which become immobilized after having travelled the so-called mean free path $\Lambda$. In pure materials, the mean free path is proportional (with a geometrical factor) to the average spacing between dislocations, $\Lambda \propto 1/\sqrt{\rho}$. Alternatively, the mean free path may be geometrically determined, e.g. by a distance between impenetrable particles or by a small grain/subgrain size. We refer to the original literature (Mecking and Kocks, 1981; Estrin and Mecking, 1984) and to the recent review by Nes (1998, p.166) for the physical interpretation of the ‘constants’ $k_1, k_2$ at the level of dislocation theory. Equation (21.10) in combination with Equation (21.9) represents actually a hardening relation corresponding to the isotropic hardening of continuum plasticity. The work-hardening rate $\theta = d\hat{\sigma}/d\varepsilon$ takes the form $\theta = \theta_h - \theta_i$ that phenomenologically describes stage II hardening evolving towards
21.2 One-parameter Models

a steady state,
\[ \frac{d\rho}{d\epsilon} = 0, \]
where the dynamic recovery compensates the athermal storage, \( \theta_h = \theta_c \). As noted by Nes (1998, p.166), the one-parameter model cannot account for any microstructural changes related to the transition between stages II and III. Roughly speaking, more than one microstructural parameter are required to describe more than one hardening stages.

The phenomenological relation between the steady-state dislocation density and stress,
\[ \rho \approx \left( \frac{\sigma}{Gb} \right)^2, \]
that is approximately valid in pure materials as well as in solid solution for a wide range of temperatures and deformation conditions (Blum, 1993; Blum et al., 2002), enables to check the order of magnitude that the internal variable \( \rho \) attains in the course of a simulation. We note in passing that if Equation (21.11) would hold exactly, then, in steady state, \( \sigma \propto \dot{\epsilon} \), Equation (21.9), and the steady-state strain rate, Equation (21.7), would not change with stress once a referential strain rate \( \dot{\epsilon}_0 \) has been chosen, which is obviously unacceptable. The same holds for the more elaborate kinetic equations, e.g. the physically-based sinh-law (21.14) below. If Equation (21.11) would hold exactly, then the dislocation velocity \( v \) would be independent of the applied stress \( \sigma \) and, accordingly, of the steady-state strain rate. However, this would be in contradiction with the observed high stress sensitivity of the steady-state strain rate. In fact, there is a slight systematic deviation from the exact form of the relation (21.11) that is, in combination with the high stress sensitivity of the glide velocity, Equation (21.14), sufficient to capture the stress dependence of the steady-state strain rate. Detailed quantitative modeling of the dislocation density evolution, e.g. in the framework of the two-parameter model (Blum et al., 2002), cf. also Chapter 31, clarifies this point.

The capability of the one-parameter model to predict transient creep behavior (primary creep) from the work hardening behavior has been demonstrated by Estrin and Mecking (1984), together with the technique of fitting the model parameters to experimentally measured stress-strain curves. Since only constant strain rate and constant stress conditions were simulated, the model equations were integrated analytically, resulting in an exponential decay towards the steady state (Estrin and Mecking, 1984). To simulate the material response under varying deformation conditions, the model equations must be solved numerically. Consider a coarse grained, pure crystalline material deformed in a stress-control mode, i.e. \( \sigma = \sigma(t) \) prescribed, as an example. Then, the constitutive law, Equations (21.7), (21.9), and 21.10, can be set up in the form of an initial value problem for a system of two coupled ordinary differential equations in the variables plastic strain \( \epsilon(t) \) and dislocation density \( \rho(t) \), in a time interval \( t \in (0, T) \),
\[ \frac{d\epsilon}{dt} = \dot{\epsilon}_0 \left( \frac{\sigma(t)}{M_GGb\sqrt{\rho}} \right)^{m}, \quad \epsilon(0) = \epsilon_{\text{ini}}, \]
\[ \frac{d\rho}{dt} = (k_1 \sqrt{\rho} - k_2 \rho) \dot{\epsilon}, \quad \rho(0) = \rho_{\text{ini}}. \]
The system (21.12) can be integrated using standard methods of numerical mathematics, where the adaptive step length of the numerical integration is directed by the required accuracy of the solution. Also, special methods for stiff systems (implicit integration) can be readily utilized if necessary (Press et al., 1989).

### 21.2.2 Materials Science Approach

As distinct from the phenomenological modeling, Section 21.2.1, it is common practice in materials science to relate the plastic strain rate to the velocity $v$ and density $\rho$ of dislocations via the Orowan equation,

$$\dot{\gamma} = b_p \rho v .$$  \hfill (21.13)

Applying the Orowan equation in the context of the one-parameter model, one implicitly assumes that the same dislocation density $\rho$ that carries the plastic deformation is at the same time stored in the crystal to annihilate, thus providing hardening and recovery mechanisms. Because of the dual role played by the dislocation density $\rho$, an Orowan-equation based model is able to account for the non-monotonic variation of the deformation resistance with the increasing dislocation density that is observed in materials with a low initial dislocation density (usually not in metals). At the beginning of deformation, the density of dislocations able to carry the plastic deformation is low, causing a high deformation resistance, Equation (21.13). The high dislocation density in advanced stages of deformation increases the hardness, Equation (21.9), thus causing a high deformation resistance again. In between, there is a point of minimum deformation resistance. In tests at constant strain rate, it corresponds to a minimum of flow stress, in tests at constant stress to a maximum of plastic strain rate (Blum, 1993).

On the contrary, in the phenomenological model, Section 21.2.1, the deformation resistance monotonically increases with increasing dislocation density.

Assuming that the dislocation glide involves thermally activated overcoming of localized obstacles, e.g. the forest dislocations, the dislocation velocity $v$ can be expressed as

$$v = v_0 \exp \left(- \frac{Q}{k_B T} \right) \sinh \left( \frac{b \Delta a^* (\tau - \hat{\tau})}{k_B T} \right) \quad \text{for} \quad \tau > \hat{\tau} ,$$  \hfill (21.14)

where $Q$ is an effective activation energy. Following Vogler and Blum (1990); Meier and Blum (1993); Blum (1993, 1996); Blum et al. (2002), we have denoted the true activation volume of the dislocation velocity by $b \Delta a^*$ to distinguish it form the effective activation volume of the stationary creep rate that is in common use. As distinct from $V$ that comprise information on the deformation microstructure, the true activation volume of the dislocation velocity is in ideal case independent of microstructural changes. For a detailed discussion of the true activation area $\Delta a^*$ and how it is assessable by experiments, see (Blum et al., 1996; Blum, 1996). As for the pre-factor $v_0$, Blum et al. (2002) suggest $v_0 = b v_D \lambda_0 / \lambda$, where $v_0 = v_D$ is the Debye frequency, $\lambda_0$ is the jump width, i.e. the mean planar-obstacle spacing, and $\lambda$ the spacing of thermal obstacles along the dislocation lines. We note that in the definition of the effective (thermal) resolved shear stress $\tau - \hat{\tau}$ that is available for dislocation motion, the hardness $\hat{\tau}$, Equation (21.9), explicitly takes over the role of yield stress.
21.3 Multi-parameter Models

21.3.1 Various Approaches

The first step towards a more realistic description of plastic deformation is to distinguish the ‘mobile’ dislocations from the ‘immobile’ or ‘stored’ ones. An update two-parameter model of this kind has been presented by Blum et al. (2002) and is reviewed by Blum, Chapter 31. The next step is to consider various cellular or subgrain structures that form from the onset of stage II hardening as well as of primary creep. A variety of microstructure-based multi-parameter models results, that allow to reflect the physics of plastic deformation more realistically compared to the one-parameter model. The models enable to improve the simulation of material response to changes in deformation conditions, and to account for successive hardening stages. Despite of many differences, the multi-parameter models share many common features. The plastic strain rate is related to the velocity of mobile dislocations via the Orowan Equation (21.13). The dislocation velocity depends on the effective stress $\tau - \tilde{\tau}$ via a sinh-law of the type (21.14) that reflects the thermal activation of the dislocation glide\(^2\). The hardness $\tilde{\tau}$ is in principle determined by Equation (21.9), although other effects may contribute to it (e.g. the Orowan stress).

In a three-parameter model, Roters et al. (2000) use an element of the composite framework, viz the rule of mixtures, Equation (21.18) below, to determine the average flow stress from the local stresses in the hard walls and soft cells. They assume the same velocity of the mobile dislocations passing through the cells and walls, which is justified at later stages of deformation, but not from the onset of plastic flow, cf. Section 21.3.2. The three parameters considered in the model are the densities of the mobile dislocations ($m$), and of two kinds of the immobile ones: those stored in cell interiors ($i$) and those stored in cell walls ($w$), both in the form of dipoles or locks. Dislocation interactions were considered in detail and structure evolution equations in the spirit of Equation (21.10) have been set up for all three dislocation densities,

$$\dot{\rho}_x = \dot{\rho}_x^+ - \dot{\rho}_x^- , \quad x = m, i, w . \quad (21.15)$$

The time dependence of the parameters is mostly implicit via the plastic strain rate $\dot{\epsilon}$, similar to Equation (21.12), but may as well be explicit for some processes (e.g. annihilation of dipoles by climb or cross slip). Equation (21.13) and a sinh-law of the type (21.14) constitute the kinetic equations in the soft and hard ‘phases’. The model is carefully set up and attention is paid to the underlying physics of dislocation interactions. Moreover, it is demonstrated that by introducing a time-dependent slip length into Equation (21.10), the model can account for Ostwald ripening of precipitates in the course of the hot-forming process that is simulated for

\(^2\) An exception is the viscous dislocation glide in high-temperature creep of solid solution with linear dependence of the dislocation velocity on the effective stress (Endo et al., 1984).
simplicity as a uniaxial test at constant strain rate. We refer to the original work (Roters et al., 2000) for further details.

In his three-parameter model, Nes (1998) considers the cellular structure as well. As distinct from Roters et al. (2000), however, he assumes the cell walls (in fact subgrain boundaries) to consist of geometrical dislocations. As distinct from dipolar cell walls, the subgrain boundaries accommodate misorientations between the subgrains, which is typical of hardening stages IV and V as well as of high-temperature creep. The subgrain boundaries are treated as obstacles to dislocation glide causing the dislocations to bow-out between them, in accord with more detailed dislocation models (Sedláček and Forest, 2000) as well as with discrete dislocation simulations (Devincre et al., 2001). The Orowan stress introduced in this way acts as a back stress in the subgrains. Being inversely proportional to the subgrain size, the Orowan stress introduces an intrinsic length scale into the model. Moreover, a stress concentration factor has been introduced, causing a forward stress in the boundaries, a phenomenon that automatically comes out in the full composite framework, Section 21.3.2. The three parameters used in the model are: dislocation density inside the subgrains, dislocation density in the boundaries that is directly related to the boundary misorientation, and the cell size that introduces a contribution to the hardness through the Orowan stress. Evolution equations of the type (21.15) are presented for all these variables. The model aims at simulating work hardening up to stage IV. Also in this model, a great deal of attention is paid to the underlying physics and we refer to the original work (Nes, 1998; Nes et al., 2000; Nes and Marthinsen, 2002) for further details.

Simulating high-temperature creep, Blum et al. (1989a,b); Vogler and Blum (1990) deployed the full composite approach of Mughrabi (1983), admittedly restricted to constant-stress deformation mode. The composite approach leads to the build up of internal stresses that enable compatible deformation of the heterogeneous microstructure, i.e. a back stress in the subgrains and a forward stress in the walls. As distinct from the Orowan stress, the stresses introduced in this way depend on volume fraction of the walls, not on the absolute dimensions, thus introducing no intrinsic length scale into the model. A detailed dislocation-based model of Sedláček and Forest (2000) clarifies the relation between the composite approach and the Orowan stress. The model of Blum makes use of four parameters: mobile dislocation density in the subgrains, stored dislocation density in the subgrain boundaries, the subgrain size and the volume fraction of the subgrain boundaries. Evolution equations have been defined in a phenomenological manner, reflecting the exponential decay of the parameters from initial to steady-state values, both experimentally justified. The model was successfully applied to simulate creep curves including abrupt changes of stress for a wide range of materials ranging from aluminium, various technical alloys, semiconductors, and steel to rock salt, cf. e.g. (Vogler and Blum, 1990; Hofmann and Blum, 1993; Meier and Blum, 1993; Blum, 1993; Henes et al., 1999; Hunsche and Hampel, 1999).

For the numerical integration of the differential equations resulting from the multi-parameter models, various ad-hoc methods with unclear mathematical properties (convergence, accuracy, stability) are often used. This is because in these models, attention is usually paid to the details of the dislocation mechanisms and to the fitting to experimental curves, but not to the kinematic framework and mathematical structure. Typically, the simplest integration method, i.e. the Euler’s method, is implemented. The stepsize control in these implementations is typically empirical, though physically motivated. For example, if the increase of
plastic strain or change in dislocation density in one step is greater than a certain limit, the stepsize is reduced. The limit is determined so that the integration of a typical deformation curve succeeds, no attention being paid to the accuracy of the computation.

21.3.2 Composite Framework

The composite model of Mughrabi (1983), cf. also (Mughrabi and Ungár, 2002), treats the dislocation microstructure as a two-phase structure consisting of plastically hard walls with high dislocation density and soft cell/subgrain interiors with comparatively low density of dislocations. The approach leads in a natural way to the build-up of internal stresses that enable a compatible deformation of the soft and hard phase. The dislocation velocity in the both phases evens out not until a sufficient magnitude of the internal stresses have been established. The internal stresses rearrange after any change in deformation conditions. The build-up of internal stresses corresponds to the \textit{kinematic hardening} of continuum plasticity (Asaro, 1975), that is related to the Bauschinger effect. Of course, this framework can be applied to truly two-phase materials (e.g. nickel-based superalloys) as well.

We suggest that the composite framework of Mughrabi (1983) is able to serve as a unified kinematic framework for the microstructure-based multi-parameter constitutive models. Following Sedláček and Blum (2002), the composite framework is presented in a general form, thus allowing for implementation of various kinetic and structure evolution equations. A general macroscopic constitutive law results in the form of an initial value problem for a set of ordinary differential equations, Equations (21.26) below, that can be solved by standard methods of numerical mathematics for any prescribed time evolution of applied stress or strain. To arrive at a mathematically tractable formulation in terms of averaged quantities, the continuum-mechanics equations describing the non-homogeneous plastic deformation at the microscale are considered in a simplified form. At the heart of the model is the decomposition (21.4). Instead of rigorous strain compatibility, iso-strain condition at the micro scale is considered, i.e. material strains in the soft (s) and hard (h) regions equal the overall material strain $\varepsilon$,

\begin{align}
\varepsilon_s + \frac{\sigma_s}{E} &= \varepsilon, \\
\varepsilon_h + \frac{\sigma_h}{E} &= \varepsilon.
\end{align}

The average stress equilibrium (Albenga’s law) takes the form of the rule of mixtures,

\begin{equation}
\sigma = f_s \sigma_s + f_h \sigma_h, \tag{21.18}
\end{equation}

with the volume fractions of the hard and soft regions fulfilling the condition $f_s + f_h = 1$.

The overall mean plastic strain then follows from Equations (21.4) and (21.16) – (21.18),

\begin{equation}
\varepsilon = f_s \varepsilon_s + f_h \varepsilon_h. \tag{21.19}
\end{equation}

The kinematic framework, Equations (21.4) and (21.16) to (21.19), must be completed by specific (microscopic) constitutive equations, similar to Equations (21.2) and (21.3) of the one-parameter model. Based e.g. on Equations (21.13) and (21.14), kinetic equations can be
set up for the local plastic strain rates in the soft and hard regions as functions of the local stresses, the current microstructure, and temperature \(T\),

\[
\frac{d}{dt} \epsilon_s = \dot{\epsilon}_s(t, \sigma_s, X, Y, T) ,
\]

\[
\frac{d}{dt} \epsilon_h = \dot{\epsilon}_h(t, \sigma_h, X, Y, T) .
\]

Evolution equations must be specified for the sets \(X\) and \(Y\) of microstructure parameters. The parameters \(X\) evolve with plastic strain (e.g. the athermal storage and dynamic recovery),

\[
\frac{d}{dt} X = \dot{X} (\epsilon, X, Y, T) ,
\]

where the symbol \((\dot{ })\) denotes plastic-strain derivative. The evolution of the microstructure set \(X\) with time,

\[
\frac{d}{dt} X = \dot{X} \epsilon ,
\]

is formulated using the plastic strain rate \(\dot{\epsilon}\), Equation (21.25) below. The parameters \(Y\) evolve explicitly with time (e.g. static recovery of dislocation density or Ostwald ripening of particles),

\[
\frac{d}{dt} Y = \dot{Y}(t, X, Y, T) ,
\]

Note that the volume fractions in general depend on the evolving microstructure parameters. For instance, the time rate of change of \(f_h = f_h(X(\epsilon(t)), Y(t))\) is

\[
\frac{df_h}{dt} = \frac{\partial f_h}{\partial \epsilon} \dot{\epsilon} + \frac{\partial f_h}{\partial \tau} ,
\]

the same for \(f_s\). Using the relation \(f_s + f_h = 1\), the plastic strain rate results from time derivative of Equation (21.19) in the form

\[
\dot{\epsilon} = \frac{f_s \dot{\epsilon}_s + f_h \dot{\epsilon}_h + \frac{\partial f_h}{\partial \epsilon} (\epsilon_h - \epsilon_s)}{1 - \frac{\partial f_h}{\partial \epsilon} (\epsilon_h - \epsilon_s)} .
\]

We aim at a macroscopic time-dependent relation between the flow stress \(\sigma(t)\) and plastic strain \(\epsilon(t)\), which can be computed by solving the kinetic Equations (21.20), (21.21) and the evolution Equations (21.23), (21.24). The kinematic framework, Equations (21.4) and (21.16) to (21.19), enables us to re-formulate these equations by specifying the dependence of \(\sigma_s, \sigma_h\) and \(\epsilon\) on \(\sigma, \epsilon, \epsilon_s, \epsilon_h, X, \) and \(Y\), so that a closed system of differential equations results in the form,

\[
\frac{d\epsilon_s}{dt} = \dot{\epsilon}_s(t, \epsilon_s, \epsilon_h, X, Y, T) , \quad \epsilon_s(0) = \epsilon_{s,ini} ,
\]

\[
\frac{d\epsilon_h}{dt} = \dot{\epsilon}_h(t, \epsilon_s, \epsilon_h, X, Y, T) , \quad \epsilon_h(0) = \epsilon_{h,ini} ,
\]

\[
\frac{dX}{dt} = \dot{X} (\epsilon_s, \epsilon_h, X, Y, T) \dot{\epsilon}(t, \epsilon_s, \epsilon_h, X, Y, T) , \quad X(0) = X_{ini} ,
\]

\[
\frac{dY}{dt} = \dot{Y}(t, X, Y, T) , \quad Y(0) = Y_{ini} ,
\]

\[
\frac{d}{dt} \epsilon_s = \dot{\epsilon}_s(t, \sigma_s, X, Y, T) ,
\]

\[
\frac{d}{dt} \epsilon_h = \dot{\epsilon}_h(t, \sigma_h, X, Y, T) .
\]

\[
\frac{d}{dt} X = \dot{X} (\epsilon, X, Y, T) ,
\]

\[
\frac{d}{dt} Y = \dot{Y}(t, X, Y, T) .
\]
where \( \dot{\epsilon} \) is given by Equation (21.25). The dependence of \( \epsilon \) in Equation (21.22) on \( \epsilon_s, \epsilon_h \) and \( X \) is given by Equation (21.19). In specifying the dependence of \( \sigma_s, \sigma_h \) in Equations (21.20) and (21.21) on \( \sigma, \dot{\epsilon}, \epsilon_s, \epsilon_h, X \) and \( Y \), we have to distinguish between the following deformation modes.

**Stress Control** In a stress controlled test, the applied stress \( \sigma(t) \) as a function of time is prescribed. The local stresses \( \sigma_s, \sigma_h \) entering Equations (21.20) and (21.21) are derived from Equations (21.4) and (21.16) to (21.18) as functions of the local plastic strains \( \epsilon_s, \epsilon_h \), and the microstructure-dependent volume fractions \( f_s, f_h \).

\[
\sigma_s = \sigma - Ef_h(\epsilon_s - \epsilon_h), \quad \sigma_s = \sigma - Ef_s(\epsilon_h - \epsilon_s) .
\]  

(21.27)  

(21.28)

The system (21.26) is solved for the local plastic strains \( \epsilon_s(t), \epsilon_h(t) \) and the set of microstructure parameters \( X(t), Y(t) \). The plastic response \( \epsilon(t) \) is found with the help of Equation (21.19). The material strain \( \epsilon(t) \) follows from Equation (21.4).

**Strain control** In a strain controlled test, the material strain \( \epsilon(t) \) as a function of time is prescribed. The local stresses \( \sigma_s, \sigma_h \) entering Equations (21.20) and (21.21) follow from Equations (21.16) and (21.17) as functions of the local plastic strains \( \epsilon_s, \epsilon_h \).

\[
\sigma_s = E(\dot{\epsilon} - \epsilon_s), \quad \sigma_s = E(\dot{\epsilon} - \epsilon_h) .
\]  

(21.29)  

(21.30)

The system (21.26) is solved for the local plastic strains \( \epsilon_s(t), \epsilon_h(t) \) and the set of microstructure parameters \( X(t), Y(t) \). The plastic strain \( \epsilon(t) \) follows from Equation (21.19) and the applied stress \( \sigma(t) \) from Equation (21.4).

The proposed framework was tested by Sedláček and Blum (2002) by employing specific kinetic and structure evolution equations and simulating several tests at varying deformation conditions in both stress and strain control modes.

### 21.4 Conclusions

In the field of materials science, a variety of microstructure-based constitutive models has been developed, primarily to deepen our understanding of plastic deformation in terms of dislocation mechanisms. As a rule, a great deal of attention has been paid to the individual physical mechanisms, whereas the kinematic framework and mathematical structure of the models have been considered only casually. Moreover, trying to capture the active mechanisms in more and more detail, most of the models have become more and more complex, thus becoming less and less attractive for a possible implementation in FE applications, where the efficiency and CPU time are often the decisive criteria for the choice of a constitutive law. To make these models more attractive for possible FE applications, a pragmatic approach is needed to identify the ingredients that are really necessary to capture the decisive features of the deformation curves for specific deformation conditions of a particular process to be simulated. Then, efficient microstructure-based models with clear mathematical structure and without abundant complexity could be tailored for specific applications. Depending on the
required accuracy of the FE simulation, more or less expensive models (and, therefore, fitting procedures) could be chosen. The unified kinematic framework presented in Section 21.3.2 is a step in this direction.

Acknowledgements

Sincere thanks are due to Prof. W. Blum for critical reading of the manuscript and for many useful suggestions.

References


22 Taylor-Type Homogenization Methods for Texture and Anisotropy

Paul Van Houtte, Saiyi Li, and Olaf Engler

22.1 Introduction

Crystallographic textures can develop at various stages in thermo-mechanical processing of sheet metals. They cause anisotropy of the mechanical and physical properties of the material. In principle this should be taken into account during the simulation of forming processes with the finite element (FE) method. In addition to that, the evolution of the texture due to the plastic deformation should be simulated as well. In this chapter, focus will be on the second aspect. The first aspect has been dealt with in Chapter 6.

A texture updating procedure incorporated in a FE model is useless unless better quantitative descriptions of the current anisotropy can be derived from the updated textures. This also implies that a good method for quantitative descriptions of textures must be available. The theory proposed by Bunge (1982) will be used for that purpose in the present chapter. A crystallographic orientation is represented by a set of three Euler angles: \( g = \{\varphi_1, \Phi, \varphi_2\} \). The crystallographic texture is then represented by an Orientation Distribution Function (ODF) \( f(g) \) which is defined as follows: \( f(g)dg \) is the volume fraction of all grains which have a crystallographic orientation belonging to an elementary volume \( dg \) around \( g \) in Euler space. The ODF can easily be obtained from X-ray pole figures or from automated EBSD (electron backscattering diffraction) scans with the help of specialized software (Bunge, 1982; Kocks et al., 1998; Van Houtte, 1991; Randle and Engler, 2000).

This chapter is about Taylor-type homogenization methods in models for the plastic deformation of polycrystalline materials. The term ‘Taylor-type homogenization method’ means here that simple assumptions are used for the relationship between macroscopic quantities (average flow stress, average plastic strain) and mesoscopic ones (stresses and plastic strains at the level of the individual grains). The best known of these assumptions has been formulated by Taylor (1938). It states that the plastic strain of all crystallites within a polycrystal is the same, and hence equal to the plastic macroscopic plastic strain. Other assumptions will also be explained, namely those used by various types of so-called relaxed constraints theories. All these models have in common that the active slip systems, the slip rates, the rate of lattice rotation, and the deviatoric stress acting in the individual grains can be calculated for each grain separately, or at most for a group of a few grains. As a result, calculation time is much smaller as compared to the Crystal-Plasticity Finite Element Method (CPFEM, see Chapter 26 and 28) which work with a finite element mesh placed over the microstructure (each element representing a single grain, or part of a single grain), or Self-Consistent models.
(SC, see Chapter 23), which assume that the interaction of a grain with all other crystals can be approximated by the interaction between the grain and a hypothetical homogeneous matrix (surrounding the grain) which has the average constitutive behavior of the entire polycrystal.

22.2 Local Constitutive Laws (Mesoscopic Scale)

At the mesoscopic scale, a small volume element (of a size which can range from a few nm to several µm) is considered. It must be small enough so that it can be regarded to have uniform properties (including crystal orientation), and that the stress and strain distributions in it can be treated as homogeneous. In Taylor-type models, such volume element is identified with a single crystallite within the polycrystal.

Taylor-type models usually neglect elastic strains. The local constitutive law consists of a kinematical equation and an energetic assumption. The kinematical equation relates the velocity gradient \( \mathbf{l} \) which describes the local deformation (per unit time) with the slip rates of all slip active slip systems

\[
\mathbf{l} = \mathbf{W}^L + \sum_{s=1}^{N} \hat{\mathbf{b}}^s \otimes \mathbf{m}^s \dot{\gamma}_s
\]  

(22.1)

Note that the lower case symbol \( \mathbf{l} \) means the velocity gradient at mesoscopic scale whereas the upper case symbol \( \mathbf{L} \) will be used for the macroscopic velocity gradient. The symbol \( \otimes \) represents the dyadic product between two vectors (\( x \otimes y \) is a second rank tensor with components \( x_i y_j \)). A similar convention will be used for other tensors (stress, strain rate etc.). The tensor equation above is valid in any frame which does not rotate with respect to the reference frame of the polycrystal. The latter will be called macroscopic frame in the rest of this chapter. In order to express the equation in scalar notation, either the macroscopic frame or the crystal frame can be used. The crystal frame is a frame which does not rotate with respect to the macroscopic frame, but which at the instant considered is parallel with a set of orthogonal crystal axes. \( \mathbf{W}^L \) is the spin of the crystal lattice; \( \dot{\gamma}_s \) is the slip rate on slip system \( s \); \( N \) is the total number of slip systems (f.c.c.: 12, i.e. \{111\}<110>; b.c.c.: 24, i.e. \{110\}+(112)<110>); \( \hat{\mathbf{b}}^s \) is the unit vector in slip direction of slip system \( s \) (proportional to the Burgers vector), and \( \mathbf{m}^s \) is the unit vector normal to the slip plane of slip system. In the full-constraints Taylor theory (see below), the local velocity gradient tensor \( \mathbf{l} \) is known whereas the slip rates \( \dot{\gamma}_s \) must be calculated. This will also be assumed here.

\( \mathbf{W}^L \) vanishes from Equation (22.1) if the symmetric part is taken:

\[
\dot{\mathbf{d}} = \sum_{s=1}^{N} \mathbf{M}^s \dot{\gamma}_s
\]  

(22.2)

in which

\[
\mathbf{M}^s = \frac{1}{2} \left( \mathbf{b}^s \otimes \mathbf{m}^s + \mathbf{m}^s \otimes \mathbf{b}^s \right) \quad \text{(no summation over } s) \]  

(22.3)

The local strain rate tensor \( \dot{\mathbf{d}} \) is the symmetrical part of \( \mathbf{l} \). At first sight, both sides of Equation (22.2) have 6 independent tensor components. However, one of these is a linear
combination of two other ones. The sums of the left-hand sides as well as of the right-hand sides is zero for \(ij = 11, 22, 33\) because

- left hand sides: it is assumed that the volume does not change;
- right-hand sides: the vectors \(\hat{b}_s^m\) and \(m^s\) are orthogonal.

As a result, Equation (22.2) only represents 5 independent scalar equations.

On the one hand, \(l\) and \(d\) are prescribed, whereas on the other hand, it is not known in advance which slip systems will be active. That means that all \(\hat{\gamma}_s\) must be treated as unknowns. Since there are more unknowns than equations, the solution of the equations will not be unique. The theory of linear algebra then teaches that there will be a large number of different sets of 5 non-zero \(\hat{\gamma}_s\)-values which all are solutions, as well as all their linear combinations. It is then clear that an additional assumption is needed in order to restrict the number of solutions. This additional assumption is called here the energetic assumption. There are several variants of it. Taylor (1938) himself assumed that, of all possible solutions, the best is the one for which the rate of internally dissipated frictional work is minimal:

\[
P^* = \sum_{s=1}^{N} \tau^c_s |\hat{\gamma}_s| = \text{Min}
\]  

(22.4)

in which \(\tau^c_s\) is the critical resolved shear stress on slip system \(s\). The \(\tau^c_s\) are directional, i.e. their value may depend on the direction of slip on the slip system \(s\). Van Houtte (1988) describes how Equations (22.2)–(22.3) can be solved by means of linear programming, for which it is appropriate to use the crystal frame for the representation of the tensors in Equations (22.2)–(22.3), because the tensors \(M^s\) always have the same representation in that frame, regardless of the crystal orientation \(g\). This, of course, does not apply to the strain rate tensor \(d\). In principle, the solution of Equations (22.2), (22.4) will be a \(\hat{\gamma}_s\)-matrix (in an \(N\)-dimensional space) containing 5 non-zero slip rates and \(N-5\) zero slip rates. It makes \(P^*\) minimal. Once the slip rates have been found, the kinematical Equation (22.2) is used to find \(W^L\), the spin of the crystal lattice, which then in turn is used to estimate the new orientation of the crystallite after a finite strain step.

As explained in any textbook on metal physics, the generalized Schmid law establishes the relation between the resolved shear stresses \(\tau^r_s\), defined as:

\[
\tau^r_s = M^s : \sigma' = \sum_i \sum_j \frac{1}{2} (\hat{b}_s^i m^s_j + \hat{b}_s^j m^s_i) \sigma'_{ij}
\]

(22.5)

and the critical resolved shear stresses \(\tau^c_s\), both for a given slip system. \(\sigma'\) is the deviatoric stress tensor. For a given time \(t\), the \(\tau^c_s\) are material constants. In general, the law says that \(|\tau^r_s| \leq \tau^c_s\). For the active slip systems, identified by the 5 non-zero \(\hat{\gamma}_s\)-values, this becomes:

\[
\sum_i \sum_j \frac{1}{2} (\hat{b}_s^i m^s_j + \hat{b}_s^j m^s_i) \sigma'_{ij} = \text{sign}(\hat{\gamma}_s) \tau^c_s
\]

(22.6)

Equation (22.4) represents a set of 5 linear equations from which the 5 independent components of the deviatoric stress \(\sigma'\) can be obtained. It can be proved (Van Houtte, 1988) that after
the solution of Equations (22.2), (22.4), \( P^* = \sigma : d \), the rate of plastic work per unit volume. The Taylor factor \( M \) can now be introduced. It is often used to characterize the plastic anisotropy of a crystal:

\[
M = \frac{P}{\tau_c d_{M}}
\]  

(22.7)

\( \tau_c \) is a reference value for the critical resolved stress (if they are all the same, then it is that value). \( d_{M} \) is the von Mises equivalent strain rate. \( M \) depends on the crystal orientation and on the strain mode, i.e. the direction of \( d \) in strain rate space. It does not depend on the magnitude of \( d \). A typical application of the Taylor factor is the calculation of flow stresses: the flow stress in a tensile test is for example equal to \( Mg\tau_c \), using the Taylor factor associated to the strain rate tensor of that test.

Some alternative formulations of the energetic assumptions have also been proposed. One is due to Bishop and Hill (1951a-b). They use the generalized Schmid law to construct a yield locus for each crystallite (often referred to as the “Bishop-Hill yield locus”) and then use the Maximum Work Principle from plasticity theory in order to find the deviatoric stress which corresponds to the prescribed strain rate. Since such yield locus has the aspect of a polyhedron in stress space, the stresses found are nearly always vertices of it. Each vertex corresponds to a certain number of active slip systems, which then makes it easy to identify the active slip systems. Bishop and Hill (1951a-b) have shown that their method is equivalent to one based on Taylor’s (1938) assumption Equation (22.4). As a result, solutions and problems are identical. Therefore, the Taylor theory and the Bishop-Hill theory are often called by a single name: the (rate insensitive) Taylor-Bishop-Hill theory.

It is possible that less than 5 slip systems are activated in the solution of this theory. The stress is then not located at one of the vertices of the yield locus, but at some intermediate point between adjacent vertices.

A third form of the energetic equation is offered by the so-called visco-plastic model (see e.g. Toth et al., 1988). The method is explained in Chapter 5 and 23.

### 22.3 The Taylor Ambiguity

The rate insensitive Taylor-Bishop-Hill model (based on Equations (22.2)–(22.4)) may lead to multiple solutions for the slip rates and the lattice spin (but not for the stress or the Taylor factor) in the following cases:

- the same values are used for the critical resolved shear stresses \( \tau_{c}^{s} \) on many or all slip systems;
- at least 5 slip systems are simultaneously activated.

As a result, the problem is not very relevant for non-cubic materials, because different \( \tau_{c}^{s} \) values are then assigned to different slip systems. It is also not very relevant for RC Taylor models (see below), because less than 5 slip systems are active in those cases.

The reason why the same \( \tau_{c}^{s} \) values are used for all potential slip systems is mainly because experimental data or theoretical methods to differentiate between them are often not available.
However, there are various physical phenomena which would cause differentiation between the \( \tau_c \). One of them is strain rate sensitivity; it is used by the visco-plastic model to choose a solution among all possible solutions of the Taylor equations. Another one is the effect of the dislocation substructure, which develops during plastic deformation. Peeters et al. (2001) have quantitatively compared the variabilities of the critical resolved shear stresses caused by these two effects, using a model for the development of dislocation substructures in polycrystalline IF steels. It was found that beyond 1% plastic strain, the effect of the strain rate sensitivity is completely overshadowed by that of the dislocation patterns. In general, the dominant effect which controls the choice of the “true” solution among all possible solutions of the Taylor model may very well depend on the material, the temperature range etc. In cases for which there is no sound information about the physics of the problem, the arbitrary choice offered by the visco-plastic model is acceptable.

### 22.4 Full Constraints (FC) Taylor Theory

So far, only the mesoscopic constitutive model has been discussed. A method must also be provided to achieve the transition from the mesoscale to the macroscale. Indeed, the boundary conditions are in fact known for the body as a whole, not for the individual volume elements. For example, in the case of rolling, the deformation at the central plane is described by a velocity gradient tensor \( L \) of which all components (in the macroscopic frame) are zero except:

\[
L_{11} = -L_{33} = D_0
\]

in which \( D_0 \) is a scalar measure of the strain rate. Taylor (1938) formulated a second hypothesis (the first one is given by Equation 22.4) which is regarded as the primary model for the transition from mesoscale to macroscale. He assumed that the plastic strain of all grains of a polycrystal is equal to the macroscopic plastic strain. Elastic strains are neglected. This theory is also called the ‘Full Constraints (FC) theory’, as it maximizes the influence of the geometrical constraints. The theory is implemented by requiring that:

\[
\mathbf{l} = \mathbf{L}
\]

in which \( \mathbf{l} \) is the local velocity gradient used in Equation (22.2). In some applications, a sequence of macroscopic velocity gradient tensors \( \mathbf{L} \) are retrieved from the results of a macroscopic finite element simulation of a forming process. Once \( \mathbf{l} \) is known, a local constitutive law (Section 22.2) is used for further elaboration.

The FC model offers a solution for the distribution of stresses and strains throughout the polycrystal for which

- geometrical compatibility is assured (there are no misfits at grain boundaries);
- stress equilibrium is normally not achieved at grain boundaries.

It is clear that this solution can at most be approximative. In a true polycrystal, there are even stress and strain gradients inside the grains. Nevertheless, texture-induced anisotropy is sufficiently captured by this solution for most applications; however, deformation texture evolution is not.
22.5 Classical Relaxed Constraints (RC) Models

About 20 years ago, it was felt that the FC theory was too strict, and that the quality of deformation texture predictions might be improved by somewhat “relaxing” the geometrical constraints (Honeff and Mecking, 1981; Kocks and Chandra, 1982; Van Houtte, 1982). This idea is in part based on the observation, that after a certain amount of rolling, grains tend to become flattened and elongated. It then seemed that a misfit caused by a difference between for example $l_{13}$ and $L_{13}$ (Figure 22.1) would not be dramatic.

A variant of the Taylor theory could then be conceived, in which Equation (22.9) would no longer be fully enforced, but the equation for $l_{13}$ would be omitted. Equation (22.9) would no longer represent 9 independent scalar equations (using the macroscopic frame), but only 8. This means that for each individual crystallite, $l_{13}$ would be unknown: As a result, $d_{13}$ would also be unknown, since

$$d_{13} = l_{13} + l_{31}/2$$  \hspace{1cm} (22.10)

(Note that $l_{31}$ can still be identified with the prescribed $L_{31}$). Suppose that the macroscopic frame is used to express Equation (22.3). The equation for $d_{13}$ can then also be omitted. Equation (22.3) has now become a set of 4 linear equations. Taylor’s assumption that the rate of internally dissipated frictional work must be minimized (Equation (22.4)) is maintained. The 4 remaining Equations (22.3) and Equation (22.4) can again be solved by linear programming. This model is called the lath model; it is one of the so-called “Relaxed Constraints” (RC) models. 4 slip systems would be active according to this model. The pancake model also adopts this relaxation, but adds another one to it: the transverse shear $l_{23}$ is also relaxed. Only 3 equations are left of Equation (22.3) (when expressed in the macroscopic frame); the theory predicts the activation of 3 slip systems.

In practice, one avoids working in the macroscopic frame. This can be achieved as follows. First, Equation (22.9) is replaced by:

$$1 = L - \sum_{r=1}^{R} \tilde{K}_{rr}^{RLX}_{\gamma_r}$$  \hspace{1cm} (22.11)
The $\tilde{K}_r$ are velocity gradient tensors which describe ‘unit relaxations’ of the type $l_{13}$ and $l_{23}$ (pancake model). The $\dot{\gamma}_r^{RLX}$ are new unknowns which describe the additional degrees of freedom given by the relaxations. Equations (22.1)–(22.6) can then be solved for this choice of $l$. The unknowns $\dot{\gamma}_r^{RLX}$ are included in the minimization of $P^*$ which is solved by linear programming (Van Houtte, 1988). The application of this technique also leads to conclusions about the stress:

$$\sigma'_{13} = 0 \quad \text{and} \quad \sigma'_{23} = 0$$

(22.12)

(expressed in the macroscopic frame). Note also that these stress conditions apply to all grains of the polycrystal. Hence the stress equilibrium is satisfied at grain boundaries as far as the stress components $13$ and/or $23$ are concerned. *Introducing relaxed constraints creates strain misfits at grain boundaries, but improves the satisfaction of the stress equilibrium.*

The FC Taylor theory and the classical RC models described in the present section treat each crystallite separately. Therefore their implementation is rather simple. However, they have some theoretical shortcomings. Unlike the FC theory, the RC models cannot be used for any deformation mode. They have been invented for materials with flat or elongated grains, and relaxation modes such as $l_{13}$ in Figure 22.1 should leave the plane with which the flat grains are parallel unrotated and undistorted (or, in the case of elongated grains, they should leave the axis of elongation unrotated and undistorted). In addition, the relaxation modes should not relax non-zero components of the prescribed macroscopic velocity gradient $L$. This makes it for example impossible to use the pancake model for the simulation of plane strain with superimposed shear, as may occur close to the surface of the sheet during the rolling process (Engler et al., 2000).

Self-consistent models (Chapter 23) may be regarded as generalized RC models which determine the relaxation on the basis of a mathematical model for the interaction grain/matrix (=surroundings). The latter is treated as a homogenous medium. Thus, in contrast to the Taylor type models, in these models each grain deforms differently, depending on its relative hardness with respect to the matrix and on the grain shape.

### 22.6 Multi-grain RC Models

#### 22.6.1 Introduction

The deformation texture predictions of the RC models are not really much better than those of the FC model (especially not if quantitative methods were used for the comparison with experimental textures). It was suspected that this was due to the neglect of local interactions between specific neighboring grains. Therefore, some multi-grain, or $N$-site, models have been proposed. The ultimate multi-grain model is of course the CPFEM model, which is discussed in Chapter 26 and 28. In the present section, we will address the Lamel Model and the Advanced Lamel Model, which look at 2 crystallites, and the GIA Model, which looks at 8 crystallites.
22.6.2 The Lamel Model

The Lamel Model was also designed for the rolling process and for a material with flat, elongated grains (Van Houtte et al., 1999, 2002). It treats two grains simultaneously. It is supposed that they form a stack of 2 brick-shaped grains (Figure 22.2). First, the ODF of the polycrystal is discretized in several thousands of orientations, e.g. using the ‘statistical method’ described by Toth and Van Houtte (1992). Two of the resulting crystallite orientations are then chosen at random from the discretized set and assigned to the top and bottom grain, respectively. The normal relaxations considered are Type I and Type II shown by Figure 22.2. They leave the rolling plane undistorted and unrotated. No grain boundary sliding is allowed at the interface between the two grains. In some cases, the Type III relaxation shown in Figure 22.2 is also used; however, it is of a different nature as the Type I and Type II relaxations because:

- the corresponding shears in the two grains are equal, instead of being opposite;
- it does not leave the interface undistorted.

Therefore, it will no longer be discussed here, although some results will be mentioned in Section 22.8.

In principle, this model is implemented mathematically by setting up the equations of the pancake RC model (Section 22.6) for the top and the bottom grain. The two sets of equations are however coupled in two ways:

- the $\dot{\gamma}_{RLX}^{l_{13}}$ associated with the $l_{13}$ relaxation in the two grains are forced to be each other’s opposite;
- the same is done with the $l_{23}$ relaxation;
- $P^*$ (Equation (22.4)) is the sum of the virtual work of the two grains taken together.

![Figure 22.2: Illustration of the three types of relaxation considered in the Lamel and the Lamel + Type III models. Type I corresponds to RD-ND shear, Type II to TD-ND shear and Type III to RD-TD shear.](image-url)
It can be shown that the minimization techniques will also assure the satisfaction of the following stress conditions (Van Houtte et al., 1999):

\[ \sigma'_{13} = \sigma'_{23} \quad \text{and} \quad \sigma'_{12} = \sigma'_{21} \]  

(22.13)

(macroscopic frame) which means that, as far as the shear stresses \( \sigma'_{13} \) and \( \sigma'_{23} \) are concerned, equilibrium is assured at the interface, however without requiring that these shear stresses would be zero, as is done by the classical RC theories.

### 22.6.3 The Advanced Lamel Model

The Lamel model has been quite successful in quantitatively predicting rolling textures, especially for steel (Van Houtte et al., 1999, 2002). However, the same theoretical objection applies as against the classical RC models: it is in principle only valid for rolling of materials with flat, elongated grains, because it has been designed for that case. To overcome that restriction, a new theory has been proposed: the Advanced Lamel Model (Alamel) (Van Houtte, 2003). It is based on the following assumptions:

(i) The local strain rate \( \dot{l} \) is not constant throughout a crystallite such as grain \( \alpha \) in Figure 22.3; as a result, the slip rates are not constant either.

(ii) In regions close to a grain boundary (such as Region 1 in Figure 22.3), the local stress and strain rate (and slip rates) are heavily affected by the interaction with the neighboring grain;

(iii) the set of crystallographic orientations which develop in the various regions close to the grain boundaries (region 1 and similar regions along grain boundaries BC, CD, DE, EF and FA in Figure 22.3) can be regarded as the contribution of the grain to the overall deformation texture of the polycrystal.

It is clear that these assumptions (especially (iii)) may still be subject to debate. At the time being, the judgment about the validity of this assumption is based upon the comparison of predicted deformation textures with experimentally observed ones.

![Figure 22.3: Schematic representation of a microstructure.](image)
The evolution of the crystallite orientations in regions 1 and 2 (Figure 22.3) can be obtained with the same equations as those of the Lamel model. Note that Equation (22.13) is now valid in the reference frame $y_1y_2y_3$ (Figure 22.3), not in the macroscopic frame.

In practice, the Advanced Lamel Model is implemented as follows:

- as before, the ODF of the initial texture is discretized in several thousands orientations;

- a *microstructure file* describing the initial microstructure is read. It consists of a large number of records which describe interfaces such as AB in Figure 22.3. This description essentially consists of the surface of the interface and the Euler angles of the frame $y_1y_2y_3$ with respect to the macroscopic frame. Normally, this file is initialized to represent an equiaxed microstructure in three dimensions. Non-equiaxed versions can be derived from it by applying a user-defined “pre-deformation” to the elements in this file. This (real or fictitious) pre-deformation should be described by a deformation gradient tensor $F$, which has to be specified by the user (deformation gradient $F$: let $x$ be a vector in the material before the deformation, and $y$ after. Then $y = F \cdot x$).

- a large number of pairs “region 1 + region 2” (Figure 22.3) are now created by random selection of an interface from the microstructure file, and (for each region) a random choice of a crystallite orientation from the discretized ODF.

- The slip rates in each “region 1 + region 2” are then calculated as described above, leading to the lattice spins $W^L$ in each region.

- The crystal orientations of the two regions at the beginning of the next strain increment are calculated by explicit integration, using the $W^L$.

- The interfaces between the regions (orientation, surface) are updated by first updating the deformation gradient tensor $\dot{F}$ (see above) by explicit integration of the classical equation over the increment:

$$\dot{F} = L \cdot F$$

(22.14)

in which $L$ is the macroscopic velocity gradient at the beginning of the increment. The updated tensor $F$ is then used to obtain new values for the orientation and the surface of the interface at the beginning of the next increment.

- At the end of the simulation of a large deformation, the deformation texture is assembled by taking the final crystallite orientations of all regions of all pairs, using the surfaces as weighting factors.

As opposed to the Lamel Model or the classical RC models, the geometrical nature of the relaxations is not related to the geometry of the prescribed deformation. In the case of rolling, this means that the shear plane of the two relaxations need not be parallel with the rolling plane. This makes the Advanced Lamel Model applicable to any deformation mode. However, the relaxations do sometimes relax non-zero components of the prescribed macroscopic strain rate tensor (when it is expressed in the $y_1y_2y_3$ frame). It was found (Van Houtte et al., 2003) that this can be tolerated, on condition that a single-crystal hardening law is used (for every region separately) during the simulation.
22.6.4 The Grain Interaction (GIA) Model

Another multi-grain RC model is the Grain Interaction Model (GIA) proposed by Crumbach et al. (2001). As the Lamel model, it takes interactions between neighboring grains into account, but instead of operating on a stack of 2 grains, it works with a cluster of 8 grains. As in the Lamel model, the cluster as a whole much satisfy the Taylor condition: the average strain of the cluster must be equal to the macroscopic strain. As long as this condition is satisfied, the grains within the cluster are allowed to relax; however, these relaxations are ‘penalized’ by an energy term. The final solution is obtained by minimizing the total work: work dissipated by the slip systems augmented by the penalty term (Crumbach et al., 2001). Calculation time is considerably higher as for the other models mentioned so far.

22.7 Validation of the Models

The results of the models will be quantitatively compared to experimentally observed deformation textures, and to the results of CPFEM models. The texture indices $I(\Delta f)$ of the difference ODFs will be used for the comparisons. Such texture index is the integral of the square of the difference between the Orientation Distribution Function (ODF) of the experimental rolling texture and the one simulated by a model:

$$I(\Delta f) = \int [f_{\text{model}}(g) - f_{\text{exp}}(g)]^2 \, dg$$

(22.15)

$g$ is a crystal orientation. The integrals are taken over entire orientation space. The value of $I(\Delta f)$ is zero if the two ODFs are identical. If desired, $I(\Delta f)$ can be normalized by dividing it by the texture index $I_{\text{exp}}$ (integral of the square of the ODF) of the experimental texture. The results of deformation texture simulations are sets of weighted discrete orientations. They are converted into continuous ODF’s by putting a Gaussian distribution of the type $\exp(\psi/\psi_0)^2$ with a spread of $\psi_0 = 7^\circ$ upon each of the orientations. Table 22.1 shows results obtained for an IF steel after 70% rolling. Some of these results have been described elsewhere in much greater detail (Van Houtte et al, 1999, 2002). It should be pointed out that prior to cold rolling, the IF steel was hot rolled in the ferritic range. As a result, the as-hot rolled microstructure was not equiaxed. It was attempted to take that into account in one of the two Alamel-simulations.

Table 22.2 shows results obtained in the framework of an European project. Some results by the VPSC model, a self-consistent model (Lebensohn et al. 1993, see also Chapter 23), have also been included. More details on most of these results have been given by Li and Van Houtte (2002).

22.8 Conclusions

The classical statistical Taylor-type models (Taylor FC and RC) are comparatively simple to implement in software, and do not require much calculation time. Although the FC model has some merit in qualitatively prediction deformation textures for a wide range of deformation processes and materials, its predictions are quantitatively incorrect. The same holds for the
Table 22.1: Texture index of difference ODFs for IF steel results. 70 % rolling reduction. $I_{\text{exp}} = 6.7$.

<table>
<thead>
<tr>
<th>Model</th>
<th>CPFEM$^1$</th>
<th>CPFEM$^2$</th>
<th>FC</th>
<th>RCP</th>
<th>GIA</th>
<th>Lamel$^1$</th>
<th>Alamel$^1$</th>
<th>Alamel$^{1,2}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1.22</td>
<td>0.90</td>
<td>4.65</td>
<td>1.40</td>
<td>0.76</td>
<td>0.59</td>
<td>1.08</td>
<td>0.65</td>
</tr>
</tbody>
</table>

CPFEM$^1$: developed by Bate (1999); CPFEM$^2$: developed by Kalidindi et al. (1992); FC = FC Taylor theory; RCP = Pancake variant of the RC Taylor theory; GIA: developed by Crumbach et al. (2001); Lamel$^1$: standard Lamel model; Alamel$^1$: Advanced Lamel Model, starting from equiaxed grains; Alamel$^{1,2}$: starting from elongated grains (as if the grains of the hot rolled microstructure already had 40% rolling reduction) and using additional $\{123\}<111>$ slip systems (see also Raabe, 1995a-b).

Table 22.2: Texture index of difference ODFs for 2 aluminum alloys.

<table>
<thead>
<tr>
<th>Material</th>
<th>AA1200</th>
<th>AA5182</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reduction</td>
<td>40%</td>
<td>63%</td>
</tr>
<tr>
<td>FC</td>
<td>0.54</td>
<td>1.43</td>
</tr>
<tr>
<td>RCP</td>
<td>4.25</td>
<td>4.03</td>
</tr>
<tr>
<td>CPFEM$^2$</td>
<td>0.42</td>
<td>0.92</td>
</tr>
<tr>
<td>VPSC</td>
<td>1.11</td>
<td>2.02</td>
</tr>
<tr>
<td>GIA</td>
<td>0.57</td>
<td>0.87</td>
</tr>
<tr>
<td>Lamel$^1$</td>
<td>0.49</td>
<td>0.99</td>
</tr>
<tr>
<td>Lamel$^2$</td>
<td>0.49</td>
<td>1.06</td>
</tr>
<tr>
<td>$I_{\text{exp}}$</td>
<td>2.50</td>
<td>2.45</td>
</tr>
</tbody>
</table>

Models: see Table I, except: Lamel$^{1,2}$: Lamel with Type I, II and III-relaxation (Figure 22.2). VPSC: Visco-Plastic Self-Consistent Model (Lebensohn et al., 1993)
classical Taylor RC models. These can in principle only be used for rolling, whereas the FC model can be used for a wide range of deformation processes.

For aluminum alloys, a visco-plastic self-consistent model has also been tested. For AA1200, it performed less well than the Taylor FC model; for AA5182, its results were slightly better.

It has been proved that a much better quantitative agreement can be obtained if models are used which take local interactions between particular grains into account. Crystal-plasticity finite element models (CPFEM) are then the obvious choice, and they do indeed lead to much better texture predictions. However, their results are not perfect either. Besides, these models require huge computational efforts. This is not the case for a new generation of Taylor-type statistical models such as the Lamel model, the Advanced Lamel model and the GIA model. They try to take short-range grain interactions into account. Of these, only the GIA and the Advanced Lamel models are in principle applicable to a wide range of deformation modes. Their results (as shown above) are even better than those of the CPFEM model, which probably demonstrates that a much finer mesh should be used by the latter, leading to even higher calculation times.

Acknowledgements

The authors gratefully acknowledge financial support by the European Union through the VIRFORM-project. They thank P. Bate, M. Crumbach and G. Gottstein for help with the simulations and for many valuable discussions.

References


23 Self Consistent Homogenization Methods for Texture and Anisotropy

Carlos Norberto Tomé and Ricardo Anibal Lebensohn

23.1 Introduction

In this chapter we are concerned with the heterogeneity and the anisotropy of the plastic response of single-phase polycrystalline aggregates. In single-phase aggregates the properties of the constituent domains (grains) are identical with respect to the crystallographic reference system. Heterogeneity arises from the presence of different grain orientations, since grain properties are different when expressed in an ‘absolute’ reference system. Polycrystalline aggregates represent a paradigm of texture and anisotropy being intertwined, because the combination of the grain orientation distribution and the single crystal anisotropy determines the anisotropy of the overall response. The reader is referred to the book of Kocks et al. (1998) for a comprehensive treatment of the subject.

The usual requirement of polycrystal models is that the loading and displacement conditions at the boundary of the polycrystal are uniform, and that the volume average of stress, strain and strain rate over all grains must coincide with the overall stress, strain and strain rate at the boundary (Hill 1967). Since the properties and the response of the aggregate are given by weighted averages over the orientation distribution, anisotropy is a consequence of accounting for the texture, and does not have to be introduced \textit{a priori}. There are, however, several configurations of stress and strain in the grains that fulfill the boundary conditions, and the difference between polycrystal models is the assumption that they use to localize the deformation in the grains.

The Taylor assumption of uniform strain throughout the polycrystal ensures compatibility and provides an upper bound to the stress. The so called Sachs model, on the other hand, assumes proportional (not uniform) loading in the grains, and satisfies neither compatibility nor stress equilibrium. The Taylor or Full Constraints (FC) model has been favored by materials scientists because it leads to multiple slip in the grains and to reasonable texture predictions. Because of its characteristics, it has been mainly used for modeling cubic materials, which are in general mildly anisotropic plastically. Enforcing local compatibility over local equilibrium, though, may be a severe assumption in some situations. The Relaxed Constraints (RC) approach of Honneff and Mecking (1981) imposes a mixture of both conditions based on grain shape considerations, and has proven a better alternative for the prediction of rolling and torsion textures at large strains. Both the FC and RC approaches are reviewed by P. Van Houtte in Chapter 22 of this book.
In the case of low-symmetry materials, plasticity is characterized by a variety of active deformation modes present in each grain (in many cases including twinning), and markedly anisotropic single crystals with ‘soft’ and ‘hard’ orientations, which induce highly directional grain interactions. These crystals cannot accommodate certain deformation components because they either lack the necessary deformation systems or because those systems require high activation stresses. In such a case the stress built up may lead to fracture, unless it can be relaxed by inducing extra plastic accommodation in the neighboring grains. As a consequence, an approach more complex than the standard Taylor model is required in order to predict the plastic response of low symmetry materials.

An empirical way of tackling the scarcity of deformation systems without sacrificing the numerical and conceptual simplicity of the Taylor assumption is the Constrained Hybrid (CH) model of Parks and Ahzi (1990), which was originally formulated to model semi-crystalline polymers. The CH model consists in projecting the externally imposed strain along the directions (in strain space) where the grain can yield. The strain tensor applied to each grain is scaled, in order for the average to give the correct overall strain. The CH model reduces to the FC approach when the single crystal yield surface is closed in the 5-dim deviatoric stress space.

A more rigorous approach to the modeling of heterogeneous deformation of polycrystals is given by Self-Consistent (SC) and FEM methods. SC methods aim at deducing the overall response of the aggregate from the known properties of the constituent grains and an Effective Medium assumption to calculate the interaction of each grain with its surroundings. Because of their statistical nature they generally do not account for the precise topology of the aggregate: as all models, they fulfill compatibility and stress equilibrium in an average sense. The first SC polycrystal models addressing elastic, elastoplastic and viscoplastic responses, were proposed by Kröner (1961), Hill (1970) and Hutchinson (1976), respectively. FEM calculations, on the other hand, fulfill local compatibility and are better suited for describing localization effects. FEM simulations, however, fulfill local equilibrium only in a weak sense and, in addition, the results may be sensitive to the spatial grain distribution used. Self-consistent polycrystal models are the subject of this section, while FEM of crystal plasticity is addressed by S. Kalidindi and F. Roters in Chapters 26 and 28 of this book.

The main characteristic of self-consistent polycrystal models is that each grain is regarded as an inclusion embedded in the Homogeneous Effective Medium (HEM) represented by the surrounding grains. Each grain is assumed to be representative of all the grains with the same orientation, and the HEM represents the average neighborhood of all these grains. The interaction between the inclusion and the HEM is found by solving the local stress equilibrium equations. When the constitutive response of the inclusion or of the medium is non-linear, some form of linearization must be enforced. In addition, the properties of the HEM are not known a priori and follow from the requirement that grain averages have to be consistent with boundary conditions.

The plan of this chapter is as follows: Section 23.2 provides a comprehensive description of a SC polycrystal model associated with a viscoplastic response. The formulation explicitly accounts for the plastic anisotropy of grain and polycrystal, grain shape effects, and their evolution with deformation. The approach that we present here for addressing the problem of large plastic deformation of polycrystals neglects elastic effects. Section 3 describes how the SC formulation is implemented in a model of large strain plasticity that includes harden-
ing, texture evolution and the contribution of twinning to deformation, texture and hardening. Section 23.4 presents practical applications of the SC model for single-phase aggregates: we analyze the viscoplastic response of cubic, hexagonal and orthorhombic polycrystals, emphasizing the correct treatment of crystal anisotropy and analyzing its influence upon the response of the material. High symmetry materials (FCC) are treated with the techniques used for modeling more anisotropic materials, in order to assess the departure from the predictions made using the simpler FC approach.

23.2 Viscoplastic Selfconsistent Formalism

In what follows, we present the basic equations of the 1-site viscoplastic selfconsistent model, originally due to Molinari et al (1987) and extended to fully anisotropic behavior by Lebensohn and Tomé (1993). The present derivation is completely general, based on the fully incompressible formulation of Lebensohn et al. (1998) and the generalized affine linearization scheme of Masson et al. (2000) and of Lebensohn et al. (2003b). For a detailed analysis the reader is referred to Lebensohn et al. (2003a).

23.2.1 Local Constitutive Behavior and Homogenization

Let us consider a polycrystalline aggregate. The deviatoric part of the viscoplastic constitutive behavior at local level (in a given grain) is described by means of the non-linear rate-sensitivity equation:

\[ d_{ij}(\bar{x}) = \sum_s m^{s}_{ij} \dot{\gamma}^s(\bar{x}) = \dot{\gamma}_o \sum_s m^{s}_{ij} \left( \frac{m^{s}_{kl} \sigma'_{kl}(\bar{x})}{\tau^s} \right)^n \tag{23.1} \]

In the above expression \( \tau^s \) and \( m^{s}_{ij} = \frac{1}{2} (n^s_i b^s_j + n^s_j b^s_i) \) are the threshold stress and the symmetric Schmid tensor associated with slip (or twinning) system (s), where \( n^s \) and \( b^s \) are the normal and Burgers vector of such slip (or twinning) system, \( d_{ij}(\bar{x}) \) and \( \sigma'_{kl}(\bar{x}) \) are the deviatoric strain-rate and stress, and \( \dot{\gamma}^s(\bar{x}) \) is the local shear-rate on slip system (s), which can be obtained as:

\[ \dot{\gamma}^s(\bar{x}) = \dot{\gamma}_o \left( \frac{m^{s}_{kl} \sigma'_{kl}(\bar{x})}{\tau^s} \right)^n \tag{23.2} \]

where \( \dot{\gamma}_o \) is a normalization factor and \( n \) is the rate-sensitivity exponent. Linearizing Equation (23.1) inside the domain of a grain gives:

\[ d_{ij}(\bar{x}) = M_{ijkl} \sigma'_{kl}(\bar{x}) + d^o_{ij} \tag{23.3} \]

where \( M_{ijkl} \) and \( d^o_{ij} \) are the viscoplastic compliance and the back extrapolated term of the grain, respectively. Same relation holds for the average strain-rate and stress in the grain:

\[ d_{ij} = M_{ijkl} \sigma'_{kl} + d^o_{ij} \tag{23.4} \]
23.2 Viscoplastic Selfconsistent Formalism

Depending on the linearization assumption, \( M_{ijkl} \) and \( d_{ij} \) can be chosen differently. For instance, if the back-extrapolated term \( d_{ij} \) is a priori set to zero, the resulting model is a secant one, and the secant modulus results:

\[
M_{ijkl}^{\text{sec}} = \dot{\gamma}_o \sum_s \frac{m_{ij}^s m_{kl}^s}{\tau_s} \left( \frac{m_{pq}^s \sigma'^{pq}_s}{\tau_s} \right)^{n-1}
\]  
(23.5)

The secant approximation has been proved to be too stiff, leading to close-to-upper-bound results. On the other hand, taking \( M_{ijkl} = \partial d_{ij} / \partial \sigma'^{kl} \), the model is a tangent one (Molinari et al., 1987; Lebensohn and Tomé, 1993), which gives a much more compliant response. Explicitly, the tangent modulus and back-extrapolated term are:

\[
M_{ijkl}^{\text{tg}} = n \dot{\gamma}_o \sum_s \frac{m_{ij}^s m_{kl}^s}{\tau_s} \left( \frac{m_{pq}^s \sigma'^{pq}_s}{\tau_s} \right)^{n-1} = nM_{ijkl}^{\text{sec}}
\]  
(23.6a)

\[
d_{ij} = \left( M_{ijkl}^{\text{sec}} - M_{ijkl}^{\text{tg}} \right) \sigma'^{kl} = (1 - n) \, d_{ij}
\]  
(23.6b)

Performing homogenization on this heterogeneous medium consists in assuming pseudo-linear constitutive relations analogous to (23.3) at the effective medium (polycrystal) level:

\[
D_{ij} = \bar{M}_{ijkl} \Sigma'^{kl} + D_{ij}^{\text{o}}
\]  
(23.7)

where \( D_{ij} \) and \( \Sigma'^{ij} \) are overall (macroscopic) magnitudes and \( \bar{M}_{ijkl} \) and \( D_{ij}^{\text{o}} \) are the macroscopic viscoplastic compliance and back-extrapolated term, respectively. The latter moduli are unknown a priori and need to be adjusted self-consistently. Invoking the concept of the equivalent inclusion (Mura 1987), the local constitutive behavior can be rewritten in terms of the homogeneous macroscopic moduli, so that the inhomogeneity is ‘hidden’ inside a fictitious eigen-strain-rate, as:

\[
d_{ij}(\bar{x}) = \bar{M}_{ijkl} \sigma'^{kl}(\bar{x}) + D_{ij}^{\text{e}} + d_{ij}^*(\bar{x})
\]  
(23.8)

\( d_{ij}^*(\bar{x}) \) is the eigen-strain-rate field, which follows from replacing the inhomogeneity by an equivalent inclusion. Rearranging and subtracting (23.7) from (23.8) gives:

\[
\bar{\sigma}_{ij}(\bar{x}) = \bar{L}_{ijkl} \left( \bar{d}_{kl}(\bar{x}) - d_{ij}^*(\bar{x}) \right)
\]  
(23.9)

The symbol “\( \sim \)” denotes local deviations of the corresponding tensor from macroscopic values and \( \bar{L}_{ijkl} = \bar{M}_{ijkl}^{-1} \). Combining (23.9) with the equilibrium condition:

\[
\sigma_{ij,j}(\bar{x}) = \bar{\sigma}_{ij,j}(\bar{x}) = \bar{\sigma}'_{ij,j}(\bar{x}) + \bar{\sigma}^m_{ij}(\bar{x})
\]  
(23.10)

using the relation between strain-rate and velocity-gradient, i.e.:

\[
\bar{d}_{ij}(\bar{x}) = \frac{1}{2} \left( \tilde{u}_{ij,j}(\bar{x}) + \tilde{u}_{j,i}(\bar{x}) \right),
\]

and adding the incompressibility condition, we obtain:

\[
\bar{L}_{ijkl} \tilde{u}_{k,ij}(\bar{x}) + \bar{\sigma}^m_{ij}(\bar{x}) + f_i(\bar{x}) = 0
\]  
(23.11a)

\[
\tilde{u}_{k,k}(\bar{x}) = 0
\]  
(23.11b)
where the fictitious volume force associated with the heterogeneity is:

\[ f_i(\bar{x}) = -L_{ijkl}d_{kl,j}^*(\bar{x}) = \sigma_{ij}^*(\bar{x}) \]  

(23.12)

The field \( \sigma_{ij}^*(\bar{x}) = -L_{ijkl}d_{kl,j}^*(\bar{x}) \) defined in (23.12) will be called in what follows eigen-stress field.

### 23.2.2 Green Function Method and Fourier Transform Solution

System (23.11) consists of four differential equations with four unknowns: three are the components of velocity deviation vector \( \tilde{\dot{\mathbf{u}}}_i(\bar{x}) \), and one is the mean stress deviation \( \tilde{\sigma}_m(\bar{x}) \). A system of \( N \) linear differential equations with \( N \) unknown functions and an inhomogeneity term, such as (23.11), can be solved using the Green function method, as explained in what follows. Let us call \( G_{km}(\bar{x}) \) and \( H_{m}(\bar{x}) \) the Green functions associated with \( \tilde{\dot{\mathbf{u}}}_i(\bar{x}) \) and \( \tilde{\sigma}_m(\bar{x}) \), which solve the auxiliary problem of a unit volume force, with a single non-vanishing \( m \)-component, and applied at \( \bar{x} = 0 \):

\[
\begin{align*}
\bar{L}_{ijkl} G_{km,lj}(\bar{x}) + H_{m,i}(\bar{x}) + \delta_{im} \delta(\bar{x}) &= 0 \\
G_{km,k}(\bar{x}) &= 0
\end{align*}
\]  

(23.13)

Here \( \delta(\bar{x}) \) is Dirac’s delta function and \( \delta_{im} \) is the Kronecker delta. Once the solution of (23.13) is obtained, the solution of (23.11) is given by the convolution integrals:

\[
\begin{align*}
\tilde{\dot{\mathbf{u}}}_k(\bar{x}) &= \int_{\mathbb{R}^3} G_{ki}(\bar{x} - \bar{x}') f_i(\bar{x}') \, d\bar{x}' \\
\tilde{\sigma}_m(\bar{x}) &= \int_{\mathbb{R}^3} H_i(\bar{x} - \bar{x}') f_i(\bar{x}') \, d\bar{x}'
\end{align*}
\]  

(23.14)

(23.15)

System (23.13) can be solved using the Fourier transform method (Lebensohn et al., 2003a). Expressing the Green functions in terms of their inverse Fourier transforms, the differential system (23.13) transforms into an algebraic system:

\[
\begin{align*}
\alpha_j \alpha_l \tilde{L}_{ijkl} k^2 \hat{G}_{km}(\hat{k}) + \alpha_i ik \hat{H}_m(\hat{k}) &= \delta_{im} \\
\alpha_k k^2 \hat{G}_{km}(\hat{k}) &= 0
\end{align*}
\]  

(23.16)

where \( k \) and \( \hat{\alpha} \) are the modulus and the unit vector associated with a point of Fourier space \( \hat{k} = k\hat{\alpha} \), respectively. Calling \( A_{ik}^* = \alpha_j \alpha_l \tilde{L}_{ijkl} \), system (23.16) can be expressed as a matrix
23.2 Viscoplastic Selfconsistent Formalism

The product \( A \times B = C \) where \( A \), \( B \) and \( C \) are matrices given by:

\[
\begin{pmatrix}
    k^2 \hat{G}_{11} & k^2 \hat{G}_{12} & k^2 \hat{G}_{13} \\
    k^2 \hat{G}_{21} & k^2 \hat{G}_{22} & k^2 \hat{G}_{23} \\
    k^2 \hat{G}_{31} & k^2 \hat{G}_{32} & k^2 \hat{G}_{33}
\end{pmatrix} = B
\]

where

\[
A = \begin{pmatrix}
    A_{11}^d & A_{12}^d & A_{13}^d & \alpha_1 \\
    A_{21}^d & A_{22}^d & A_{23}^d & \alpha_2 \\
    A_{31}^d & A_{32}^d & A_{33}^d & \alpha_3 \\
    \alpha_1 & \alpha_2 & \alpha_3 & 0
\end{pmatrix}
\]

The 4 \( \times \) 4 matrix \( A \) is real and symmetric. As a consequence, its inverse will also be real and symmetric. Using the explicit form of matrix \( C \), we can write the solution of (23.17) as:

\[
B = A^{-1} \times C
\]

Finally, comparing (23.17) and (23.18):

\[
k^2 \hat{G}_{ij} = A^{-1}_{ij}
\]

\[
\begin{pmatrix} \alpha \\ \alpha \\ \alpha \\ 0 \end{pmatrix} = \begin{pmatrix} k \hat{H}_1 \\ k \hat{H}_2 \\ k \hat{H}_3 \end{pmatrix}
\]

Since the components of \( A \) are real functions of \( \alpha_i \), so are the components of \( A^{-1} \), and so are \( k^2 \hat{G}_{ij} \) and \( ik \hat{H}_i \). This property leads to real integrals in the derivation that follows.

23.2.3 Viscoplastic Inclusion and Eshelby Tensors

Now that we have a solution for the Green tensors, we can write the solution of our eigen-strain-rate problem using the convolution integrals (23.14)–(23.15). Taking partial derivatives to Equation (23.14) we obtain:

\[
\tilde{\dot{u}}_{k,t}(\bar{x}) = \int_{\Omega} G_{ki,t}(\bar{x} - \bar{x}') f_i(\bar{x}') d\bar{x}'
\]

Replacing (23.12) in (23.21), recalling that \( \partial G_{ij}(\bar{x} - \bar{x}') / \partial \bar{x} = -\partial G_{ij}(\bar{x} - \bar{x}') / \partial \bar{x}' \), integrating by parts, and using the divergence theorem, we obtain:

\[
\tilde{\dot{u}}_{k,t}(\bar{x}) = \int_{\Omega} G_{ki,t}(\bar{x} - \bar{x}') \sigma^*_{ij}(\bar{x}') d\bar{x}'
\]

Equation (23.22) provides an exact implicit solution to the problem. Such solution requires to know the local dependence of the eigen-stress tensor. However, we know from the elastic
Eshelby inclusion formalism that if the eigen-strain is uniform over an ellipsoidal domain where the stiffness tensor is uniform, then the stress and the strain are constant over the domain of the inclusion. The latter suggests us to assume an eigen-stress of constant value (a priori unknown) within the volume $\Omega$ of the inclusion, and zero outside. This allows us to average the local field (23.22) over the domain $\Omega$ and obtain an average strain-rate inside the inclusion of the form:

$$\tilde{\dot{u}}_{k,l} = \left( -\frac{1}{\Omega} \int \int G_{ki,jl} (\bar{x} - \bar{x}') \ d\bar{x} \ d\bar{x}' \right) \tilde{L}_{ijmn} d^*_{mn}$$

(23.23)

where $\tilde{\dot{u}}_{k,l}$ and $d^*_{mn}$ have to be interpreted as average quantities inside the grain. Expressing the Green tensor in terms of the inverse Fourier transform and taking derivatives we obtain:

$$\tilde{\dot{u}}_{k,l} = \left( \frac{1}{8\pi^3} \int \int \alpha_j \alpha_l \left( k^2 \hat{G}_{ki} (k) \right) \exp \left[ -i k (\bar{x} - \bar{x}') \right] \ dkd\bar{x}d\bar{x}' \right) \tilde{L}_{ijmn} d^*_{mn}$$

(23.24)

Writing $d\bar{k}$ in spherical coordinates: $d\bar{k} = k^2 \sin \theta \ d\theta \ d\phi$ and using relation (23.19), the Green interaction tensor $T_{kij}$ can be expressed as:

$$T_{kij} = \frac{1}{8\pi^3} \int \int \alpha_j \alpha_l A_k^{-1} (\bar{\alpha}) \Lambda (\bar{\alpha}) \sin \theta \ d\theta \ d\phi$$

(23.25)

where $\theta$ and $\phi$ are the spherical coordinates of the Fourier unit vector $\bar{\alpha}$ and:

$$\Lambda (\bar{\alpha}) = \int \left( \int \int \Omega \exp \left[ -i k (\bar{x} - \bar{x}') \right] \ d\bar{x} \ d\bar{x}' \right) k^2 dk$$

(23.26)

Integration of (23.26) inside an ellipsoidal grain of radii $(a, b, c)$ is given by (Bervellier et al. 1987):

$$\Lambda (\bar{\alpha}) = \frac{8\pi^3}{3} \left( \frac{abc}{\rho (\bar{\alpha})} \right)^2$$

(23.27)

where $\rho (\bar{\alpha}) = \left[ (a\alpha_1)^2 + (b\alpha_2)^2 + (c\alpha_3)^2 \right]^{1/2}$. Replacing (23.27) in (23.25), the expression of $T_{kij}$ for an ellipsoidal grain results:

$$T_{kij} = \frac{abc}{4\pi} \int \int \alpha_j \alpha_l A_k^{-1} (\bar{\alpha}) \sin \theta \ d\theta \ d\phi$$

(23.28)
23.2 Viscoplastic Selfconsistent Formalism

The convolution integral over the Green tensor $\hat{H}(\bar{x})$ allows us to obtain an expression for the mean stress deviation $\tilde{\sigma}^m(\bar{x})$, which is the fourth unknown function in differential system (23.11). This way of computing the hydrostatic pressure field has been used by Lebensohn et al. (1998) in a particular application of VPSC, to make a transition from viscoplastic incompressible loading to elastic unloading.

Expression (23.28) has to be integrated numerically using, for instance, a Gauss-Legendre technique. The evaluation of the integrand requires us to invert the 4x4 linear system (23.17) for each integration point. The symmetric and skew-symmetric Eshelby tensors are defined as:

$$S_{ijkl} = \frac{1}{4} \left( T_{ijmn} + T^+_{jimn} T_{ijnm} + T_{jimn} \right) \bar{L}_{mnkl}$$  \hspace{1cm} (23.29)

$$\Pi_{ijkl} = \frac{1}{4} \left( T_{ijmn} - T_{jimn} + T_{ijnm} - T_{jinm} \right) \bar{L}_{mnkl}$$ \hspace{1cm} (23.30)

Taking symmetric and skew-symmetric components to (23.24) and using (23.29)–(23.30), we obtain the strain-rate and rotation-rate deviations in the ellipsoidal domain:

$$\tilde{d}_{ij} = S_{ijkl} d^*_{kl}$$  \hspace{1cm} (23.31)

$$\tilde{w}_{ij} = \Pi_{ijkl} d^*_{kl} = \Pi_{ijkl} S^{-1} \tilde{d}_{mn}$$  \hspace{1cm} (23.32)

23.2.4 Interaction and Localization Equations

Expressions similar to Equation (23.9), relating deviations with respect to overall quantities, also holds for the average stress, strain-rate, and eigen-strain-rates in the grains:

$$\tilde{\sigma}'_{ij} = L_{ijkl} \left( \tilde{d}_{kl} - d^*_{kl} \right)$$  \hspace{1cm} (23.33)

Replacing the eigen-strain-rate given by (23.31) into the deviation equation (23.33), we obtain the following interaction equation:

$$\tilde{d}_{ij} = -\tilde{M}_{ijkl} \tilde{\sigma}'_{kl}$$  \hspace{1cm} (23.34)

where the interaction tensor is given by:

$$\tilde{M}_{ijkl} = (I - S)_{ijkl}^{-1} S_{mnpq} \bar{M}_{pqkl}$$  \hspace{1cm} (23.35)

Replacing the local and overall deviatoric constitutive relations (23.3) and (23.7) into the interaction equation (23.35) we can write, after some manipulation, the following localization equation:

$$\sigma'_ij = B_{ijkl} \tilde{\sigma}'_{kl} + \Phi_{ij}$$  \hspace{1cm} (23.36)

where the localization tensors are defined as:

$$B_{ijkl} = \left( M + \tilde{M} \right)_{ijkl}^{-1} \left( M + \bar{M} \right)_{mnkl}$$  \hspace{1cm} (23.37)

$$\Phi_{ij} = \left( M + \tilde{M} \right)_{ijkl}^{-1} \left( D^*_{kl} - d^*_{kl} \right)$$  \hspace{1cm} (23.38)
23.2.5 Selfconsistent Equations

The derivation presented in the previous sections solves the problem of a viscoplastic incompressible inclusion embedded in a viscoplastic incompressible effective medium being subject to external loading conditions. In this section we are going to use the previous result to construct a polycrystal model, consisting in regarding each grain as an ellipsoidal inclusion embedded in an effective medium which represents the polycrystal. The properties of such medium are not known a priori but have to be found thorough an iterative self-consistent procedure. Replacing the stress localization equation (23.36) in the local constitutive equation (23.3) we obtain:

\[ d_{ij} = M_{ijkl} \sigma'_{kl} + d^o_{ij} = M_{ijkl} B_{klmn} \Sigma'_{mn} + M_{ijkl} \Phi_{kl} + d^o_{ij} \]  (23.39)

Enforcing the condition that the weighted average of the strain-rate over the aggregate has to coincide with the macroscopic quantities, i.e.:

\[ D_{ij} = \langle d_{ij} \rangle \]  (23.40)

In what follows the brackets “\( \langle \rangle \)” denote grain averages, weighted by the associated volume fraction. Using (23.39) and the macroscopic constitutive equation (23.7) we obtain:

\[ \bar{M}_{ijmn} \Sigma'_{mn} + D^o_{ij} = \langle M_{ijkl} B_{klmn} \rangle \Sigma'_{mn} + \langle M_{ijkl} \Phi_{kl} + d^o_{ij} \rangle \]  (23.41)

Equating the linear and independent terms leads to the following self-consistent equations for the homogeneous compliances and back-extrapolated term:

\[ \bar{M}_{ijmn} = \langle M_{ijkl} B_{klmn} \rangle \]  (23.42a)

\[ D^o_{ij} = \langle M_{ijkl} \Phi_{kl} + d^o_{ij} \rangle \]  (23.42b)

The self-consistent Equations (23.42) are derived imposing the average of the local strain-rates to coincide with the applied macroscopic strain-rate (Equation (23.40)). If the grain ellipsoids have the same shape and orientation, it can be shown that the same equations are obtained from the condition that the average of the local stresses coincides with the macroscopic stress. If the grains have each a different shape, they have associated different Eshelby tensors, and the interaction tensors cannot be factored from the averages. In this case, the following general self-consistent expressions should be used (Walpole 1969; Lebensohn et al. 1996; Lebensohn et al 2003a):

\[ \bar{M}_{ijkl} = \langle M : B \rangle_{ijkl} \langle B \rangle_{mnkl}^{-1} \]  (23.43a)

\[ D^o_{ij} = \langle M : \Phi + d^o \rangle_{ij} - \langle M : B \rangle_{ijkl} \langle B \rangle_{klmn}^{-1} \langle \Phi \rangle_{mn} \]  (23.43b)

The self-consistent Equations (23.42) are a particular case of (23.43). Both sets constitute fix-point equations that provide improved estimates of \( \bar{M}_{ijkl} \) and \( D^o_{ij} \), when they are solved iteratively starting from an initial guess. From a numerical point of view, Equations (23.43) are more robust and improve the speed and stability of the convergence procedure, even when solving a problem where all the inclusions have the same shape.
23.2 Viscoplastic Selfconsistent Formalism

23.2.6 Secant, Tangent and Intermediate Approximations

The selfconsistently adjusted macroscopic modulus $\bar{M}$ (Equation (23.7)) that results from a secant (Equation (23.5)) or tangent (Equation (23.6)) local linearization, can be called $\bar{M}_\text{sec}$ or $\bar{M}_\text{tg}$, respectively. If the rate sensitivity exponent $n$ is the same for all deformation systems, the above moduli are related by: $\bar{M}^{tg} = n\bar{M}^{sec}$ (Hutchinson 1976), i.e.: a relation analogous to the one that holds at local level (see Equation (23.6a)). The corresponding interaction tensors (Equation (23.41)) are:

$$\tilde{M} = (I - S)^{-1} : S : \bar{M}^{sec} \quad \text{(secant approximation)} \quad (23.44a)$$

$$\tilde{M} = (I - S)^{-1} : S : \bar{M}^{tg} = n (I - S)^{-1} : S : \bar{M}^{sec} \quad \text{(tangent approximation)} \quad (23.44b)$$

An inspection of the interaction equation (23.34) indicates that the smaller the compliance, the smaller is going to be the local deviation of the strain-rate with respect to the average. As a consequence, the secant interaction is stiff and gives results closer to Taylor. Intermediate approximations that give polycrystal’s responses in between the stiff secant and the compliant tangent approaches, can be obtained introducing an adjustable parameter $n^{eff}$, such that $1 < n^{eff} < n$. The interaction tensor is therefore given by:

$$\tilde{M} = n^{eff} (I - S)^{-1} : S : \bar{M}^{sec} = (n^{eff}/n) (I - S)^{-1} : S : \bar{M}^{tg} \quad (23.45)$$

23.2.7 Algorithm

To illustrate the use of this formulation, we describe here the steps required to predict the local and overall viscoplastic response of a polycrystal, for an applied macroscopic velocity gradient $\dot{U}_{ij} = D_{ij} + W_{ij}$ (decomposed here into the symmetric strain-rate $D_{ij}$ and the skew-symmetric rotation-rate $W_{ij}$). In order to start an iterative search of the local states, one should assume initial values for the local deviatoric stresses and moduli. Starting with an initial Taylor guess, i.e.: $d_{ij} = D_{ij}$ for all grains, the inversion of Equation (23.1) and the use of Equations (23.6) allow us to calculate initial values of $\sigma_{ij}',M_{ijkl}',\text{and } d_{ij}'$, respectively, for each grain. Next, initial guesses for the macroscopic moduli $\bar{M}_{ijkl}'$ and $D_{ij}'$ (usually simple averages of the corresponding local moduli) are obtained. With them, and the applied strain-rate $D_{ij}$, the initial guess for the macroscopic stress follows from the inversion of the macroscopic constitutive law (Equation (23.7)), while the Eshelby tensors $S_{ijmn}$ and $\Pi_{ijmn}$ can be calculated using the macroscopic moduli and the grain shape by means of the procedure described in Sections 23.2.2 and 23.2.3. Subsequently, the interaction tensor $\tilde{M}_{ijkl}'$ (Equation (23.34)), and the localization tensors $B_{ijkl}$ and $\Phi_{ij}$ (Equations (23.35) and (23.36)), can be obtained as well. With these tensors, new estimates of $\bar{M}_{ijkl}'$ and $D_{ij}'$ are obtained by solving iteratively the selfconsistent Equations (23.42) (for unique grain shape) or (23.43) (for a distribution of grain shapes). After achieving convergence on the macroscopic moduli (and, consequently, also on the macroscopic stress $\Sigma_{ij}'$ and the interaction tensor $\tilde{M}_{ijkl}'$), a new estimate of the grain stress can be obtained combining the local constitutive equation (23.1)
and the interaction equation (23.34) as follows:

\[
\dot{\gamma}_o \sum_s m^s_{ij} \left( \frac{m^{s}_{pq} \sigma'_{pq}}{\tau^s} \right)^n - D_{ij} = -\tilde{M}_{ijkl} (\sigma'_{kl} - \Sigma'_{kl}) \tag{23.46}
\]

Equation (23.46) constitutes a $5 \times 5$ non-linear system of algebraic equations, where the unknowns are the five independent components of the deviatoric stress tensor $\sigma'_{kl}$ of the grain. If the recalculated local stresses are different from the input values for any of the grains that constitute the polycrystal, a new iteration should be started. Otherwise, the iterative scheme is completed and the shear-rates on the slip (or twinning) systems and the strain-rate of each grain are calculated respectively by means of Equations (23.2) and (23.1), while the rotation-rate of each grain is obtained as:

\[
w_{ij} = W_{ij} + \tilde{w}_{ij} \tag{23.47}
\]

where $\tilde{w}_{ij}$ is given by Equation (23.32).

The above numerical scheme can be used either to obtain the anisotropic response of the polycrystal, probing it along different strain-paths (i.e.: applying different strain-rates $D_{ij}$ and obtaining the corresponding stress response $\Sigma'_{ij}$), or to predict texture development, by applying incremental deformation steps. The latter case requires the incremental updating of the shape and the orientation of the grains (due to both slip and twinning reorientation) and updating the critical stress of the deformation systems, due to strain hardening, as well. Details of these updating schemes are given in the next section.

### 23.3 Implementation of a Texture Development Calculation

#### 23.3.1 Kinematics

We provide here a brief overview of the equations used in kinematics. The reader is referred to the textbook of Gurtin (1981) for a comprehensive treatise on kinematics of continuum media. The present formulation is not exclusive to selfconsistent models, but applies to any polycrystal model. We define $\bar{x}^{(o)}$ as the initial coordinates of a point in the undeformed crystal and $\bar{x} (\bar{x}^{(o)})$ as the final coordinates of the same material point in the deformed crystal. The displacement field is therefore given by: $u_i (\bar{x}) = x_i - x_i^{(o)}$. Let us assume that the deformation of a grain is characterized by a given velocity gradient tensor $l_{ij} = \dot{u}_{i,j}$ and a deformation gradient tensor $f_{ij}$, defined as:

\[
l_{ij} = \frac{\partial \dot{u}_i}{\partial x_j} \tag{23.48}
\]

\[
f_{ij} = \frac{\partial x_i}{\partial x_j^{(o)}} \tag{23.49}
\]

The above quantities are related by:

\[
l_{ij} = \dot{f}_{ik} \dot{f}_{kj}^{-1} \tag{23.50}
\]
and:

\[ x_i = f_{ij}x^{(o)}_j \]  

(23.51)

Since plastic deformation is accommodated by crystallographic shear, which does not reorient the crystal axes, it is useful to identify the initial system with the initial position of the crystal axes. This allows us to decompose the deformation gradient into a plastic component \( f^{(o)}_{ij} \) referred to the initial state, and a lattice rotation \( R_{ij} \) that rotates the crystal to the final (current) configuration:

\[ f_{ij} = R_{ik}f^{(o)}_{kj} \]  

(23.52)

This decomposition applies at arbitrary large deformation. What makes this decomposition appealing is that in the initial configuration reference the crystal axes do not rotate, \( \vec{n}^s \) and \( \vec{b}^s \) remain invariant, and the velocity gradient is given by a linear superposition of shear rates on all active slip and twinning systems as:

\[ l_{ij}^{(o)} = \sum_s \dot{\gamma}_s b^s_{ik} n^s_j \]  

(23.53)

In the initial axes the plastic deformation tensor component obeys a relation analogous to (23.50):

\[ l_{ij}^{(o)} = \dot{\gamma} f^{(o)}_{ik} f^{(o)-1}_{kj} \]  

(23.54)

Decomposing the Schmid tensor \( b^s \otimes n^s \) into the symmetric and skew symmetric components:

\[ m_{ij}^s = \frac{1}{2} \left( b^s_{ik} n^s_j + b^s_{jk} n^s_i \right) \]  

(23.55a)

\[ q_{ij}^s = \frac{1}{2} \left( b^s_{ik} n^s_j - b^s_{jk} n^s_i \right) \]  

(23.55b)

allows us to decompose \( l_{ij}^{(o)} \) into a plastic strain-rate and a plastic rotation-rate:

\[ l_{ij}^{(o)} = d_{ij}^{(o)} + w_{ij}^{(o)} = \sum_s \dot{\gamma}_s m_{ij}^s + \sum_s \dot{\gamma}_s q_{ij}^s \]  

(23.56)

Combining (23.50), (23.52) and (23.54) we obtain:

\[ l_{ij} = \left( R_{ik} f^{(o)}_{kj} + R_{ik} f^{(o)}_{kj} \right) \left( f^{(o)-1}_{ik} R_{jk} \right) = \dot{R}_{ik} R_{jk} + R_{ik} R_{jl} d_{kl}^{(o)} + R_{ik} R_{jl} w_{kl}^{(o)} \]  

(23.57)

Hence, decomposing \( l_{ij} = d_{ij} + w_{ij} \), it holds that:

\[ d_{ij} = R_{ik} R_{jk} d_{kl}^{(o)} \]  

(23.58a)

\[ w_{ij} = \dot{R}_{ik} R_{jk} + R_{ik} R_{jl} w_{kl}^{(o)} \]  

(23.58b)
The strain rate, given by Equation (23.58a), is simply a transformation from initial to final crystal axes. The spin or rotation rate of the continuum, given by Equation (23.58b), allows us to obtain the rate of change of the crystal orientation matrix as:

$$\dot{R}_{ij} = (w_{ik} - w_{ik}^{(o)}) R_{kj}$$  \hspace{1cm} (23.59)

The latter can be used to update incrementally the orientation of the crystal and, as a consequence, to follow the texture evolution. Combining (23.59) with (23.47) the crystallographic rotation-rate of the crystal lattice associated with each grain is given by:

$$w_{ij}^{\text{cryst}} = W_{ij} + \dot{w}_{ij} - w_{ij}^{(o)}$$  \hspace{1cm} (23.60)

In addition, the velocity gradient and the incremental form of (23.50) can be used for updating the deformation gradient of the grain, after a time increment $\Delta t$:

$$f_{ij}(t + \Delta t) = (f_{ij}(t) + \dot{f}_{ij}(t) \Delta t) = (\delta_{ik} + l_{ik} \Delta t)f_{kj}(t)$$  \hspace{1cm} (23.61)

The above updated deformation gradient can be in turn used to update the shape of the grain as follows: a spherical locus of points in the initial configuration of the crystal is described by the equation:

$$x_{i}^{(o)} x_{i}^{(o)} = 1$$  \hspace{1cm} (23.62)

The corresponding locus in the deformed state can be calculated using (23.51) as:

$$x_{j}(f^{-1}T_{ji}(f^{-1})_{ik}x_{k} = x_{j}(f \cdot f^{T})_{j}^{-1} x_{k} = 1$$  \hspace{1cm} (23.63)

This is the equation of an ellipsoid. The eigenvectors and the inverse square root of the eigenvalues of $(f \cdot f^{T})$ define the direction and length of the axes of the ellipsoid representing the grain.

To summarize, after a time increment we have two distinctive rotations at grain level: one is the rotation of the crystallographic lattice associated with the grain, given by $w_{ij}^{\text{cryst}} \Delta t$ (Equation (23.60)), and the other is the rigid rotation of the continuum, given by Equation (23.58b). As for the ellipsoid that represents the grain shape, it is uniquely characterized by the eigenvalues and eigenvectors of $(f \cdot f^{T})$.

### 23.3.2 Hardening

The threshold stress $\tau^{*}$ that appears in Equation (23.1) describes (in an average way) the resistance for activation that the slip or twinning modes experience, and it generally increases with deformation. In the applications of Section 23.4, and in most of our polycrystal simulations, we employ an empirical extended Voce law (Tomé et al. 1984) to describe the variation of $\tau^{*}$ in each system of each deformation mode. The expression that characterizes the evolution of the threshold stress with accumulated shear strain in each grain is:

$$\dot{\tau}^{*} = \tau^{*}_{0} + (\tau^{*}_{1} + \theta^{*}_{1} \Gamma)(1 - \exp(-\Gamma |\theta^{*}_{0}/\tau^{*}_{1}|))$$  \hspace{1cm} (23.64)
23.3 Implementation of a Texture Development Calculation

where $\Gamma$ is the accumulated shear in the grain; $\tau_0, \theta_0, \theta_1$ and $(\tau_0 + \tau_1)$ are: the initial threshold stress, the initial hardening rate, the asymptotic hardening rate and the back-extrapolated threshold stress, respectively. In addition, we allow for the possibility of self and latent hardening by defining coupling coefficients $h^{ss'}$ which empirically account for the obstacles that dislocations on system ($s'$) pose to propagation of dislocations in system ($s$). Eventually, the increase in the threshold stress in each slip or twinning system, after a time increment $\Delta t$, is calculated as:

$$\Delta \tau_s = \frac{\hat{d}\tau_s}{d\Gamma} \sum_{s'} h^{ss'} \gamma_{s'} \Delta t$$  (23.65)

We use self-hardening as a reference and set $h^{ss} = 1$. When self hardening and latent hardening are indistinguishable, then $h^{ss'} = 1$. This hardening law allows us to describe the high hardening rate observed at the onset of plasticity, and its decrease towards a constant rate, at large strains. The condition $\theta_0 \geq \theta_1 \geq 0, \tau_1 \geq 0$ corresponds to increasing yield stress and decreasing hardening rate tending to linear saturation. Linear hardening is a limit case of this law, and takes place when $\tau_1^* = 0$, while a rigid perfectly plastic response is given by $\theta_0 = \theta_1 = \tau_1 = 0$.

It is straightforward to implement more complex hardening laws in polycrystal codes. Peeters et al (2001) and Toth et al. (2002) describe the evolution of $\tau_s$ with strain using dislocation models. Kok et al. (2002), extend the Mechanical Threshold Strength model to describe the dependence of $\tau_s$ with temperature, shear rate, and accumulated shear in the grain.

23.3.3 Twinning Reorientation

While here we assume that twinning has associated, as slip, a critical resolved shear of activation in the twinning plane and along the twinning direction, it differs from slip in its directionality, which we model by allowing activation only if the resolved shear stress is positive (along the Burgers vector of the twin). Another aspect of twinning that needs to be incorporated into the models is the fact that the twinned fractions are regions (usually of lamellar morphology) with a different orientation than the surrounding matrix. These twinned regions not only contribute to the texture of the aggregate but, most important, act as effective barriers for the propagation of dislocations or of other twin lamellae. The hardening induced by the twins is empirically enforced here by assigning high values to the latent hardening coefficients $h^{ss'}$ describing slip-twin and twin-twin interactions.

As for the effect on texture of the twinned fractions, here we use the Predominant Twin Reorientation Scheme (PTR) proposed by Lebensohn and Tomé (1993), which works as follows: within each grain ($g$) we keep track of the shear strain $\gamma_{t,g}$ contributed by each twin system ($t$), and of the associated volume fraction $V_{t,g} = \gamma_{t,g} / \gamma^{tw}$ as well, where $\gamma^{tw}$ is the characteristic twin shear. The sum over the $T^{tw}$ twin systems associated with twin mode ($m$), and over all grains, represents the accumulated twinned fraction $V_{acc,m}$ in the aggregate for the particular twin mode (the one that one would measure by SEM), namely:

$$V_{acc,m} = \sum_g \sum_{t=1}^{T^{tw}} \gamma_{t,g} / \gamma^{tw}$$  (23.66)
Since it is not numerically feasible to consider each twinned fraction as a new orientation, the PTR scheme adopts a statistical approach. At each incremental step some grains are fully reoriented by twinning, provided certain conditions are fulfilled. We call effective twinned fraction \( V_{\text{eff},m} \) the volume associated with the fully reoriented grains for that mode, and define a threshold volume fraction \( V_{\text{thr},m} \) as:

\[
V_{\text{thr},m} = A_{\text{th}1} + A_{\text{th}2} \frac{V_{\text{eff},m}}{V_{\text{acc},m}}
\]

where \( A_{\text{th}1} \) and \( A_{\text{th}2} \) are appropriate threshold values (typically 0.1 and 0.5, respectively). After each deformation increment we pick a grain at random and identify the twin system with the highest accumulated volume fraction. If the latter is larger than the threshold \( V_{\text{thr},m} \) then the grain is allowed to reorient and \( V_{\text{eff},m} \) and \( V_{\text{thr},m} \) are updated. The process is repeated until either all grains are randomly checked or until the effective twin volume exceeds the accumulated twin volume. In the latter case we stop reorientation by twinning and proceed to the next deformation step. Two things are achieved in this process: a) only the historically most active twin system in each grain is considered for reorienting the whole grain by twinning; b) the twinned fraction is consistent with the shear activity that the twins contribute to deformation.

The PTR scheme described above was implemented by Tomé et al. (2001) for Zr, and an example is presented in Section 23.4. Other twin reorientation models implemented in polycrystal codes are: the Monte-Carlo scheme used by Van Houtte (1978) to simulate plastic deformation of brass, also based on reorienting whole grains by twinning; the Volume Fraction Transfer (VFT) scheme proposed by Tomé et al. (1991), which accounts for all twin fractions and accumulates them in Euler space cells; Kalidindi (1998) allows for an increasing number of orientations as twinning takes place (see also Chapter 27 in this book).

### 23.4 Applications

#### 23.4.1 Tension and Compression of FCC

Axisymmetric tensile and compressive deformation of FCC aggregates is one of the most familiar and simpler applications of a polycrystal code. Despite its simplicity, however, this example allows us to illustrate the basic features of a simulation of texture evolution, and also the differences in the predictions of different polycrystal models. A comprehensive discussion of the subject can be found in Kocks et al. (1998) and in Tomé (1999). The conditions used in the simulations that follow are:

a) The initial texture consists of 500 initially random orientations.

b) Slip takes place on (111) planes along \(<110>\) directions (i.e.: 12 systems plus opposites).

c) The single crystals have no strain-hardening, i.e.: \( \tau_0 = 1, \tau_1 = 0, \theta_0 = 0 \) and \( \theta_1 = 0 \) (see Equation (23.64)). This feature allows us to analyze the ‘texture hardening’ effect, which is a geometric effect associated with the orientation of the grains with respect to the loading axis, and their reorientation under tension (compression).
23.4 Applications

d) A rate sensitivity exponent \( n=20 \) is used in the viscoplastic constitutive law (Equation (23.1)).

e) The applied deformation is axial tension and compression along \( x_3 \), up to an accumulated longitudinal true strain \( \varepsilon_{33} = 1.0 \). All components of the macroscopic velocity gradient are imposed to the aggregate, e.g., in the tensile case:

\[
\dot{U}_{i,j} = \begin{pmatrix}
-0.5 & 0 & 0 \\
0 & -0.5 & 0 \\
0 & 0 & 1.0 \\
\end{pmatrix} \, s^{-1}
\]  

(23.68)

f) The time increment adopted is \( \Delta t = 0.02 \, s \)  \( (\Delta \varepsilon_{33} = D_{33} \Delta t = 0.02) \). The stress-strain state, the crystallographic orientation and the shape of each grain are updated after every incremental deformation step.

The results are illustrated in terms of inverse pole figures, showing the tensile or compressive axis in the crystal reference frame (Figure 23.1) and the stress-strain response and the slip system activity as a function of the deformation strain (Figure 23.2).

The simulations were done using different approaches, each one characterized by the degree of stiffness in the grain-matrix interaction: the FC approach is infinitely stiff, the secant formulation is fairly stiff, and the tangent formulation is fairly compliant. We also used an intermediate interaction scheme (see Equation (23.45)) using \( n_{\text{eff}} = 10 \) (Tomé, 1999). It is evident from Figure 23.1 that the different approximations have associated different final textures and slip activities. In what concerns axisymmetric tension, the FC and secant approaches predict similar final textures, with a strong (111) component and a weaker (100) component. The more compliant approaches, instead, predict a fiber joining the (111) and (100) texture components. The latter result is a consequence of the grains accommodating deformation with less than 5 independent systems (Figure 23.2). As for the macroscopic stress (Figure 23.2), it is clear that, although the final texture leads to geometric hardening, less stress is required to deform the aggregate when the grain-matrix interaction is more compliant. FC provides an upper bound to the yield stress.

The case of axisymmetric compression is even more interesting, because the more compliant cases correctly predict the (110) component observed experimentally, while the stiffer FC and secant cases fail to fully reorient the compression axis along (110). Once again, the reason is that the more compliant cases (tangent and \( n_{\text{eff}} = 10 \)) are more efficient in accommodating deformation: they require less active systems per grain because grains tend to deform by plane (as opposed to axial) strain, and they also require lower yield stress.

23.4.2 Torsion (Shear) of FCC

In this application we simulate torsion of a thin walled tube using the VPSC polycrystal model. Locally, the deformation can be regarded as simple shear, and we analyze two cases: a fix-end and a free-end torsion test. Assume that \( x_1 \) is the circumferential direction, \( x_2 \) is the axial direction and \( x_3 \) is the radial direction of the tube. The macroscopic velocity gradients
Figure 23.1: Inverse pole figure of the tensile axis (top) and the compressive axis (bottom) after 1.0 longitudinal true strain. Simulations were done using the following approaches: Full Constraints (FC), secant (SEC), tangent (TAN) and an intermediate $n^{eff} = 10$ in the interaction equation.

associated with fix and free-end torsion are, respectively:

$$
\dot{U}_{i,j} = \begin{vmatrix}
0 & 1 & 0 \\
0 & 0 & 0 \\
0 & 0 & 0
\end{vmatrix} s^{-1} \quad \text{(fix-end torsion)} (23.69)
$$

$$
\dot{U}_{i,j} = \begin{vmatrix}
? & 1 & 0 \\
0 & ? & 0 \\
0 & 0 & ?
\end{vmatrix} s^{-1} \quad \text{(free-end torsion)} (23.70)
$$

In the latter equation, the question marks indicate that the component is not a priori imposed. In the first case (fix-end), the tube is not allowed to deform either in the through-thickness ($D_{11} = 0$) or in the axial direction ($D_{22} = 0$). From incompressibility it follows that $D_{33} = 0$. In the second case (free-end), the tube is not constrained in the through-thickness and the axial directions ($\sigma_{11} = 0$ and $\sigma_{22} = 0$). As a consequence, the corresponding macroscopic rates are unknown until the deformation of the aggregate is solved. Within the VPSC formulation it is straightforward to enforce mixed boundary conditions on stress and strain rate components inside the interaction equation (23.46). As in the previous example, we start the simulation using a discrete distribution of 500 random orientations, and assume that slip takes place in the $(111)<110>$ systems. For crystal hardening we choose a linear hardening law characterized by: $\tau_0 = 1$, $\tau_1 = 0$, $\theta_0 = 1$ and $\theta_1 = 1$. We enforce time increments of $\Delta t = 0.04 s$ ($\Delta \varepsilon_{12} = \frac{D_{12}}{D_{11}} \Delta t = 0.02$) until we accumulate a total shear strain of $\varepsilon_{12} = 2.0$. The stress and orientation of each grain is updated after every step. Interestingly, the calculation of the resulting macroscopic stress and strain components shows that an axial stress develops during the fix-end torsion, while an axial strain evolves in the case of free-end torsion. Both effects are a direct consequence of the texture that develops during torsion,
which does not have the same symmetry as the test. These effects can be discussed using the concept of polycrystal yield surface (PCYS) defined as the locus of stress states which induce plastic yield. The associated strain-rate is normal to such locus, as prescribed by the associative flow law or normality rule. Numerically, points on the PCYS can be obtained by probing the aggregate (accounting for current texture and hardening) in different strain-rate directions, and calculating the associated stress state (for PCYS’s derived from experimental textures and a more comprehensive discussion, see Kocks et al., 1998, Ch. 10).

Figure 23.3 displays the (111) pole figures associated with fix-end and free-end torsion test simulations, up to $\varepsilon_{12} = 2.0$. The fix-end one is a typical FCC shear texture. The free-end pole figure differs from the fix-end one only in a clockwise rotation of about 10 deg. The final textures, however, do not provide any information about the evolution of texture and stress. Figure 23.4 shows a $(\sigma_{22},\sigma_{12})$ section of the yield locus, calculated at $\varepsilon_{12} = 1.0$ and 2.0. It also shows the evolution of the axial stress $\sigma_{22}$ as a function of the shear strain $\varepsilon_{12}$. Observe that the axial stress becomes compressive as torsion increases, and eventually
reverses the trend at large shear strains. Such situation is graphically illustrated by the yield locus in Figure 23.4. The fix-end condition indicates that the normal at the yield point must have a zero $D_{22}$ component which, in this case, gives a non-zero negative stress component $\sigma_{22}$. The free-end case, on the other hand, is characterized by a monotonically increasing axial true strain $\varepsilon_{22}$ (tube elongation). In the yield locus section of the $(\sigma_{22}, \sigma_{12})$ subspace, such testing condition requires the stress vector to have zero axial component ($\sigma_{22} = 0$). As a consequence, the normal to the yield surface at such point has a positive non-zero component $D_{22}$. Also note that, although the final textures are not very different for fix-end and free-end torsion, the magnitude of the axial stress or the axial strain is comparable to the corresponding shear components (Figure 23.4).

**Figure 23.3:** (111) pole figures associated with fix-end and free-end torsion testing after an accumulated shear strain $\varepsilon_{12} = 2.0$.

**Figure 23.4:** $(\sigma_{22}, \sigma_{12})$ sections of the Polycrystal Yield Surface after a shear strain of $\varepsilon_{12} = 1.0$ and $2.0$. Evolution of axial and shear components associated with fix-end and free-end torsion simulations.
23.4 Applications

23.4.3 Twinning and Anisotropy of HCP Zr

Zirconium exhibits HCP structure below about 900°C. This example is meant to illustrate the response of an aggregate which exhibits slip and twinning deformation modes, and which is highly anisotropic both, at the single crystal and the polycrystal level (for a comprehensive discussion of simulations and experiments done in clock-rolled Zr, see Tomé et al. 2001 and Kaschner et al. 2001).

The initial texture of the material is represented using 377 orientations, and corresponds to a clock-rolled and annealed Zr plate, with a strong c-axis component (hard) along the Normal Direction (ND, direction x3) of the plate (Figure 23.5). We consider that deformation takes place at room temperature. Experimental evidence indicates that under this condition HCP crystals deform by prismatic <a> slip, pyramidal <c+a> slip and tensile twinning. The single crystal hardening parameters were adjusted to experimental tensile and compression data. They are:

a) \( \tau_0 = 5 \), \( \tau_1 = 30 \), \( \theta_0 = 1500 \), \( \theta_1 = 50 \) MPa, for (1010)<1120> prismatic slip,

b) \( \tau_0 = 70 \), \( \tau_1 = 270 \), \( \theta_0 = 3000 \), \( \theta_1 = 25 \) MPa, for (10-11)<1123> pyramidal slip,

c) \( \tau_0 = 50 \), \( \tau_1 = 0 \), \( \theta_0 = 75 \), \( \theta_1 = 75 \) MPa, for (1012)<1123> tensile twinning.

The following deformation modes are imposed to the sample:

a) Axial through-thickness compression (TTC), up to a longitudinal true strain of \( \varepsilon_{33} = 0.2 \) at a rate of \( 10^{-2} \) s\(^{-1}\). The lateral components \( \sigma_{11} \) and \( \sigma_{22} \) are imposed to be zero, while \( \varepsilon_{11} \) and \( \varepsilon_{22} \) are left unconstrained.

b) Axial in-plane compression (IPC), up to longitudinal true strain of \( \varepsilon_{11} = 0.2 \), at a rate of \( 10^{-2} \) s\(^{-1}\). The lateral components \( \sigma_{22} \) and \( \sigma_{33} \) are imposed to be zero, while \( \varepsilon_{22} \) and \( \varepsilon_{33} \) are left unconstrained. This boundary condition, coupled with the stiffness of the texture along the ND, leads to a prediction of strong ovalization in the section of the sample.

c) Axial in-plane tension (IPT), up to \( \varepsilon_{11} = 0.2 \), at a rate of \( 10^{-2} \) s\(^{-1}\). The lateral components \( \sigma_{22} \) and \( \sigma_{33} \) are imposed to be zero.

Figure 23.6 depicts the experimental stress-strain response of this material. The difference between the TTC test and the IPC test highlights the strong anisotropy of the aggregate. In addition, even the compressive and the tensile results for the In Plane tests are different. Superimposed, we show in Figure 23.6 the simulated response using the SC polycrystal model and the hardening parameters listed above for the room temperature deformation modes. As a matter of fact the hardening parameters were adjusted for reproducing the experimental response (Tomé et al. 2001), and the extended Voce hardening (Equation (23.64)) for each of the deformation modes is shown in Figure 23.7. It is clear that prism slip is the easiest to activate, followed closely by tensile twinning. The latter can only accommodate tensile deformation along the c-axis and, as a consequence, the much harder pyramidal slip is required for providing c-axis compression. The reader should be aware, though, that once the parameters are fixed, the constitutive law should correctly predict ovalization, deformation texture, and the response of Zr subjected to other deformation modes. The mode activity predicted by the
model, depicted in Figure 23.8, provides an explanation for the differences in the mechanical response among the different tests. In TTC the hard pyramidal slip dominates, which explains the much higher macroscopic stress, and the relatively little texture evolution (Figure 23.5). Prismatic systems are favorably oriented for IPC, and they dominate this deformation mode, although there is also a non-negligible contribution of tensile twins (Figure 23.8). The latter lead to a reorientation by twinning of about 20% of the total volume, which shows in the basal pole figures as a new component along sample axis $x_1$ (Figure 23.5). The sample develops strong ovality, which practically amounts to plane strain deformation ($\varepsilon_{22} = 0.196$, $\varepsilon_{33} = 0.004$). In the IPT case the deviatoric stress induced in the grains contains a compressive component along the c-axis and a favorable resolved shear on the prismatic planes. As a consequence, deformation is entirely accommodated by prism slip (Figure 23.8) since tensile twins cannot be activated and pyramidal slip, which is too hard, can be circumvented. The fact that for IPT the texture is unchanged confirms the overwhelming prism activity (which does not reorient the c-axis).

We want to emphasize that the possibility of accommodating deformation by activating mostly soft modes when they are available is a characteristic of SC polycrystal models. The interaction equation (Equation (23.34)) accounts for the relative directional stiffness of grain and surrounding medium, and allows for the medium to accommodate more strain than the grain in certain directions. As a consequence, the condition of having five independent systems active per grain, is not required within SC schemes, and hard systems tend to be less active, when kinematically possible.

### 23.4.4 Compression of Olivine (MgSiO$_4$)

Olivine is a geologic material of orthorhombic symmetry, found in the composition of the Earth mantle. Olivine textures are a result of convective deformation of the upper mantle, and affect the characteristics of seismic wave propagation. This example is meant to present the reader with an application of polycrystal simulation for a material which lacks five independent slip systems. While the FC formalism would be inapplicable to this problem, the SC
Figure 23.6: Experimental (dot line) and predicted (solid line) stress-strain curves of clock-rolled Zr deforming at room temperature in through-thickness compression (TTC), in-plane compression (IPC) and in-plane tension (IPT) as a function of longitudinal true strain.

Figure 23.7: Extended Voce law describing the evolution of CRSS for prism slip, pyramidal slip and tensile twins as a function of the accumulated shear strain in the grain. Room temperature fits to curves in Figure 6.
formalism can still be used for describing plastic deformation of olivine. For a comprehensive discussion of simulation and experiments see Kocks et al. (1998).

The conditions used in the following olivine simulations are:

a) The initial texture consists of 500 initially random orientations.

b) Olivine’s orthorhombic crystals deform by slip on (010)<100>, (001)<100> and (010)<001> systems. Here we assume relative threshold stresses \( \tau_{(010)<100>} = \tau_{(001)<100>} = 0.4 \) and \( \tau_{(010)<001>} = 1.0 \), and no strain hardening.

c) The applied deformation was axial compression along \( x_3 \), up to a longitudinal true strain \( \varepsilon_{33} = 0.5 \), at a rate of \( 10^{-2} \text{ s}^{-1} \).

There is a numerical issue for materials with open yield surfaces as this one. Since the initial guess of stress in each grain is done via FC, it is necessary to close the single crystal yield surface temporarily to avoid numerical singularities. This may be done by adding extra deformation modes with high CRSS. Once convergence has been achieved, they are permanently switched off after the first deformation step. For this particular application slip modes (110)<110> and (021)<112> were added, in order to close the yield surface.

The final texture and the system activity are shown in Figures 23.9 and 23.10, respectively. The final texture resembles the experimental one (Kocks et al, 1998, Ch. 6) and the deformation is accommodated using less than 3 active systems per grain.

23.5 Further Selfconsistent Models and Applications

In this article we limit our discussion of selfconsistent polycrystal models to the case of plastic forming of single-phase aggregates, described by a viscoplastic constitutive law. Our purpose is to illustrate the basic approach in which selfconsistent models are based, and how their predictions of texture development and system activity compare with the more traditional
23.5 Further Self-consistent Models and Applications

Taylor model. We also present some applications where the plastic anisotropy of the single crystal requires a self-consistent approach to the problem. Since elastic effects can be neglected during plastic forming, and since creep is usually of no concern, we have deliberately omitted here the description of elastoplastic and viscoelastic self-consistent polycrystal models. In what concerns the former models, the interested reader may refer to Hutchinson's (1970) implementation of Hill's (1965) elasto-plastic formulation, and the approach of Berveiller and Zaoui (1979), both simplified to isotropic applications. The more recent implementations of Hill's model by Turner and Tomé (1994) and Lebensohn et al (1996) account for fully anisotropic crystals and polycrystals. In what concerns viscoelasticity, several models have been proposed by French researchers, with varying degrees of complexity (Kouddane et al., 1993; Rougier et al., 1994; Sabar et al., 2002), but all the implementations address only isotropic phases. Turner and Tomé (1993) and Turner et al. (1994) implement fully anisotropic viscoelastic self-consistent formulations for linear and non-linear viscosity, respectively.

Several extensions of the 1-site VPSC formulation can be found in the literature. These extensions address the following problems and applications: calculation of internal stresses (Lebensohn et al. 1998a), kinematics effects due to grain interaction (Bolmaro et al. 1997; Tomé et al. 2002), deformation-based recrystallization (Wenk et al. 1997; Lebensohn et al.
1998b; Takeshita et al 1999), solution of the inverse VPSC problem (Signorelli et al. 2000), non-monotonic deformation histories (Wenk et al. 2000), coupling of VPSC and FE methods (Tomé et al. 2001, Kaschner et al. 2001) and VPSC for voided polycrystals (Lebensohn et al. 2003a,b). In addition, among the several multi-site extensions of the VPSC formulation it is worth mentioning: 2-site VPSC formulation (Lebensohn and Canova 1997), n-site VPSC formulation (Canova et al. 1992, Solas and Tomé 2000), VPSC for lamellar structures (Lebensohn 1999), n-site deformation-based recrystallization (Solas et al. 2001).

Finally, it is worth mentioning the recent variational SC formulations for viscoplastic polycrystals of Ponte Castañeda, based on the use of variational linear comparison methods, which express the effective potential of the nonlinear viscoplastic polycrystal in terms of that of a linearly viscous polycrystal with properties that are determined from suitably designed variational principles. The main advantage of these methods is that they make explicit use of the local field fluctuations in the determination of the constitutive behavior of the polycrystal. The first of these formulations is Ponte Castañeda’s (1991) variational method and makes use of the SC approximation for linearly viscous polycrystals to give bounds on the SC estimate for viscoplastic polycrystals. The second is based on the recently proposed second-order method (Ponte Castañeda 2002) and makes use of the SC approximation for a more general class of linearly viscous polycrystals, incorporating residual stresses to generate more accurate SC estimates for viscoplastic polycrystals (Liu and Ponte Castañeda 2003).

References


References


Solas D.E., Tomé C.N., 2001. Texture and strain localization prediction using a N-site polycrystal model" Int. J. Plasticity 17, 737-753


References


24 Phase-field Extension of Crystal Plasticity with Application to Hardening Modeling

Bob Svendsen

The purpose of this chapter is the brief discussion and application of recently formulated extensions of standard crystal plasticity to account for the possible effects of incompatibility in the local inelastic deformation on material behaviour. In particular, attention is focused here on additional hardening caused by the development of so-called geometrically-necessary dislocations, and more generally, of dislocation substructures, accommodating such incompatibility. In the context of crystal plasticity, such local inelastic deformation is represented by the glide-system (scalar) slips. As it turns out, this quantity and its spatial gradient can be used as a basis on which to formulate a continuum thermodynamic, multifield, multiscale generalization of standard crystal plasticity in which these slip can be interpreted as phase fields. Indeed, in this extended context, one derives a generalized Ginzburg-Landau or Cahn-Allen-type evolution/field relation for these slips. The form of this and the other field relations of the model is determined by that of the free energy, by that of a dissipation potential, and by that of the evolution relations for the local inelastic state variables. Modeling and simulation results for the constrained simple shear of a thin crystalline strip demonstrates the ability of the approach to predict the additional hardening due to local deformation incompatibility observed for fcc-systems.

24.1 Introduction

Standard micromechanical modeling of the inelastic material behaviour of metallic single crystals and polycrystals (e.g., (Hill and Rice, 1972; Asaro, 1983; Cuitino and Ortiz, 1992)) is commonly based on the premise that resistance to glide is due mainly to the random trapping of mobile dislocations during locally homogeneous deformation. Such trapped dislocations are commonly referred to as statistically-stored dislocations (SSDs), and act as obstacles to further dislocation motion, resulting in hardening. As anticipated in the work of Nye (1953) and Kröner (1960), and discussed by Ashby (1970), an additional contribution to the total density of dislocations and so to hardening can arise when the continuum lengthscale (e.g., grain size) approaches that of the dominant microstructural features (e.g., mean spacing between precipitates relative to the precipitate size, or mean spacing between glide planes). Indeed, in this case, the resulting deformation incompatibility between, e.g., “hard” inclusions and a “soft” matrix, is accommodated by the development of so-called geometrically-necessary dislocations (GNDs), and more generally, by that of dislocation substructures. Experimentally-observed effects in a large class of materials such as increasing material hardening with de-
creasing (grain) size (i.e., the Hall–Petch effect) are commonly associated with the development of such GNDs. As standard crystal plasticity accounts (at the glide-system level) for the effects of SSDs alone on the hardening behaviour of single and polycrystals, there arises the need to extend crystal plasticity in some fashion if one is interested in modeling such size-dependent aspects of the material behaviour. Such considerations and results have motivated a number of workers over the last few years to formulate various extensions of phenomenological and/or crystal plasticity (e.g., Fleck and Hutchinson (1993); Naghdi and Srinivasa (1994); Le and Stumpf (1996); Steinmann (1996); Dai and Parks (1997); Forest, Cailletaud and Sievert (1997); Shizawa and Zbib (1999); Shu and Fleck (1999); Acharya and Bassani (2000); Gurtin (2000); Menzel and Steinmann (2000); Ortiz, Repetto and Stainier (2000); Svendsen (2000); Gurtin (2002); Svendsen (2002)) in order to account in various ways for the dependence of the material behaviour on deformation inhomogeneity (e.g., strain gradients), or in the crystal plasticity context, deformation incompatibility. It is not the purpose of the current work to compare and contrast all of these approaches; rather, the focus here is on a brief formulation and discussion of one such approach and model ((Svendsen, 2000),(Svendsen, 2002)) in the crystal plasticity context. Like most, if not the others, this approach applies at least to fcc-based systems. Model predictions of additional hardening due to deformation incompatibility are then examined in the context of the simple shear of a thin crystalline strip (e.g., Shu, Fleck, Van der Giessen and Needleman (2001)). Such a strip can serve for example as a first approximation of a thin film. Again, the focus here is on hardening; other predictions of the model, such as those for dislocation substructure development, will be reported on elsewhere.

### 24.2 Basic Considerations and Results

Phenomenological or continuum crystal plasticity is based on the assumption that any material point in the material or structure of interest is endowed with a (finite) set of \( n \) glide systems (see Chapter 5). The geometry and orientation of each such glide system is described as usual by an orthonormal basis \((s_a, n_a, t_a)\) \((a = 1, \ldots, n)\). Here, \( s_a \) represents the direction of glide in the plane, \( n_a \) the glide-plane normal, and \( t_a := s_a \times n_a \) the direction transverse to \( s_a \) in the glide plane. Since we neglect in this work the effects of any processes involving a change in or evolution of either the glide direction \( s_a \) or the glide-system orientation \( n_a \), these unit vectors, and so \( t_a \) as well, are constant. Assuming then that inelastic slip in the glide planes due to dislocation motion is the only source of local inelastic deformation \( \dot{F}_p \), in the material, the usual form

\[
\dot{F}_p = \sum_{a=1}^{n} (s_a \otimes n_a) F_p \gamma_a \tag{24.1}
\]

for the evolution of \( F_p \) applies (e.g., Asaro (1983), Teodosiu (1997)). Here, \( \gamma_a \geq 0 \) is interpreted as the (positive accumulated scalar inelastic) slip deformation\(^1\) in \( a^{th} \) glide plane. From (24.1), one sees that standard crystal plasticity does not account for the effect of any spatial

\(^1\) In the usual case, \( \gamma_a \) in (24.1) represents the net inelastic slip (i.e., of either sign). In the current case, \(-s_a\) and \(s_a\) are distinct glide systems (e.g., Ortiz, Repetto and Stainier (2000)).
inhomogeneity in \( \gamma_a \), e.g., due to the development of GNDs, or more generally to that of dislocation substructures, on the local inelastic state in the material as gauged by \( F_p \). As such, one requires an additional measure to take this effect into account. To this end, note that GND and more generally dislocation substructure development is driven by (short-range) stress build-up during loading resulting from, e.g., orientation mismatch between distinct glide systems in grains, or between grains at grain boundaries. Alternatively, as already mentioned above, the development of these may be driven by stress build-up resulting from deformation-state mismatch between, e.g., “hard” inclusions and a “soft” matrix in composites. In turn, their development results in a minimization of such stress build-up and so of the energy stored in the material during loading. As such, the resulting inhomogeneous local deformation state is energetically-favoured over the corresponding homogeneous one (e.g., Ortiz, Repetto and Stainier (2000)). As recognized in essence by Nye (1953), an effective continuum measure of the accommodation of such local deformation inhomogeneity by dislocations is given by the incompatibility

\[
\text{curl} \ F_p = \sum_{a=1}^{n} (s_a \otimes n_a)(\text{curl} \ F_p) \dot{\gamma}_a + s_a \otimes (\nabla \dot{\gamma}_a \times F_p^T n_a) \quad (24.2)
\]

consistent with (24.1). This interpretation of \( \text{curl} \ F_p \) is based on that of the magnitude of the line integral \( \oint_C F_p \cdot t \) of \( F_p \) around any closed material curve \( C \) with direction \( t \) as a measure of the effective length of GNDs in the material. In addition, the projection of this measure onto any material surface can be related to the effective surface density of such dislocations (e.g., Svendsen (2000), Svendsen (2002)). Indeed, one can derive the form

\[
\text{curl} \ F_p = b \sum_{a=1}^{n} s_a \otimes g_{oa} \quad (24.3)
\]

for the incompatibility of \( F_p \) in terms of the magnitude \( b \) of the Burgers vector and a set \( (g_{o1}, \ldots, g_{on}) \) of vector-valued GND densities whose evolution is given by

\[
g_{oa} = \sum_{b=1, b \neq a}^{n} (n_a \cdot s_b) g_{ob} \dot{\gamma}_a + b^{-1} \nabla \dot{\gamma}_a \times F_p^T n_a \quad (24.4)
\]

consistent with (24.2). The first term in this relation depending in particular on the orientation mismatch between distinct glide systems leads to latent-hardening-like effects, and the second to self-hardening-like ones. An alternative or related interpretation of \( \text{curl} \ F_p \) arises in the the context of the line-tension model for dislocations, which was exploited by Ortiz, Repetto and Stainier (2000) to relate the magnitude of \( \text{curl} \ F_p \) to the dislocation core energy.

As shown, e.g., by Dai and Parks (1997), measures of local deformation incompatibility and GND density such as \( \text{curl} \ F_p \) or \( g_{oa} \) can be incorporated into standard finite-element-based crystal plasticity modeling formally as additional internal variables by evaluating the gradient-dependent terms numerically. To this end, one is forced in particular to integrate (24.2) or (24.4) explicitly and obtain the gradient terms via post-processing. More generally, however, the above relations imply that the set \( \gamma := (\gamma_1, \ldots, \gamma_n) \) of glide-system slips represent additional continuum degrees-of-freedom, while \( F_p \) and \( \text{curl} \ F_p \) are measures of the local inelastic state of the material. To discuss this further, note first that (24.1) depends (linearly) on \( F_p \) and the set \( \dot{\gamma} := (\dot{\gamma}_1, \ldots, \dot{\gamma}_n) \). Similarly, (24.2) depends (linearly) on
492

24 Phase-field Extension of Crystal Plasticity with Application to Hardening

curl \( F_p, F_p, \gamma \), and the set \( \nabla \gamma := (\nabla \gamma_1, \ldots, \nabla \gamma_n) \). As such, both (24.1) and (24.2) represent special cases of the quasi-linear evolutionary constitutive form

\[
\dot{\alpha} = K(\alpha) \gamma + J(\alpha) \nabla \gamma
\]

(24.5)

for a general set \( \alpha \) of internal inelastic/dislocation state variables determined by constitutive quantities \( K \) and \( J \). For example, these could include \( F_p \) as well as set of glide-system densities for SSDs and GNDs (i.e., such as \( g_{g1}, \ldots \)), or alternatively, \( F_p, \gamma, \) and \( \text{curl} F_p \). In any case, we then have the situation the evolution of such internal state variables becomes dependent on the \( \nabla \dot{\gamma} \) of the accumulated slip rates \( \gamma \) now requires that we model the \( \gamma \) as time-dependent fields. Consequently, in the current thermomechanical context, the absolute temperature \( \theta \), the motion/deformation \( \xi \), and the set \( \gamma \) of glide-system accumulated slips, represent the principal time-dependent fields, \( \alpha \) being determined constitutively by the history of \( \gamma \) and \( \nabla \gamma \). In (24.1) and (24.5), respectively. Restricting attention to quasi-static processes for simplicity, these considerations then imply the general material frame-independent constitutive form

\[
C = C(\theta, C, \nabla \theta, \alpha, \gamma, \nabla \gamma)
\]

(24.6)

for all dependent constitutive quantities, e.g., stress, heat flux, internal energy, or entropy, of the model. Here, \( C = F^T F \) represents the right Cauchy-Green deformation tensor, and \( F = \nabla \xi \) the deformation gradient. Applying methods developed in previous work (Svendsen (1999); Svendsen (2001)), Svendsen (2000) and Svendsen (2002) showed that the \( \gamma \) can be modeled from a phenomenological, continuum thermodynamic point of view as either (i), generalized internal variables, or as (ii), internal degrees of freedom. Indeed, either of these approaches facilitate in particular the derivation of a field relation for \( \gamma \). In the case of (ii), this relation follows in particular from the invariance of the total energy balance. With the help of either of these approaches, exploitation of the dissipation or entropy principle toward fulfillment of the second law of thermodynamics results in particular in the reduced form

\[
\psi = \psi(\theta, C, \alpha)
\]

(24.7)

from (24.6) for the free energy density \( \psi \). Further, one obtains the usual hyperelastic form

\[
S = 2 \psi C
\]

(24.8)

for the second Piola-Kirchhoff stress \( S \) appearing in the derived form

\[
0 = \text{div}(FS)
\]

(24.9)

for momentum balance, i.e., restricted to quasi-static conditions and neglecting any momentum supplies for simplicity. In addition, the entropy balance together with (24.7) and the reduced dissipation-rate density yield the field relation

\[
e \dot{\theta} = \frac{1}{2} \theta S_{\theta} \dot{C} + (d_{\gamma} \gamma + \theta K^T \psi_{\theta \alpha} \gamma) \nabla \gamma + (d_{\nabla \gamma} \gamma + \theta J^T \psi_{\theta \alpha} \nabla \gamma) + \text{div} d_{\nabla \ln \theta}
\]

(24.10)

for temperature (i.e., assuming no internal energy supplies) in terms of a dissipation potential \( d \) of the form (24.6), with \( e := -\theta \psi_{\theta \theta} \) the heat capacity at constant \( \gamma, C, \) and so on. Lastly, one derives the form

\[
d_{\gamma} = \text{div} (J^T \psi_{,\alpha} + d_{\nabla \gamma}) - K^T \psi_{,\alpha}
\]

(24.11)
for the evolution of $\gamma$ determined solely by the forms of the free energy density $\psi$ and the dissipation potential $d$. In effect, this result represents a non-local generalization of the glide-system flow rule, with the right-hand side being the effective thermodynamic force driving glide-system dislocation motion and so slip. As will become quite apparent in the special case of small strain to be discussed in the next section, (24.11) represents a type of generalized Ginzburg-Landau or Cahn-Allen-type field relation for $\gamma$ (e.g., Cahn and Allen (1977)). As such, $\gamma$ is also interpreted as a set of phase fields in this work, i.e., for inhomogeneous dislocation motion and glide-system slip.

From a physical point of view, the field relation (24.11) represents a generalized, non-local flow rule for the set $\gamma$ of accumulated glide-system slips. In particular, this relation implies that processes associated with GND development affect “local” or short-range hardening behaviour through the term $K^T\psi_\alpha$. In addition, they influence non-local or longer-range hardening behaviour via the flux-like quantity $J^T\psi_\alpha$. To look into this briefly in more detail, consider for example the form taken by the terms in (24.11) determined by the free energy for the simplest case in the phenomenological context, i.e., the choice $\alpha = (F_p, \gamma, \text{curl} F_p)$.

Then we have

$$\psi(\theta, C, \alpha) = \psi_C(\theta, C, F_p, \gamma, \text{curl} F_p) \quad (24.12)$$

from (24.7) for $\psi$. In turn, this form for $\psi$ results in that

$$-(K^T\psi_\alpha)_a = \tau_a - x_a - r_a,$$

$$J^T\psi_\alpha)_a = F_p^T n_a \times (\psi_C, \text{curl} F_p, \text{curl} F_p) s_a, \quad (24.13)$$

for the local and non-local contributions, respectively, to (24.11) resulting from energy stored in the system. Here,

$$\tau_a := -s_a \otimes n_a \cdot \psi_C, F_p F_p^T \quad (24.14)$$

represents the glide-system Schmidt stress,

$$x_a := s_a \otimes n_a \cdot \psi_C, \text{curl} F_p, (\text{curl} F_p)^T \quad (24.15)$$

a local contribution to glide resistance due to GND development, and

$$r_a := \psi_C, \gamma_a \quad (24.16)$$

one due to SSD development. Note that, in the context of the interpretation of each glide system as an effective dislocation, the combination $\tau_a - x_a$ can be shown to correspond to the effective Peach-Koehler force for glide on this system. Furthermore, as shown by these results, deformation incompatibility contributes in general, i.e., at large deformation, to both the local (24.13) and non-local (24.13) terms, and so the corresponding hardening effects. This is because the rate of $\text{curl} F_p$ depends on both $\dot{\gamma}$ and $\nabla \dot{\gamma}$ from (24.2). As will be seen in the next section, at small deformation, however, the local contribution (24.15) to (24.13) vanishes, leaving only the non-local one.
24.3 The Case of Small Deformation

Intuitively, one might expect the effects of GNDs on the material behaviour in real systems with little initial orientation mismatch to become predominant only at large deformation when misorientation development is most pronounced. From this point of view, all contributions from GNDs should vanish at small deformation. For “small” systems, however, the non-local term (24.13) is not negligible and leads to an additional hardening effect. To discuss this in more detail, consider the specialization of the results from the previous section to small deformation. In particular, in this case, the flow rule (24.1) becomes integrable, yielding the form

\[ F_P = I + \sum_{a=1}^{n} (s_a \otimes n_a) \gamma_a \]  

(24.17)

for \( F_P \), assuming no initial inelastic deformation in the material. This measure is related, e.g., to the slip tensor \( \gamma^{(I)} = \sum_{a=1}^{n} s_a \otimes n_a \) of Shizawa and Zbib (1999). In turn, (24.17) yields the expression

\[ \text{curl} \ F_P = \sum_{a=1}^{n} s_a \otimes (\nabla \gamma_a \times n_a) \]  

(24.18)

for the incompatibility of \( F_P \), related to the dislocation density tensor \( \text{curl} \ \gamma^{(I)} \) of Shizawa and Zbib (1999). On this basis, one obtains in particular the reduced small-strain form

\[ \psi = \psi(\theta, E, \alpha) \]  

(24.19)

for free energy density \( \psi \) analogous to (24.7), as well as that

\[ d = d(\theta, E, \alpha, \nabla \theta, \dot{\gamma}, \nabla \dot{\gamma}) \]  

(24.20)

for the dissipation potential \( d \) analogous to (24.6) for \( d \) at large deformation. Here,

\[ E = \frac{1}{2} (F + F^T) - I \]  

(24.21)

represents the small-strain measure. In turn, (24.19) determines the small-deformation form

\[ T = \psi, E \]  

(24.22)

of the stress analogous to (24.8), and so that

\[ 0 = \text{div} T \]  

(24.23)

for momentum balance. As in the large deformation case, the simplest form of the remaining relations is obtained for the choice \( \alpha = (F_P, \gamma, \text{curl} F_P) \). On the other hand, since \( F_P \) via (24.17) and \( \text{curl} F_P \) via (24.18) are now functions of \( \gamma \) and \( \nabla \gamma \), one actually has \( \alpha = (\gamma, \nabla \gamma) \).

In particular, for this choice, the evolution relation (24.11) then reduces to the set

\[ \frac{\delta \varphi}{\delta \gamma_a} : = \varphi, \gamma_a - \text{div} \varphi, \nabla \gamma_a = 0, \quad a = 1, \ldots, n \]  

(24.24)
of glide system relations in terms of the generalized Euler-Lagrange derivative of the rate potential \( \varphi := \dot{\psi} + d \) from (24.19) and (24.20). In the context of small deformation, then, the generalized the Ginzburg-Landau or Cahn-Allen form of this field relation for \( \gamma \) becomes evident. In particular, this result is referred to as a generalized such form because the dependence of this relation on the free energy in the standard case is generalized here to one on the rate potential \( \varphi \), reflecting the extended modeling of more general inelastic processes relevant here.

Consider lastly the special form of \( \psi \) in the small-deformation context analogous to that (24.12) in the case of large deformation, i.e.,

\[
\psi(\theta, E, \gamma, \nabla \gamma) = \psi_C(\theta, E, \gamma, \nabla \gamma, \text{curl} F_\gamma) ,
\]

where

\[
E_p := \frac{1}{2} (F_p + F_p^T) - I
\]

is the inelastic strain analogous to (24.21). As with (24.12) before, note that (24.25) represents the simplest possible form for \( \psi \) accounting for energy storage due in particular to accumulated inelastic deformation and deformation incompatibility. In the context of (24.25), one obtains in particular the expressions

\[
\begin{align*}
-\psi_{,\gamma a} &= \tau_a + r_a , \\
\psi_{,\nabla \gamma a} &= n_a \times (\psi_{,\text{curl} F_\gamma})^T s_a ,
\end{align*}
\]

for the local and non-local contributions to (24.24) from energy storage via (24.26) and (24.18) by analogy with (24.13). Here, the Schmidt stress takes the form

\[
\tau_a := - s_a \cdot \psi_{,\text{curl} F_\gamma} n_a ,
\]

whereas the flow stress \( r_a \) is still given by (24.16). In the case of small deformation, then, \( x_a \) vanishes identically. As mentioned above, this follows from the fact that, in contrast to the rate of \( \text{curl} F_\gamma \) in the case of large deformation, which depends on both \( \dot{\gamma} \) and \( \nabla \dot{\gamma} \) from (24.2), the time derivative of (24.18) depends only on \( \nabla \dot{\gamma} \). As such, only the non-local contribution (24.27) from deformation incompatibility to hardening behaviour arises here.

### 24.4 Simple Shear of a Crystalline Strip

In order to investigate the consequences of the approach and formulation discussed in the previous section further, consider its application to the case of the (isothermal) simple shear of a crystalline strip containing one or two glide planes. This example has been used in the recent work of Shu, Fleck, Van der Giessen and Needleman (2001) in order to compare the predictions of discrete dislocation computer simulation of Needleman and Van der Giessen (1995) and non-local strain-gradient approach of Fleck and Hutchinson (1993) and applied to crystal plasticity (e.g., Shu and Fleck (1999)). As discussed by them, it represents a model problem for the type of plastic constraint found at grain boundaries of a polycrystal, or the surface of a
thin film, or at interfaces in a composite. To be compatible with the assumptions made in Shu, Fleck, Van der Giessen and Needleman (2001), we work with the model for small deformation from the previous section and assume, (i), isothermal conditions, (ii), elastic isotropy, (iii), linear self-hardening, and (iv), no latent hardening. In particular, these assumptions result in the form

\[
\psi = \frac{1}{2} \lambda (I - E_\varepsilon)^2 + \mu E_\varepsilon : E_\varepsilon + \sum_{a=1}^n \gamma_a + \frac{1}{2} h_0 \gamma_a^2 + g^{-1} \mu \ell^\theta |\text{curl} F_p|^\theta \tag{24.29}
\]

for the free energy density \( \psi \). Here, \( E_\varepsilon := E - E_0 \) represents the (small) elastic strain. In addition, \( \lambda \) and \( \mu \) represent the Lamé constants, \( \gamma_a \) a characteristic glide-system yield stress, and \( h_0 \) a characteristic hardening modulus for linear self-hardening. Further, \( \ell \) and \( g \) represent a characteristic length and exponent, respectively, for energy stored in the material due to local deformation incompatibility. Lastly, the simple power-law form

\[
d = \frac{n}{n+1} \varsigma \gamma_0 \sum_{a=1}^n \left( \frac{\gamma_a}{\gamma_0} \right)^{(n+1)/n} \tag{24.30}
\]

for the dissipation potential \( d \) is relevant. Here, \( \varsigma \) represents a characteristic energy scale for activation of dislocation glide motion with units of \( \text{J m}^{-3} \) or \( \text{Pa} \), and \( \gamma_0 \) a characteristic value of \( \gamma_a \). With the forms for \( \psi \) and \( d \) given, the form of the evolution relation (24.24) for \( \gamma \) is determined. In particular, the simplest form of this relation is obtained for the choices \( g = 2 \) and \( n = 1 \), yielding

\[
\varsigma \gamma_0^{-1} \gamma_a = \mu \ell^2 \sum_{b=1}^n A_{ab} : \nabla(\nabla \gamma_b) + 2\mu s_a : E_\varepsilon n_a - (\gamma_0 + h_0 \gamma_a) , \tag{24.31}
\]

with \( A_{ab} := (s_a : s_b) \cdot [(n_a \cdot n_b) I - n_a \otimes n_b] \). In particular, note that \( A_{ab} : \nabla(\nabla \gamma_a) = [I - n_a \otimes n_a] \cdot \nabla(\nabla \gamma_a) = [s_a \otimes s_a + t_a \otimes t_a] \cdot \nabla(\nabla \gamma_a) \) represents the divergence of \( \nabla \gamma_a \) projected onto the \( ab \)-th glide plane. It is worth emphasizing that the form of this projection results from the dependence of \( \psi \) on \( \text{curl} F_p \). For comparison, note that \( A_{ab} := (s_a : s_b) \cdot (n_a \cdot n_b) I \), and so \( A_{ab} : \nabla(\nabla \gamma_a) = (s_a : s_b) \cdot (n_a \cdot n_b) \text{div}(\nabla \gamma_a) \), would hold if \( \psi \) depended on the inhomogeneity \( \nabla F_p \) instead of on the incompatibility \( \text{curl} F_p \). In the crystal plasticity and current context, at least, the distinction is significant in the sense that no additional hardening results in the current context when \( \psi \) depends directly on \( \nabla F_p \).

The specific geometry and corresponding boundary-value problem for the case of double slip are depicted in Figure 24.1. In particular, the discrete dislocation simulations of this case discussed in Shu, Fleck, Van der Giessen and Needleman (2001) are carried out under the assumption that dislocation motion is blocked at the lower (\( x_2 = 0 \)) and upper (\( x_2 = H \)) boundaries of the strip. Since the development of \( \gamma_a \) is due to dislocation mobility, such boundary conditions are realized here by assuming that \( \gamma_a \) vanishes there. The simulations discussed in Shu, Fleck, Van der Giessen and Needleman (2001), as well as those to be discussed here in what follows, were carried out for plain-strain conditions at an average strain rate of \( \dot{\Gamma} = 10^4 \text{ s}^{-1} \), with \( \dot{\Gamma}(t) = \dot{\Gamma} t \). In addition, they modeled the strip as unbounded in the \( x_1 \) and \( x_3 \) directions via the use of periodic boundary conditions. For simplicity, in the current preliminary study, this situation was approximated by working with a 10x100 finite-element mesh of thickness or height \( H \) and by focusing on conditions in the middle of the strip in
order to minimize the effects of the vertical boundaries. The parameter values used by Shu, 
Fleck, Van der Giessen and Needleman (2001) are representative for aluminum, implying in 
particular $\lambda = 51053$ MPa and $\mu = 26300$ MPa for the elasticity constants. Because 
the respective model relations are quite different, the remaining parameters were chosen for 
the preliminary investigation here in order to obtain approximately the same effective yield 
stress and hardening level as obtained by them for double slip and $H \gg \ell$, yielding for the choices 
g = 2 and $n = 1$ the values $\varsigma = 30$ MPa, $\dot{\gamma}_0 = 2 \times 10^3$ s$^{-1}$, $\eta_0 = 0.5$ MPa, and $h_0 = 50$ MPa.

\[ u_z(t) = H \Gamma(t), \quad u_y = 0, \quad \dot{\gamma}_y = 0, \quad \eta_y = 0 \]

\[ \alpha = 120^\circ, \quad \alpha_y = 60^\circ \]

\[ u_x = 0, \quad u_y = 0, \quad \gamma_x = 0, \quad \gamma_y = 0 \]

Figure 24.1: Geometry and boundary conditions for double slip.

Figure 24.2: Profiles of the shear displacement gradient $u_{1,2}$ for double slip.

Consider first the results for the shear displacement gradient profiles across the strip (Figure 
24.2). The non-uniformity of the deformation is due here to slip being blocked at the upper 
and lower boundaries of the strip, resulting in dislocation pile-up. In both cases, the 
thickness of the boundary layer increases, and the deformation inhomogeneity becomes more 
pronounced, in the range as $H/\ell$ decreases from infinity down to a value of about about 25. 
Interestingly, a further decrease of $H/\ell$ results in a reverse of this trend, leading to a pro-
gressive disappearance of the boundary layer and to the development of uniform shear strain 
across the layer. The increase of the boundary-layer thickness with decreasing strip thickness 
is consistent with the discrete dislocation simulation; the subsequent reversal of the bound-
ary layer development and trend toward uniform shear strain across the layer with decreasing
$H/\ell$, however, is not observed in these, at least for the range of parameters investigated. Also evident in Figure 24.2 is a thickening of the boundary layer with increasing deformation, again for values of $H/\ell$ larger than about 25, after which it begins to disappear. Such a thickening of the boundary layer with increasing deformation is also observed in the discrete dislocation simulations, but not predicted by the non-local strain-gradient-based theory.

Turning next to the dependence of the material behaviour on specimen size, consider the stress-strain behaviour predictions for double slip shown in Figure 24.3. In the case of single slip, for example, dislocation pile-up at the boundaries results in substantial hardening which is rather insensitive to other model details, in agreement with discrete dislocation results. As for the double slip case, the results imply that, at larger strip thicknesses, the additional slip resulting from a second active slip system is able to relax dislocation pile-up and counteract the blockage of slip at the boundaries, resulting in a substantial reduction in the level of hardening relative to the single slip case. As the strip thickness decreases further, however, dislocation pile-up again becomes dominate, resulting again in substantial hardening analogous to the single slip case. Note that this effect correlates with the reduction of boundary-layer thickness with decreasing strip thickness discussed above, and like this latter effect, is not seen in the discrete dislocation simulation for the range of parameters investigated.

Consider finally the dependence of the flow stress on the Hall–Petch ratio $\sqrt{H/\ell}$. Preliminary results imply that, for values of this ratio between 0.3 and 0.5, the dependence of the flow stress on this ratio is linear, reminiscent of the Hall–Petch effect. Below and above this region, however, the scaling appears to become non-linear. In addition, it appears to depend on the level of deformation. Investigation of these and other aspects and consequences of the current formulation, including three-dimensional effects (e.g., parabolic hardening, dislocation microstructure) and those at large deformation, represent work in progress.

References

25 Generalized Continuum Modelling of Single and Polycrystal Plasticity

Samuel Forest

25.1 Introduction

25.1.1 Scope of this Chapter

Several so-called strain-gradient plasticity models are available today for both metal single and polycrystals. They are able to account for some observed size effects like Hall–Petch effects, constrained plasticity in thin layers and around precipitates. These models can be cast in fact into a single thermomechanical framework described in this chapter and based on the notion of generalized continuum. Generalized continua rely on the introduction of additional degrees of freedom or of higher order derivatives of displacement or internal variable. They exhibit non local effects like the influence of constituent size on material behaviour. Modifications of the classical expressions for the power of internal and contact forces and of the dissipation potential are presented in a systematic way for three main kinds of generalized crystal plasticity theories: Cosserat, second gradient and gradient of internal variable approaches (Section 25.2). Extended homogenization techniques are then necessary to deduce the size-dependent polycrystalline behaviour from the knowledge of the nonlocal constitutive behaviour of single crystals. They are reported in Section 25.3. Applications to well-known problems of mechanical metallurgy follow in Section 25.4 for both single and polycrystals.

Each model presentation follows the three following main steps. The principle of virtual power is applied to derive balance equations and appropriate boundary conditions. This requires the choice of the number of degrees of freedom (order of the theory) and the precision (grade of the theory) of the retained theory. The key functions to be defined in each theory are the power density of internal forces \( p^i \) and the power density of contact forces \( p^c \), body forces being excluded here for simplicity. The principle of virtual power then states than the integral of \( p^i \) over any subdomain of a given body \( V \) balances the contribution of contact forces at its boundary. The energy principle introduces the Helmholtz free energy \( \psi \) as a function of state and internal variables. The entropy principle is exploited to derive state laws and intrinsic dissipation \( D \) (Germain et al., 1983).

Throughout this chapter, vectors and tensors of higher order are represented by boldface letters. Scalar quantities are lower case. The dot denotes contraction with respect to one index and \( \otimes \) means tensor product. Index notation according to a Cartesian basis \( e_i \) is used at some places to explain the meaning of intrinsic expressions. The third order permutation tensor is denoted by \( \epsilon \) and its components give the sign of the permutation \( (ijk) \) of \( (1, 2, 3) \). The theories are presented within the small strain and small rotation framework for simplicity.
although the extensions of all results to finite strain plasticity are available for instance in (Forest and Sievert, 2003).

### 25.1.2 Motivations for Generalized Continuum Crystal Plasticity

The motivations for enhanced continuum crystal plasticity theories are to be found in the development of finer descriptions of dislocation populations inside a deformed single crystal volume element \( V \). The necessarily incomplete information about the dislocation state suggests the use of statistical mechanics, as explained in Chapter 21. Kröner, E. (1969) proposed that this information be given in terms of \( n \)-point dislocation correlation tensor functions. If the vectors \( \xi(x) \) and \( b(x) \) describe the line vector and Burgers vector of a dislocation located at \( x \), the first and second correlation functions read:

\[
\alpha = \langle b \otimes \xi \rangle, \quad A(x, x') = \langle (b \otimes \xi)(x) \otimes (b \otimes \xi)(x') \rangle = A(x - x') \tag{25.1}
\]

where the brackets denote ensemble averaging. Statistical uniformity is assumed. Second and fourth rank tensors \( \alpha \) and \( A \) are indeed related to classical plastic state indicators used in classical crystal plasticity. For a large enough volume element \( V \), it is resorted to the ergodic hypothesis so that ensemble averaging is replaced by volume averaging over \( V \). In that case, \( \alpha \) turns out to be identical to the so-called dislocation density tensor, or Nye’s tensor, which is the basic variable of the continuum theory of dislocations (Nye, 1953). The components of \( \alpha \) can be related to the densities of so-called geometrically necessary dislocations (GND, see Chapter 24). On the other hand, one invariant of the tensor \( A \) can be shown to be

\[
A_{ijkl}(0) = b^2 L/V = b^2 \rho \tag{25.2}
\]

where \( L \) is the total length of dislocation lines inside \( V \). The dislocation density \( \rho \) is the usual scalar dislocation density measure in physical metallurgy.

Classical continuum crystal plasticity (see Mandel, 1973 and Chapter 26) relies on the use of internal variables that are more or less related to \( \rho \) rather than \( \alpha \). It has proved to be efficient in describing the main features of the deformation behaviour of single crystals under tensile, shear and, to some extent, non-homogeneous loading conditions. However, quantities \( \alpha \) and \( \rho \) are independent moments of the same distribution and, in principle, should both enter the constitutive framework (Forest et al., 1997). As shown in Chapter 24, the dislocation density tensor can be related to the plastic incompatibility by:

\[
\alpha = -\text{curl} \ f^p = -\epsilon_{jkl} f^p_{ikl} e_i \otimes e_j \tag{25.3}
\]

where the displacement gradient \( f = f^e + f^p \) has been decomposed into elastic and plastic parts. It means that, if \( \alpha \) is introduced as an additional internal variable in the constitutive theory of crystals, it will have a non-local character since it is related to the gradient of the plastic deformation. This is the main incentive for the development of a second gradient theory of crystal plasticity as done by Fleck and Hutchinson (1997). On the other hand, equation (25.3) can be rewritten in terms of the gradient of lattice rotation:

\[
\alpha = \kappa^T - (\text{tr} \ \kappa) \mathbf{1} + \text{curl} \ e^\varepsilon \tag{25.4}
\]
where \( \varepsilon^e \) is the symmetric part of \( f^e \). If \( \Phi \) is the rotation vector associated to the skew-symmetric part of \( f^c \) (so-called lattice rotation vector), the lattice torsion-curvature is defined by

\[
\kappa = \Phi \otimes \nabla = \Phi_{i,j} e_i \otimes e_j
\]  

(25.5)

Accordingly, the manifestation of \( \alpha \) turns out to be the existence of lattice curvature which can be decomposed into elastic and plastic parts \( \kappa^e \) and \( \kappa^p \). The thermodynamic force associated with \( \kappa^e \) is a couple stress tensor that can influence the equilibrium state of the solid and should therefore enter the equation of balance of moment of momentum of the medium. This can be done within the framework of a Cosserat theory, depicted in Section 25.2.1.

Thus, the motivations for the sketched Cosserat and second gradient theories are definitely associated with the notion of dislocation density tensor, or in other words GND. However, GND are not the only source of size effects in plasticity (Kubin and Mortensen, 2002). In particular, the class of models initiated by Aifantis (1987), based on the introduction of the gradient of internal variables like plastic slip or dislocation density, aims at describing the formation of dislocation structures and deformation patterns especially during cycling loading. Such models are presented in Section 25.2.3.

### 25.2 Generalized Continuum Crystal Plasticity Models

Three enhancements of classical crystal plasticity, based on the mechanics of generalized continua, are presented in this section. For simplicity, the formula are written for single slip but the straightforward extension to multislip crystal plasticity can be found in the quoted references. The slip direction and normal to the slip plane are \( l \) and \( n \) respectively. The cross product \( n \times l = \xi \) denotes the line vector of edge dislocations with Burgers vector \( b_l \).

#### 25.2.1 Cosserat Single Crystal Plasticity

The links between crystal plasticity and the Cosserat continuum have been seen since the early stages of the continuum theory of dislocations (McClintock et al., 1958; Mura, 1965; Kröner, 1967). It makes it possible to introduce in a natural way the influence of resistance to lattice curvature on crystal hardening. The degrees of freedom and modeling variables of the Cosserat theory are the displacement \( u \), the micro-rotation \( \Phi \) and their gradients: \( (u, \Phi, u \otimes \nabla, \Phi \otimes \nabla) \). The Cosserat deformation and curvature second rank tensors are defined by

\[
e = u \otimes \nabla + \varepsilon \Phi = e^e + e^p, \quad \kappa = \Phi \otimes \nabla = \kappa^e + \kappa^p
\]  

(25.6)

The power densities of internal and contact forces are used to derive two balance equations, namely balance of momentum and balance of moment of momentum (Eringen, 1999). They involve the second rank tensors of force stresses \( \sigma \) and couple stresses \( M \):

\[
p^i = \sigma : \dot{e} + M : \dot{\kappa}, \quad p^c = t.\dot{u} + m.\dot{\Phi}
\]

(25.7)

\[
div \sigma = 0, \quad div M - e : \sigma = 0, \quad t = \sigma.n, \quad m = M.n
\]  

(25.8)
The traction and couple vectors acting on a surface element are denoted by $t$ and $m$. The free energy $\psi(e^e, \kappa^e, q)$ is a function of elastic deformation and curvature, and internal variable(s) $q$. The associated thermodynamical forces are:

$$
\sigma = \rho \frac{\partial \psi}{\partial e^e}, \quad M = \rho \frac{\partial \psi}{\partial \kappa^e}, \quad R = \rho \frac{\partial \psi}{\partial q}
$$

(25.9)

This leads to the following intrinsic dissipation:

$$
D = \sigma : \dot{e}^p + M : \dot{\kappa}^p - R \dot{q}
$$

This constitutes equations require two symmetric four-rank tensors of elastic moduli, $C$ and $D$ having the dimension of $MPa$ and $MPa.mm^2$ respectively:

$$
\sigma = C : e, \quad M = D : \kappa
$$

(25.10)

The plastic flow rules can be derived from the choice of a dissipation potential depending on all thermodynamical forces. Following (Forest et al., 1997), the dissipation potential is split into two contributions $f_1(\sigma, R, q)$ and $f_2(M, R_c, q_c)$. This corresponds to a multi-criterion framework involving two plastic multipliers:

$$
\dot{\epsilon}^p = \lambda_1 \frac{\partial f_1}{\partial \sigma}, \quad \dot{\kappa}^p = \lambda_2 \frac{\partial f_2}{\partial M}, \quad \dot{q} = -\lambda_1 \frac{\partial f_1}{\partial R}
$$

(25.11)

Note that the choice of the proper plastic multiplier to be used in the evolution equation for hardening variable (25.11) depends on the specific type of hardening laws and number of internal variables. A possible choice for the potentials is:

$$
f_1 = |\tau| - R, \quad f_2 = |m| - l_p R_c, \quad \tau = \sigma : (l \otimes n), \quad m = M : (\xi \otimes l)
$$

(25.12)

where $l_p$ is a characteristic length. The yield function $f_2$ represents an extension of Schmid law $f_1$ to couple-stresses. If the resolved couple-stresses $m$ reaches a critical value $l_p R_c$, plastic curvature will develop with respect to two $\xi$ and $l$. The development of torsion accommodated by arrays of screw dislocations is also possible and is described in (Forest et al., 1997). Accordingly, this model requires at least two hardening variables $q$ and $q_c$, the thermodynamical forces of which being $R$ and $R_c$. A coupling must exist between hardening variables for the model to give rise to size effects, as explained in the example of Section 25.4.3. Within the present Cosserat model, the flow rules become:

$$
\dot{\epsilon} = \gamma^p l \otimes n, \quad \dot{\kappa}^p = \frac{\dot{\theta}}{l_p} \xi \otimes l, \quad \dot{\lambda}_1 = \gamma^p \text{sign } \tau, \quad \dot{\lambda}_2 = \frac{\dot{\theta}}{l_p} \text{sign } m
$$

(25.13)

The meaning of the second plastic multiplier therefore is that of a scalar curvature rate, i.e. the ratio of a rotation rate by a characteristic length.

Until now, the physical meaning of micro-rotation $\Phi$ has not been specified. It is in fact intended to coincide with the notion of crystal lattice rotation, so that the Cosserat directors are simply lattice directions defined in the released isoclinic configuration introduced in (Mandel, 1973). The definitions of Cosserat relative deformation (25.6) and plastic deformation rate (25.13) imply that the skew-symmetric part of elastic deformation $e^e$ measures the difference between lattice and Cosserat rotations. The constraint that $e^e$ be symmetric must therefore be added for the microrotation to have the wanted physical meaning. This can be achieved by choosing very high elastic constants linking the skew-symmetric parts of stress and of elastic deformation. Skew-symmetric reaction stresses then remain.
25.2.2 Second Gradient Single Crystal Plasticity

Following the second gradient modelling framework settled by Mindlin (see (Mindlin and Eshel, 1968)), Fleck and Hutchinson (1997) have proposed a crystal plasticity model incorporating characteristic lengths associated with the effect of so-called geometrically necessary dislocations, and also of gradient of slip normal to the slip plane. The degrees of freedom and modeling quantities of the theory are the displacement field and its first and second gradients: \( (u, u \otimes \nabla, u \otimes \nabla \otimes \nabla) \). The corresponding strain measures are the second order strain tensor \( \varepsilon \), with its classical definition, and its gradient \( \eta = \varepsilon \otimes \nabla \). They are split into elastic and plastic contributions:

\[
\varepsilon = \varepsilon^e + \varepsilon^p, \quad \eta = \eta^e + \eta^p
\]

(25.14)

Note that it is not assumed that \( \eta^p \) is equal to the gradient of plastic strain: **plastic strain gradient is not the gradient of plastic strain** in the theory developed in (Fleck and Hutchinson, 1997) (see (Forest and Sievert, 2003) for an alternative gradient of plastic strain theory).

Two stress tensors are associated with the previous strain measures and appear in the power densities of internal and contact forces:

\[
p^i = \sigma : \dot{\varepsilon} + M : \dot{\eta}, \quad p^c = t.\dot{u} + m.D_n.\dot{u}
\]

(25.15)

A detailed derivation of these equations can be found in Germain (1973). \( G \) is the mean curvature of the surface at point \( x \). The free energy density is a function of elastic strain and strain gradient and possibly of hardening variable(s) \( q \): \( \psi(\varepsilon^e, \eta^e, q) \). The state laws are deduced from the analysis of the entropy principle:

\[
\sigma = \rho \frac{\partial \psi}{\partial \varepsilon^e}, \quad M = \rho \frac{\partial \psi}{\partial \eta^e}, \quad R = \rho \frac{\partial \psi}{\partial q}
\]

(25.18)

The remaining intrinsic dissipation reads: \( D = \sigma : \dot{\varepsilon} + M : \dot{\eta}^p - R \dot{q} \). Fleck and Hutchinson (1997) propose a simplified version of elasticity constitutive equations including the usual four-rank tensor of elasticity moduli \( C \) and an additional intrinsic length scale \( l_e \):

\[
\sigma = C : \varepsilon^e, \quad M = l_e^2 C : \eta^e
\]

(25.19)

The dissipative processes are modeled by a potential function \( f(\sigma, M, R) \) of the thermodynamical forces:

\[
\dot{\varepsilon}^p = \dot{\lambda} \frac{\partial f}{\partial \sigma}, \quad \dot{\eta}^p = \dot{\lambda} \frac{\partial f}{\partial M}, \quad \dot{q} = -\dot{\lambda} \frac{\partial f}{\partial R}
\]

(25.20)
25.2 Generalized Continuum Crystal Plasticity Models

where \( \dot{\lambda} \) is the plastic multiplier. The theory presented in (Fleck and Hutchinson, 1997) considers the general case of viscoplasticity, we translate it here in the particular case of rate-independent plasticity.

The usual Schmid law is now extended to incorporate not only resolved shear stresses but also resolved hyperstresses, related by a constitutive length \( l_p \):

\[
f = |\tau| + |m| l_p = R, \quad \text{with} \quad \tau = \sigma : (l \otimes n), \quad m = M : (l \otimes n \otimes l) \quad (25.21)
\]

The plastic strain and strain gradient rates follow from the normality rules (25.20):

\[
\dot{\varepsilon}^p = \dot{\gamma}^p l \otimes n, \quad \dot{\eta}^p = \dot{\gamma}^S l \otimes n \otimes l \quad (25.22)
\]

For the sake of simplicity, we have introduced only the effect of slip gradient in the slip direction \( l \), although the original model also contains a contribution of slip gradient in the normal direction \( n \). The first type of strain gradient effect is associated with the density of so-called geometrically necessary dislocations, the second one remains unclear. It must be noticed that the slip gradient variable \( \lambda^S \) does not coincide with the gradient of slip \( \gamma^p \otimes \nabla \). Both are related to the plastic multiplier \( \dot{\lambda} \) according to (25.20):

\[
\dot{\lambda} = \dot{\gamma}^p \text{sign} \tau = l_p \dot{\gamma}^S \text{sign} m.
\]

For monotonous loading, the slip gradient variable \( \gamma^S \) turns out to be equal to the amount of slip \( \gamma^p \) divided by \( l_p \).

25.2.3 Gradient of Internal Variable Approach

The pioneering model of Aifantis (1987) introduces the gradient of an internal variable into the modelling. The thermomechanical setting of the theory can be shown to have a form very similar to that of the previous Cosserat and second gradient theories. In particular, additional boundary conditions are also necessary and are clarified here. In the previous Aifantis models, the internal variable is the scalar dislocation density (so-called density of statistically stored dislocations) or the cumulative plastic strain. When applied to the slip variable \( \gamma^p \), so indirectly to a density of geometrically necessary dislocations, it builds a link with the models of previous sections. In particular, it appears that the gradient of internal variable approach gives rise, as a special case, to the well-known Laplacian term \( \Delta \gamma^p \), the physical meaning of which remains to be explained in each specific situation. This model remains different from the second grade and Cosserat models in so far as it is not directly related to the dislocation density tensor.

The same structure of model presentation as for the second grade and Cosserat models is kept in this section, so that direct comparison is possible at each step. The degrees of freedom and modeling quantities are the displacement, the internal variable \( \gamma^p \) and their gradients: \((u, \gamma^p, u \otimes \nabla, \gamma^p \nabla) = (\gamma^p, \epsilon_i)\). The classical strain tensor \( \epsilon \) is decomposed into elastic and plastic contributions as usual. However, the power density of internal forces must be extended as a linear form of the modeling quantities:

\[
p^i = \sigma : \varepsilon + \alpha \gamma^p + B : \gamma^p \nabla, \quad p^c = t \cdot u + \alpha^c \gamma^p \quad (25.23)
\]

A generalized stress vector \( B \) must be introduced as the dual quantity of the internal variable gradient. This leads to two sets of balance equations and associated boundary conditions:

\[
\text{div} \sigma = 0, \quad \alpha = \text{div} B, \quad t = \sigma \cdot n, \quad \alpha^c = B \cdot n \quad (25.24)
\]
The term $\alpha \dot{\gamma}^p + B \cdot \nabla \dot{\gamma}^p = \text{div} (\dot{\gamma}^p B)$ represents a nonlocal power of internal forces of first degree. The second balance equation $\alpha = \text{div} B$ can also be interpreted as a definition of $\alpha$. The value of $\gamma^p$ or the dual force $\alpha^c$ must be prescribed at the boundary of the considered solid. The free energy is a function of elastic strain, $\gamma^p \nabla$ and internal variables $q$ (including $\gamma^p$ if necessary): $\psi(\varepsilon^e, \gamma^p \nabla, q)$. The state laws and intrinsic dissipation take the form:

$$\sigma = \rho \frac{\partial \psi}{\partial \varepsilon^e}, \quad B = \rho \frac{\partial \psi}{\partial \gamma^p \nabla}, \quad R = \rho \frac{\partial \psi}{\partial q}, \quad D = \sigma : \dot{\varepsilon}^p + \alpha \dot{\gamma}^p - R \dot{q} \quad (25.25)$$

The classical elastic constitutive relations are unchanged. A potential function $f(\sigma, \alpha, R)$ is used to derive the flow and hardening rules:

$$\dot{\varepsilon}^p = \dot{\lambda} \frac{\partial f}{\partial \sigma}, \quad \dot{\gamma}^p = \dot{\lambda} \frac{\partial f}{\partial \alpha}, \quad \dot{q} = -\dot{\lambda} \frac{\partial f}{\partial R} \quad (25.26)$$

In a case of a single crystal undergoing single slip, the yield criterion is again an extension of Schmid law, from which the flow rules are deduced:

$$f = |\tau + \alpha| - R = |\tau + \text{div} B| - R, \quad \text{with} \quad \tau = \sigma : (l \otimes n) \quad (25.27)$$

The free energy can be taken as a quadratic function of $\gamma^p \nabla$, so that, in the case of isotropic or cubic material symmetry: $B = c \gamma^p \nabla$, $\text{div} B = c \Delta \gamma^p$. Under monotonous loading conditions such that $\tau$ remains positive, the well-known relation found in Aifantis models is retrieved ($c > 0$): $\tau = R - c \Delta \gamma^p$.

### 25.3 From Single to Polycrystals: Homogenization of Generalized Continua

The polycrystal can now be regarded as a heterogeneous Cosserat or strain gradient material since it is an aggregate of Cosserat or strain gradient single crystal grains. As a result, some homogenization procedures must be designed to study the resulting properties of the polycrystal and extend standard homogenized polycrystal models like Taylor and self-consistent schemes (see Chapters 22 and 23). The first question is to determine the nature of the resulting homogenized medium: Is it still a Cosserat or strain gradient continuum or something else? The multiscale asymptotic method is a reliable technique to answer this general question. It has been designed for periodic microstructures but the main results can in fact be shown to be valid also for random materials using powerful mathematical convergence theorems.

#### 25.3.1 Introduction to Multiscale Asymptotic Method

In the case of periodic microstructures, a unit cell $V$ can be defined and asymptotic methods are well-adapted for deriving the form of the effective balance and constitutive equations (Sanchez-Palencia, 1974). The key-point is the choice of a small parameter $\varepsilon$ introduced in the multiscale asymptotic developments. Two different choices are possible in the case of periodic
heterogeneous Cosserat media. Three characteristic lengths must be distinguished: the size $l$ of the unit cell $V$, a typical characteristic length $l_c$ of the constituents of the heterogeneous Cosserat material, and a typical wave length $L_\omega$ associated with the applied macroscopic loading conditions. In classical homogenization theory, one usually assumes that $l \ll L_\omega$ and, in this case, the small parameter is $\varepsilon = l / L_\omega$. In the present situation, one may first consider a limiting process $\mathcal{H}^I$ with $\varepsilon \rightarrow 0$, for which the Cosserat length scale $l_c$ varies in the same way as $l$, so that the small parameter can also be written as $\varepsilon = l_c / L_\omega$. In this case, the effective medium can be shown to be a Cauchy continuum (Forest et al., 2001a). According to a second limiting process $\mathcal{H}^{II}$, $l_c$ is kept constant, which corresponds to $\varepsilon = l / l_c$. The main fields are now treated as functions of two variables: the macroscopic space variable $x$ and the local one $y = x / \varepsilon$. The local fields can be expanded in power series of $\varepsilon$,

$$u^\varepsilon(x) = u^0(x, y) + \varepsilon u^1(x, y) + \varepsilon^2 u^2(x, y) + \cdots$$  \hspace{1cm} (25.29)

$$\Phi^\varepsilon(x) = \Phi^0(x, y) + \varepsilon \Phi^1(x, y) + \varepsilon^2 \Phi^2(x, y) + \cdots$$  \hspace{1cm} (25.30)
where the $u^i$ and $\Phi^i$ are assumed to have the same order of magnitude and are periodic in $y$. Similar expansions exist for the force and couple stresses. The form of the constitutive equations is different for each homogenization procedure. In the case of linear elasticity, they read:

$$\mathcal{H}^I: \quad \sigma^\varepsilon = C(y) : e^\varepsilon(x) \quad \text{and} \quad M^\varepsilon = \varepsilon^2 D(y) : \kappa^\varepsilon(x)$$  \hspace{1cm} (25.31)

$$\mathcal{H}^{II}: \quad \sigma^\varepsilon = C(y) : e^\varepsilon(x) \quad \text{and} \quad M^\varepsilon = D(y) : \kappa^\varepsilon(x)$$  \hspace{1cm} (25.32)

After noting that $\nabla = \nabla_x + 1/\varepsilon \nabla_y$ (with obvious notations), we compute the gradient of the kinematic variables and the divergence of the stresses in order to introduce them in the balance equations (25.8) and in the constitutive equations. Ordering the terms according to powers of $\varepsilon$ leads to a set of equations from which a series of auxiliary boundary value problems to be solved on the unit cell can be defined (see (Boutin, 1996) for the case classical periodic homogenization). The first auxiliary problem for the procedure $\mathcal{H}^{II}$ consists in determining vector fields $v$ and $\psi$ such that:

$$u = E.y + v \quad \text{and} \quad \Phi = K.y + \psi, \quad \forall y \in V$$  \hspace{1cm} (25.33)

$$\sigma = C : (u \otimes \nabla_y) \quad \text{and} \quad M = D : (\Phi \otimes \nabla_y)$$  \hspace{1cm} (25.34)

$$\text{div}_y \sigma = 0 \quad \text{and} \quad \text{div}_y \mu = 0$$  \hspace{1cm} (25.35)

where $v$ and $\psi$ take the same values on opposite sides of the cell and the traction and surface couple vectors $\sigma.n$ and $\mu.n$ are anti-periodic. The solution of this problem gives in fact the terms $u_1, \Phi_1, \sigma^0$ and $M^0$ of the expansions.

Finally, one computes the mean work of internal forces:

$$\langle \sigma : e + \mu : \kappa \rangle = \langle \sigma \rangle : \langle e \rangle + \langle \mu \rangle : \langle \kappa \rangle$$  \hspace{1cm} (25.36)

This defines the effective (macroscopic) deformation, curvature, force and couple stress tensors as mean values of the local ones. This expression has the structure of the work of internal forces of an effective Cosserat medium. It is the extension of the classical Hill-Mandel condition of macro-homogeneity to heterogeneous Cosserat media. In contrast, the homogenization procedure $\mathcal{H}^I$ provides a classical Cauchy macroscopic continuum due to the $\varepsilon^2$ term in equation (25.31) (Forest et al., 2001a).

Asymptotic methods are difficult to extend properly to the case of nonlinear local behaviour of phases. We propose here a formulation of a boundary value problem on a given polycrystalline volume element with periodic boundary conditions, similar to the auxiliary problem (25.33) to (25.35), in which the linear constitutive equations (25.34) are replaced by the unaltered nonlinear ones. This problem is solved on a single unit cell $V$. Alternative Dirichlet conditions providing the same energy transfer (25.36) can be deduced from the previous initial boundary value problem: $u = E.x, \quad \Phi = K.x \quad \forall x \in \partial V$. They provide the same definitions of the macroscopic Cosserat deformation, curvature, stress and couple-stress tensors. Tensile boundary conditions have been prescribed to the polycrystalline volume element of figure 25.1(a) made of 50 3D zinc grains. The Figure 25.1(b) shows the activation of basal slip inside the grains. A internal slice of the aggregate is shown on Figures 25.1(c) and (d)illustrating the the lattice curvature field by a norm of $\kappa$. It can be seen that plastic curvature preferentially develops near grain boundaries. Larger polycrystalline volume elements are studied in (Barbe et al., 2001).
25.3.2 Extension of Classical Homogenization Schemes

Classical homogenization schemes based on variational formulations like the Hashin-Shtrikman bounds for effective properties can be extended to size-sensitive material models. This is done by Smyshlyaev and Fleck (1994) and Smyshlyaev and Fleck (1996) in the case of heterogeneous couple-stress medium. The couple-stress medium is a constrained Cosserat continuum for which the Cosserat rotation coincides with material rotation.

The extension of estimation methods of the effective properties like the self-consistent or Mori-Tanaka schemes, requires the solving of the problem of the inclusion in an infinite matrix when inclusion and matrix are generalized continua. Let us consider a spherical Cosserat elastic inclusion having elastic properties \((E^I, \nu^I, \mu^I_c, l^I_c)\) embedded in an infinite matrix endowed with different moduli \((E, \nu, \mu_c, l_c)\). The matrix is loaded at infinity and a tensile test is considered for example. For a given set of Cosserat elastic parameters for the inclusion and the matrix, the influence of the inclusion size \(R\) on the mean stress within the inclusion was investigated using the finite element method (Forest et al., 1999; Forest, 2001). For a large ratio \(R/l_c\), the solution tends towards the classical one. For lower ratios, in the case of an inclusion more rigid than the matrix, higher values of the mean stress are reached. Again, the stress tends to an asymptotic value for ever decreasing inclusion sizes. In the transition domain, generalized stresses and deformations are in fact non-homogeneous within the inclusion, contrary to the classical Eshelby case. The characteristic length of the Cosserat medium acts in fact as a resolution parameter (as in microscopy). When the geometry interacts with this length scale, size effects will be observed.

![Figure 25.2: Deformed states of a two-phase material element under simple shear: (a) initial state; (b) solution according to classical crystal plasticity; (c) Cosserat elastoplasticity. The solutions have been computed using the finite element method and each phase is divided into elements: hard phase in white, soft phase in gray; the mean deformation is the same in (b) and (c).]
25.4 Simulations of Size Effects in Crystal Plasticity

This Section mentions recent applications of Cosserat and strain gradient crystal plasticity to handle size effects observed in metals.

25.4.1 Constrained Plasticity in Two-Phase Single Crystals

The case of plastic flow in thin layers has been investigated using dislocation models (theoretical and numerical). These solutions can be compared to the responses of generalized crystal plasticity models. The collective behaviour of dislocations in narrow channels in multiphase materials or films is associated with size effects like the dependence of Orowan stresses on channel width (Ashby, 1971). This situation is encountered for instance in PSB of fatigued single crystals or in the $\gamma$-channels of nickel-base single crystal superalloys. In (Forest and Sedláček, 2003), an idealized laminate microstructure is considered under shear, for which one phase remains elastic whereas the second one deforms plastically. A simple dislocation analysis shows that plastic strain is not uniform due to the bowing of screw dislocations and the pileup formation near the interface. In contrast, classical crystal plasticity predicts uniform deformation, which is the actual limit state for thick layers. Due to the existence of interface conditions, deformation is not uniform when using a Cosserat or strain gradient model, and some additional parameters can be identified explicitly looking at the dislocation solution. This is illustrated in Figure 25.2 by finite element simulation of the periodic two-phase material. Similarly, boundary layers in a sheared single crystal layer has been analysed in (Shu et al., 2001) using both second gradient theory and discrete dislocation dynamics. The additional interface or boundary conditions play a central role in the development of non-homogeneous deformation in non-local models, so that explicit models including higher order gradients but keeping the standard structure of the boundary value problem as in (Acharya and Bassani, 2000), are not able to account for it. Continuum modelling of dislocation pile-ups around precipitates is proposed in Forest et al. (2000). A similar situation is investigated in (Bassani et al., 2001) combining nonlocal crystal plasticity and dislocation dynamics.

The case of PSB formation in fatigued crystals is adequately handled by Aifantis model involving dislocation densities (Aifantis, 1987), because this phenomenon has a priori nothing to do with the concept of geometrically necessary dislocations.

25.4.2 Plasticity at the Crack Tip in Single Crystals

The stress-strain field at the crack tip in a single crystal can be simulated using dislocation dynamics and continuum crystal plasticity (Van der Giessen et al., 2001). The problem of the crack in an infinite medium has been addressed very often during the development of the mechanics of generalized continua in the 60s and 70s. Solutions are available for the elastic couple stress medium and fully non local elasticity. In the case of crystal plasticity, numerical solutions with the Cosserat continuum have been compared to Rice’s solution for the classical continuum in (Forest et al., 2001b). The classical solution, derived for an elastoplastic single crystal without hardening, predicts the existence of sectors with constant stress separated by localization bands. The physical nature of the bands depends on crystal structure and orientation. Intense slip bands, kink bands and multislip bands can coexist. The use of the Cosserat
continuum leads to the weakening and even elimination of kink bands at the crack tip when additional hardening associated to lattice curvature is introduced. The Figure 25.3 shows the profiles of equivalent plastic slip along half a circle close to the crack tip for three different values of extra-hardening modulus $H$ assuming a simple extra-hardening model linear in the variable $\theta$ (see Equation (25.13)). For the considered crystal orientation, kink bands lie at 55° and 125° from the horizontal axis, whereas the vertical band is a pure slip band. The intensity of the kink bands decreases for increasing $H$ values. Note that asymptotic analyses of the crack tip field are possible within classical crystal plasticity that exclude kink banding (Drugan, 2001) but they are difficult to implement numerically. Recent experimental investigations on nickel-base superalloys at room temperature confirms in fact the existence of kink bands at the crack tip under monotonous loading (cf. Figure 25.4 and reference (Flouriot et al., 2003)). Extra-hardening parameter $H$ can be identified from the found intensity of the kink bands.

The use of physically more realistic models at the crack tip than classical crystal plasticity is an important issue for the prediction of subsequent crack growth, especially in fatigue (Deshpande et al., 2001).

### 25.4.3 Grain Size Effects in Polycrystalline Aggregates

The generalized homogenization methods presented in Section 25.3 can be used to reproduce Hall–Petch effects starting from reliable hardening rules at the single crystal level. The transition in the response of aggregates made of large to ultra-fine grains is therefore possible. The case of polycrystalline aggregates with prediction of the influence of grain size on the
Figure 25.4: Slip lines at the crack tip of a nickel-base superalloy single crystal in a CT specimen. The crack plane is (110) and the crack propagation direction is [001].

tensile curves is tackled in (Forest et al., 2000) for Cosserat elastoplasticity, and in (Acharya and Beaudoin, 2000) for viscoplasticity. In the latter case, the authors resort to a model with explicit use of higher order gradients in the hardening law but a classical structure of the boundary value problem. This is sufficient for this type of applications.
25.5 Conclusion

The validity range of generalized continuum crystal plasticity models is limited. They can be used to tackle crystal plasticity problems up to grain or precipitate sizes of 1 to 0.1 µm depending on the dislocation density. It bridges the gap between classical crystal plasticity and dislocation dynamics. The case of nano-grains may well remain out of the range of application of such models. However, extensions of continuum models that mimic the response of discrete systems (atomistics or dislocation dynamics) are still needed.

References

References


26 Micro-Mechanical Finite Element Models for Crystal Plasticity

Surya R. Kalidindi

26.1 Introduction

“Crystal Plasticity” refers to study of plastic deformation in single crystal and polycrystalline materials while taking into account explicitly the details of physics and geometry of deformation at the crystal (also called grain) level. At low homologous temperatures, the dominant mode of plastic deformation in crystalline materials is slip on specific crystallographic planes in specific directions. Crystal plasticity aims to incorporate these and other such details at the grain scale in the description and formulation of elastic-plastic constitutive models for single crystals and for polycrystalline materials.

Detailed accounts of the development of the crystal plasticity theory and its many successes (especially in predicting anisotropic stress-strain response of high stacking fault energy polycrystalline cubic metals together with the concurrent evolution of the underlying averaged texture) have been presented in other chapters. In many of these models, simplifying assumptions are routinely made in making the transition from the response of a constituent single crystal to the response of the representative polycrystalline aggregate. A typical consequence of such assumptions is that the solutions for the polycrystals often violate either equilibrium or compatibility or both. Consequently, although these simplistic models provide reasonably accurate predictions in the averaged sense at the polycrystal level, their predictions at the constituent single crystal level can be quite erroneous.

With the advances made in finite element techniques, it is now possible to build micro-mechanical polycrystalline finite element models that can provide accurate solutions of stress and strain fields in a given polycrystalline aggregate that automatically satisfy both equilibrium and compatibility (in a weak numerical sense). Such models therefore raise the exciting possibility of predicting accurately the underlying physical phenomena occurring at the grain scale. This chapter expounds on the development and use of micro-mechanical finite element models for elastic-plastic response of polycrystalline aggregates.

26.2 Theoretical Background

The necessary theoretical background is briefly reviewed next. For further details, the reader is referred to Chapter 5 and to the references listed therein.
26.2.1 Crystal Plasticity Framework

For finite deformations, one starts with a multiplicative decomposition of the total deformation gradient \( F \) into elastic and plastic components (Asaro and Rice 1977) as

\[
F = F^*F^p \tag{26.1}
\]

where \( F^* \) contains deformation gradients due to both elastic stretching as well as the lattice rotation, while \( F^p \) denotes the deformation gradient due to plastic deformation alone. Of particular importance is the fact that the hypothetical intermediate relaxed configuration defined by the above decomposition is an isoclinic configuration, i.e. the lattice orientation in this intermediate configuration is assumed to be identical to that in the initial configuration.

The constitutive equation for stress in the crystal can be expressed (Kalidindi et al. 1992) as

\[
\begin{align*}
T^* &= CE^*, & T^* &= F^*^{-1}\{\det F^*\}F^*^{-T}, & E^* &= \frac{1}{2}\left\{F^*^TF^* - 1\right\} \\
\end{align*}
\tag{26.2}
\]

where \( C \) is the fourth-order elasticity tensor reflecting the inherent anisotropy of the single crystal, \( T^* \) and \( E^* \) are a pair of work conjugate stress and strain measures, and \( T \) is the Cauchy stress in the crystal. The evolution of \( F^p \) can be expressed in a rate form as

\[
\dot{F}^p = LF^p, \quad L = \sum_\alpha \dot{\gamma}^\alpha S^\alpha_o, \quad S^\alpha_o = m^\alpha_o \otimes n^\alpha_o \tag{26.3}
\]

where \( \dot{\gamma}^\alpha \) is the shearing rate on the slip system \( \alpha \), and \( m^\alpha_o \) and \( n^\alpha_o \) denote the slip direction and the slip plane normal of the slip system, respectively, in the initial unloaded configuration. The shearing rate on the slip system is dependent on the resolved shear stress \( \tau^\alpha \) on the slip system and the slip resistance \( s^\alpha \) of that slip system, and can be expressed in a power-law relationship (Asaro and Needleman 1985) as

\[
\dot{\gamma}^\alpha = \gamma_0 \left| \tau^\alpha \right|^{1/m} \text{sgn}(\tau^\alpha), \quad \tau^\alpha \approx T^* \cdot S^\alpha_o, \tag{26.4}
\]

where \( \gamma_0 \) denotes a reference value of the slip rate and \( m \) represents the strain rate sensitivity parameter. The evolution of the slip system resistance can be expressed in the following general form

\[
\dot{s}^\alpha = \sum_\beta h^\alpha\beta \left| \gamma^\beta \right|, \tag{26.5}
\]

where \( h^\alpha\beta \) describes the hardening of slip system \( \alpha \) due to slip activity on slip system \( \beta \).

26.2.2 Total Lagrangian vs. Updated Lagrangian Schemes

An important distinction in the numerical approach used here (compared to those described in other chapters) is the use of a total Lagrangian scheme for the integration of the constitutive relations, as opposed to an updated Lagrangian scheme. In deed, this is the main reason
why the theoretical framework presented above is formulated somewhat different from that presented in the other chapters. Figure 26.1 illustrates schematically the main features of both the total Lagrangian and the updated Lagrangian methods. The updated Lagrangian scheme is
centered around updating the end configuration of a given time step to become the initial configuration for the next time step. For example, in Figure 26.1(a), configuration corresponding to time $t$ serves as the initial configuration for the time step $(t, \tau)$. As one marches forward in an updated Lagrangian scheme, there is no need to carry forward the information related to the initial configuration corresponding to the beginning of the last time step. On the other hand, a total Lagrangian scheme continues to use always the initial configuration corresponding to time zero as the reference configuration, and carries forward only the necessary state variables from the end of one time step to the next. The state variables contain enough information to define uniquely the relaxed configuration at the end of each time step. There are several advantages of using a total Lagrangian scheme. For example, by using total Lagrangian schemes in crystal plasticity models, one can avoid the need to compute the current lattice orientation of the crystal at the end of each time step; instead the orientation of the crystal can be computed only when it is required. Also, the total Lagrangian scheme for crystal plasticity model described in Section 2.1 allows easy formulation of a fully implicit time integration procedures (Kalidindi et al. 1992). Finally, although deformation twinning as a mode of plastic deformation is not addressed in this chapter, it has been demonstrated that the total Lagrangian scheme offers significant advantages when deformation twinning needs to be incorporated into crystal plasticity models (Kalidindi 1998; see also Chapter 27).

### 26.2.3 Fully Implicit Time Integration Procedures

The computational problem at hand is posed as one with fully prescribed deformation history. However, it can be easily extended to situations where only certain components of the deformation gradient history are prescribed along with certain other components of the stress history. It is assumed that at any time $t$, the deformation gradient $F(t)$ and the list of variables $\{F^p(t), s^\alpha(t)\}$ are known at a given material point in a single crystal region. Also given is the deformation gradient, $F(\tau)$, after a small increment of time $\Delta t$ (i.e. $\tau = t + \Delta t$). The objective of the computational scheme is to find the Cauchy stress $T(\tau)$ and update the list of variables $\{F^p(\tau), s^\alpha(\tau)\}$ at the given material point.

Fully implicit time integration of Equation (26.3a) leads to (Kalidindi et al. 1992)

$$F^p(\tau) \equiv \{1 + \Delta t \mathbf{L}^p(\tau)\} F^p(t), \quad (26.6)$$

and

$$F^{p-1}(\tau) \equiv F^{p-1}(t) \{1 - \Delta t \mathbf{L}^p(\tau)\}. \quad (26.7)$$

For brevity of notation, Equation (26.3b) may be conveniently expressed as

$$\Delta t \mathbf{L}^p(\tau) = \sum_{\alpha} \Delta \gamma^\alpha S^\alpha_\sigma \quad (26.8)$$

In Equation (26.8), $\Delta \gamma^\alpha$ is a function of the resolved shear stresses and the slip resistances defined by equations (26.4) and (26.5). Therefore, it is reasonable to express $\Delta \gamma^\alpha$ as

$$\Delta \gamma^\alpha = \Delta \hat{\gamma}^\alpha (T^*(\tau), s^\alpha(\tau)) \quad (26.9)$$
26.2 Theoretical Background

Substituting Equations (26.7), (26.8), (26.1) and (26.2c) in Equation (26.2a), we can derive (Kalidindi et al. 1992) the following equation for $T^*(\tau)$:

$$T^*(\tau) \equiv T^{*\text{tr}} - \sum_{\alpha} \Delta \gamma^\alpha C^\alpha$$

(26.10)

where

$$T^{*\text{tr}} = C \left[ \frac{1}{2} \{A - 1\} \right]$$

(26.11)

$$C^\alpha = C \left[ \frac{1}{2} B^\alpha \right]$$

(26.12)

$$A = F_{\text{p}}^{-T}(t) F^T(\tau) F(\tau) F_{\text{p}}^{-1}(t)$$

(26.13)

$$B^\alpha = A S^\alpha_o + S_o^{\alpha T} A$$

(26.14)

For the case where the slip systems are assumed to be non-hardening, Equation (26.10) represents a set of six non-linear equations with six unknowns. The non-linearity is due to the dependence of $\Delta \gamma^\alpha$ on $T^*(\tau)$. Equation (26.10) can be solved for the stress components using a modified Newton-type algorithm (Kalidindi et al. 1992).

For situations involving hardening of the slip systems, additional equations have to be formulated by integrating Equation (26.5). However, it is suggested that Equation (26.10) be solved first using the best guesses of the various slip resistances and then to solve for the new values of the slip resistances using a two-level iterative procedure. Such a two level iterative procedure was found to work very well for the type of models described here. After solving for $T^*(\tau)$, and updating the slip resistances $\{s^\alpha(\tau)\}$, $\Delta \gamma^\alpha$ can be computed using Equations (26.4) and (26.5), $F_{\text{p}}(\tau)$ using Equation (26.6), $F^*(\tau)$ using Equation (26.1), $E^*(\tau)$ from Equation (26.2c), and finally $T(\tau)$ from Equation (26.2b). These calculations mark the end of the time-step, and by repeating these calculations successively, one can march forward in time and simulate a large deformation process. Note that there is no need to update the crystal orientations at the end of each time step, since the initial configuration is kept fixed in these calculations (corresponding to time zero).

At any point during the simulation of a given deformation process, the lattice orientations in the current configuration can be computed using the rotation component $R^*$, defined through polar decomposition of $F^*$. Let $[g]$ represent the coordinate transformation matrix from a frame aligned with the initial crystal lattice to the fixed global reference frame, i.e.

$$\{m\}^g = [g] \{m\}^c$$

(26.15)

where $\{m\}^c$ represents a specific crystal direction in the crystal coordinate frame and $\{m\}^g$ describes the same vector in the global reference frame. The rotation $R^*$, whose components are denoted as $[R^*]$ in the fixed global reference frame, causes the same vector to take the new position $m_t$ given by

$$\{m_t\}^g = [R^*] \{m\}^g = [R^*] [g] \{m\}^c$$

(26.16)

Therefore the product $[R^*][g]$ then provides us the new coordinate transformation matrix between the new lattice orientation of the crystal and the fixed global reference frame.
26.2.4 Polycrystal Homogenization Theories

A number of polycrystal homogenization theories are used in practice to scale up to the response of polycrystals from the response of the constituent single crystals. Taylor (Taylor 1938) proposed a simple isostrain model for predicting the macroscopic yield strength of polycrystals, and this model with relatively minor improvements continues to be the most widely used theory even today. Major improvements to the basic Taylor model have been proposed in literature. These include the self-consistent models (Molinari et al. 1987; Toth et al. 1994) and the LAMEL model (Van Houtte et al. 1999; Van Houtte et al. 2002) that have been discussed in Chapter 22. A common feature of all these analytical approaches to the polycrystal homogenization theory is that they violate either equilibrium or compatibility or both in the given polycrystalline aggregate. As an alternative to these analytical approaches, micro-mechanical finite element models have been proposed (Bronkhorst et al. 1992; Kalidindi et al. 1992; Beaudoin et al. 1993; Beaudoin et al. 1995; Sarma and Dawson 1996) that satisfy both equilibrium and compatibility in a weak numerical sense in the given polycrystal. These are described in the next section.

26.3 Micro-Mechanical Finite Element Models

In a typical non-linear finite element stress analyses problem, the weak form of the principle of virtual work (including equilibrium and the boundary conditions) is satisfied in a given domain, discretized into a finite element mesh, by evaluating the constitutive response at specific locations in each element called integration points. In the present case, the given polycrystalline aggregate is discretized into a finite element mesh, such that each grain is modeled by one or more finite elements to allow for non-uniform deformations between the grains and within the grains. The single crystal constitutive behavior, together with the fully implicit time integration procedure described above, can be incorporated into a finite element code. In particular, it has been successfully incorporated into the commercial finite element code ABAQUS, through use of a user-defined subroutine for material behavior called UMAT (ABAQUS 2003). The code has been written such that it keeps track of the necessary solution dependent internal state variables at each of the integration points in the mesh, and updates them during an imposed deformation increment using the approach described earlier. In addition, the implicit finite element codes also require a computation of the Jacobian matrix at each integration point in the mesh in order to compute a better guess of the overall non-uniform deformation (and strain) field in its iterative attempts to satisfy the principle of virtual work during an imposed increment of deformation. This Jacobian matrix can be computed numerically (Kalidindi et al. 1992) or analytically for the crystal plasticity model described here.

The algorithms used for coding the user subroutine in ABAQUS are beyond the scope of this chapter. Interested readers can find the details in (Kalidindi et al. 1992). Instead, a number of different applications using these codes are described in the next section.
26.4 Examples

26.4.1 Predictions of Deformation Textures

One of the most elementary finite element models that was used in modeling a polycrystalline aggregate (Kalidindi et al. 1992; Bachu and Kalidindi 1998) is shown in Figure 26.2. In these models, the polycrystalline aggregate is given a rectangular parallelepiped geometry. Each element in the finite element mesh (shaded region in Figure 26.2) is cube-shaped, and represents an individual crystal or grain in the aggregate. The different finite elements in the mesh are given different initial crystal orientations such that the average initial texture in the model is representative of the initial texture in the sample. Figure 26.4.1 shows the measured and predicted textures for a representative high stacking fault energy cubic polycrystalline metal (the measurement is for copper) from both a Taylor-type model as well as a simple micro-mechanical finite element model of the type shown in Figure 26.2. Although the grossly simplified Taylor-type model captures quite well all of the major features of the measured texture in the deformed sample, there are definitely certain subtle features that are better captured by the micro-mechanical finite element model. Quantitative comparisons based on the position and intensities of the skeletal lines in the orientation distribution functions corresponding to the textures shown in Figure 26.4.1, along with the corresponding results from other polycrystal homogenization theories have been presented in literature (Delannay et al. 2002; Van Houtte et al. 2002; see also Chapter 22). These studies concluded that the micro-mechanical finite element models provide significantly improved predictions of the deformation textures, compared to the highly simplified Taylor-type model.

The coarse finite element models of the type shown in Figure 26.2 have also been used successfully to predict shapes of anisotropic yield surfaces for cubic polycrystals (Kalidindi and Schoenfeld 2000).

26.4.2 Predictions of Micro-Texture

In spite of the remarkable success of the modern crystal plasticity theory in predicting reasonably accurately the anisotropic stress-strain response and the averaged texture evolution during large plastic strains in polycrystalline metals, there does not yet exist a comprehensive theory that is currently capable of predicting accurately the crystal-scale evolution of the microstructure, sometimes referred to as micro-texture. Micro-texture is extremely important in developing recrystallization models. It is also intuitive that understanding the accommodation
of plastic deformation at the local scale has important bearing on fracture and other failure related properties of a number of brittle solids. The original experiments of (Barrett and Levinson 1940), and the more recent repetitions of these experiments by (Panchanadeeswaran et al. 1996) and by (Bhattacharyya et al. 2001), have cast serious doubts on the veracity of the currently used crystal plasticity theory. Below, we summarize the results from a direct comparison between measurements in an Aluminum polycrystal and the corresponding predictions from both the Taylor-type model and the micro-mechanical finite element model.

A sample was cut from a high purity (99.9%) Aluminum slab having a columnar grain structure. Figure 26.4 shows the initial sample dimensions, the orientation of the columnar crystals with respect to the sample geometry, and the surfaces that were scanned by the OIM (Orientation Image Microscopy) technique. In this study, the plastic deformation was imposed using a channel-die set-up, shown schematically in Figure 26.4. Although an overall reduction of about 40% was obtained, the deformation was imposed in about four stages. At each stage of deformation, the sample microstructure was characterized using the OIM technique.

The evolution of lattice orientations in the polycrystalline sample was modeled using both a grossly simplified Taylor-type model and the micro-mechanical finite element model. The finite element simulation was carried out assuming that the grains were perfectly columnar. ABAQUS/CAE was used for generating grain boundaries in the sample. Figure 26.5 shows the finite element mesh used in this study. It comprised of 615 C3D8 (three-dimensional 8-
Figure 26.4: A schematic description of the channel-die set-up used for imposing plastic deformation on the quasi-columnar grained sample of polycrystalline aluminum. Also shown are the sample initial dimensions and the final dimensions and the orientation of the columnar axis of the constituent crystals with respect to the sample geometry.

noded solid) elements (ABAQUS 2003). The three dimensional mesh was compressed along Side 2 (see Figure 26.5) by 40% (keeping Side 1 fixed). Sides 3 and 4 were left unconstrained. Figure 26.6(a) shows the deformed mesh. It was found that the shape of the deformed mesh from the finite element model matched well with the experimentally observed shape of the deformed sample. In particular, it was observed both in the experiment and in the finite element simulation that the deformed sample had undergone a shear of 0.16 along positive RD on a plane normal to TD.

In spite of the grossly simplifying assumptions, the Taylor-type model performed remarkably well in predicting the averaged texture for the polycrystal as shown in Figure 26.7. The finite element model predictions for the averaged final texture are better than the Taylor-type model predictions, in that they are more diffuse as in the experiment. It should be noted that the Taylor model, by its inherent assumption, could not predict the development of misorientations in the individual crystals. More importantly, the Taylor-type model also failed completely in predicting the rotations of several individual crystals, even in the average sense (see Figure 26.8). On the other hand, the finite element model provided significantly better predictions of both the rotations of individual grains as well as the development of average misorientations (Figure 26.8). However, the finite element model failed to predict the formation of deformation bands in any of the grains, as can be seen from Figure 26.6(b), which shows the contours of the magnitude of the neo-Euler angle between the initial measured and the FE predicted orientations. The regions with similar orientation will show similar shading in the plot. Experimentally, grains 5 and 10 showed development of banded orientation fields, but the corresponding grains in the FE model showed no signs of similar deformation banding within the grain.

However, in this experiment, it was possible to dig a little deeper by successfully estimating the deviations in local deformation history compared to the macroscopically imposed deformation. This was possible in our experiment because of the channel-die configuration, which allowed only a smaller number of additional shears both at the sample scale as well as the crystal scale. Using the crystal plasticity models mentioned earlier and examining all
possible trial local strain rate tensors allowed by the sample geometry and imposed boundary conditions, we were able to identify for each crystal of interest the local strain rate tensor that resulted in the lattice rotations that matched with the OIM measurements during all stages of the deformation history. In this analysis, we also carefully tracked the changes in the local energy dissipation rate (due to plasticity).

This combined experimental-modeling approach resulted in establishing new insights into the underlying physics behind the deformation of polycrystals. Certain grains in the deformed sample exhibited more or less uniform lattice orientation (with minimal misorientations) and were labeled as homogeneously oriented crystals. Our analyses indicated that all these grains were relatively hard grains compared to their neighbors. In all the homogeneously oriented crystals, there were additional shears locally in the crystal (in addition to the macroscopically imposed plane strain deformation history) that resulted in a decrease of the plastic energy dissipation rate. Consequently, the Taylor factors of all these grains in the final deformed state decreased significantly compared to those corresponding to their initial orientations. Therefore, the local deformation history of the homogeneously oriented crystals is governed largely by the need to minimize its plastic energy dissipation rate.

Certain other grains in the deformed sample were found to fragment into regions with repeating alternating orientation fields (Bhattacharyya et al. 2004). These grains were found to exhibit inherent instability in crystal lattice rotations. That is, it was found for these grains that
with a slight perturbation of the orientation before fragmentation (from different regions of the crystal), different slip system activities and consequently different crystal rotations were preferred in different regions of the grain. Finally, a certain number of crystals in the deformed sample showed grain fragmentation, but the lattice orientations in the deformed grains did not exhibit repeating orientation fields, i.e. the deformation bands in these crystals showed non-repeating orientation fields. These grains were found to be relatively soft crystals (lower Taylor factors than neighbors) with larger grain-sized and harder neighbors. There is a strong positive correlation between the additional shears experienced in these bands and the additional shears experienced by their adjacent harder neighbors. Furthermore, in these crystals, it was often noted that the Taylor factor corresponding to the final orientation was higher than

Figure 26.6: (a) The deformed FE mesh on side A after 40% compression and (b) the corresponding contours of neo-Euler angles between the initial orientation and the final orientation of the grains.
Figure 26.7: (100) pole figures of the columnar-grained sample after 40% deformation in channel-die compression. (a) Experimental measurement, (b) Taylor-type model prediction, and (c) Finite element model prediction.

that corresponding to the initial orientation of these grains. We therefore conclude that different regions of these crystals are responding to different influences from different neighbors. Further details of this study can be found in (Bhattacharyya et al. 2004).

In summary, it is noted that using a combination of sophisticated experimental and modeling techniques, it is now possible to follow the local evolution of the microstructure at the individual crystal scale in polycrystalline metals and extract the underlying physics behind how individual grains respond to a specific macroscopically imposed deformation history. The reader would also benefit from the recent work of (Raabe et al. 2001) where a novel method to experimentally measure the local deformation history is described. These developments open tremendous new possibilities for undertaking similar studies in a large number of other material systems with more complicated micromechanisms of deformation and failure. It should also be recognized that it would also be important in future to further expand the methods described here to extend the investigations to real microstructures with more realistic grain sizes and morphologies.
Figure 26.8: Comparison of evolution of lattice rotations in individual grains depicted by (100) pole figures. Shown are the initial orientations as well as the final measurements after 40% reduction in plane strain compression along with the predictions from both the Taylor-type model and the finite element model prediction.

Acknowledgements

The author is grateful for his collaborations and many discussions with his colleague, Professor Roger D. Doherty, and his ex-PhD student, Dr. Abhishek Bhattacharya, and for their
many contributions to this work. This work was supported by NSF-DMR9612343 and NSF-DMR0303395.

References

27 A Crystal Plasticity Framework for Deformation Twinning

Surya R. Kalidindi

27.1 Introduction

Crystallographic slip and deformation twinning are the dominant modes of plastic deformation at low homologous temperatures in a number of commercially important metals (Christian and Mahajan 1995; Asgari et al. 1997; Salem et al. 2002; Kalidindi et al. 2003; Salem et al. 2003). As described in other chapters of this series, significant progress has already been made in formulating a rigorous constitutive framework for finite strain elastoplastic models for crystalline materials deforming by slip alone, and implementing such models in finite element codes to predict the anisotropic stress-strain response, the macroscopic shape changes, and the averaged crystallographic texture evolution during non-homogenous bulk deformation processing operations. The modeling and simulation tools developed have been demonstrated to be fairly accurate for single phase, medium to high stacking fault energy (SFE), cubic polycrystalline metals, in which plastic deformation occurred predominantly by slip (Mathur et al. 1989; Bronkhorst et al. 1992; Kalidindi et al. 1992; Beaudoin et al. 1993; Kalidindi and Anand 1994; Surma and Dawson 1996; Delannay et al. 2002; Van Houtte et al. 2002); see also Chapters 5 and 26). This chapter summarizes some of the recent advances made in extending this modeling framework to include deformation twinning as a mode of plastic deformation in addition to crystallographic slip.

27.1.1 Slip vs. Deformation Twinning

Figure 27.1 shows a schematic of the rearrangement of atomic positions associated with the creation of a deformation twin. At least two major characteristics of deformation twins are apparent from this figure: (i) Unlike crystallographic slip, deformation twinning produces a distinct rotation of crystal lattice relative to the untwinned region (referred to as the matrix in this chapter). (ii) The twin-matrix boundaries (twin habit planes) are planes of mirror symmetry for the lattice orientations inside the twin and in the matrix.

Deformation twinning competes with crystallographic slip as a fundamental mode of plastic deformation in many cubic metals. In other words, in materials where deformation twinning is relatively easier compared to slip, deformation twinning is expected to occur extensively. It is widely believed that the stress required for formation of deformation twins is controlled mainly by the value of the stacking fault energy (SFE) of the metal. In particular, from an analysis of data from single crystals of low SFE metals, Venables (Venables 1964)
concluded that there exists a parabolic relationship between the SFE and the twin nucleation stress. It should be noted, however, that there is considerable scatter in the data presented by Venables. Furthermore, the different alloys considered by Venables had different solute concentration levels, and therefore the solid solution strengthening contributions for these different metals can be expected to be significantly different from each other, which has not been factored into the analyses. More recently, the above idea has been refined in carefully planned experimental investigations on a number of cubic polycrystalline metals. It is now clear that the microstructural variables that control directly the deformation twinning stress in polycrystalline FCC metals are the dislocation density and the average homogeneous slip-length. This dependence can be expressed empirically as a power-law relation (El-Danaf et al. 1999). Deformation twinning in cubic polycrystals is, however, promoted indirectly by lower values of stacking fault energy because of its influence on enhancing strain hardening rates and in reducing the grain breakup at least up to the strains where deformation twins appears.

In hexagonal close-packed metals, depending on the c/a ratio, either prism slip or basal slip along the close-packed basal directions (a) generally dominates the plastic response of the metal. Slip on pyramidal planes along the a-directions is significantly harder, and even more difficult on the pyramidal planes along the (c + a) directions. In these metals, in order to attain five independent shear modes (to accommodate arbitrary local plastic strains at the crystal scale), it is therefore essential to invoke either pyramidal slip along the (c + a) directions or one of the various possible deformation twinning modes. This competition between possible slip and twin modes has been brought out clearly in carefully planned experimental investigations (Salem et al. 2003) on textured samples of high purity α-Ti.

### 27.1.2 Major Consequences of Deformation Twinning

The most dominant effect, and perhaps the most important one from an engineering viewpoint, is the influence deformation twinning has on the strain hardening response of materials. It is now well established that deformation twinning contributes significantly to the high strain
27.1 Introduction

Figure 27.2: True stress-True strain responses during simple compression of high purity $\alpha$-Titanium, 70/30 Brass, and Copper.

hardening rates (and consequently to high tensile ductility) of several commercially important metals and their alloys. Figure 27.2 shows the measured true stress-true strain responses of certain fcc and hcp polycrystalline metals deformed to large plastic strains in simple compression (Kalidindi et al. 2003). In these experiments, the brass and titanium samples exhibited deformation twinning, while the copper samples did not show any signs of deformation twinning in detailed microscopy. The contrast in the mechanical responses of these representative metals becomes even more discernable in the strain hardening plots of the type shown in Figure 27.3. In both 70/30 brass and in $\alpha$-Ti, the initiation of deformation twinning was found to correlate with a distinct change in the strain hardening response of the material, thereby validating the hypothesis that deformation twinning is primarily responsible for the drastically enhanced strain hardening rates exhibited by these materials. The precise physical mechanisms contributing to the strengthening due to formation of deformation twins have been elucidated in a recent experimental investigation (Salem et al. 2002).

Deformation twinning is also known to play an important role in several aspects of microstructure evolution. Most important of these effects is perhaps the effect on the evolution of deformation textures. In fcc polycrystals, it was initially hypothesized (e.g. (Hirsch and Lucke 1988)) that the observed transition from a copper-type texture in medium and high SFE metals to a brass-type texture in low SFE metals in rolled samples was largely a direct effect of the lattice reorientation of the twinned regions in the later. In recent years, this hypothesis has been abandoned and it is now widely believed that this transition in rolled textures is largely
an indirect consequence of deformation twinning (Leffers et al. 1988; El-Danaf et al. 2000). In contrast, the effect of deformation twinning on texture evolution in several hcp polycrystals is quite unambiguous. In hcp metals, where the extent of deformation twins in the microstructure is substantial (as high as 40% by volume in room temperature deformation of high purity $\alpha$-Ti), there is a significant contribution to evolution of deformation textures from the lattice orientation of the twinned regions (Kalidindi et al. 2003).

### 27.2 Historical Perspective

Incorporation of deformation twinning in the crystal plasticity modeling framework as an additional mode of plastic deformation has been an outstanding problem. The similarity between the shearing modes involved in slip and twinning have long been recognized and exploited in these efforts (Chin et al. 1969). The potential twin systems for different classes of crystal structures have also been identified and summarized in literature (Barrett and Massalski 1966). As discussed by Van Houtte (Van Houtte 1978), the major hurdle in extending the crystal plasticity framework to include deformation twinning is the lack of an efficient procedure to handle the large number of orientations created by the twinned regions which are different from the untwinned regions of the grain. In particular, consider the following example by Van Houtte (Van Houtte 1978). Suppose that in each time step each grain produces one set of twins (i.e. one twin system is activated). If there are $N$ starting grain orientations, the number of grain
Historical Perspective

orientations after \( n \) time increments will be \( N2^n \). For example, with 150 initial grain orientations and forty time steps (assuming a strain of 0.05 in each step to reach a total strain of 2.0), the number of final grain orientations will be \( 1.65 \times 10^{14} \). As a result, it is highly impractical to carry out such simulations. This problem gets greatly compounded when such models are to be incorporated into finite element codes to facilitate simulations of non-homogeneous deformations and these orientations need to be stored and handled at each integration point in the finite element mesh.

The two major strategies proposed previously in literature to address the problem described above are described next.

27.2.1 Twin Reorientation Schemes

Van Houtte (Van Houtte 1978) proposed an efficient scheme where the evolution of the volume fractions of the twinned regions in each grain are tracked carefully, but the twinned regions in the grain are not re-oriented at the end of each time step. Instead, a statistical criterion based on the volume fractions of the twinned regions in the grain and in the entire polycrystalline aggregate is used to decide if the entire grain should be re-oriented into a dominant twin orientation. In this technique, since the entire grain is either re-oriented or not, the number of total grain orientations in the calculations is kept fixed. Van Houtte used this methodology and demonstrated that it is possible to predict the “brass-type” textures observed in rolling and simple shear of low SFE fcc metals. The disadvantages of this scheme are: (i) a large number of grain orientations are required in order for the statistical criterion used in this procedure to be meaningful, and (ii) the decision to re-orient the grain was based on the increments of the twin volume fractions in a given time-step independent of the previous deformation history, and therefore the selected orientation to which the grain may be twinned may not be the most dominant one. Although the later deficiency was addressed by Tome and co-workers (Tome et al. 1991) with their predominant twin reorientation scheme, this technique still has the disadvantage that the grain is re-oriented to only its dominant twin orientation; the orientations of the other twinned regions in the grain are lost. The twin reorientation technique was also employed successfully for simulating stress-strain curves and texture evolution in low stacking fault energy fcc polycrystals (Staroselsky and Anand 1998).

27.2.2 Volume Fraction Transfer Scheme

Tome and co-workers (Tome et al. 1991) suggested another novel scheme, which involved the use of weighted grain orientations. In this scheme, the twinned parts of the grains are re-oriented at the end of each time step, but since weighted grain orientations are used there is no need to create new orientations; only the weights of the grain orientations need to be suitably modified to reflect the orientation changes caused by twinning. The usefulness of this scheme was demonstrated by simulating a variety of textures in Zr alloys (primarily in rolling) while using different values for the ratios of the slip system resistances to the twin system resistances in these materials. The disadvantage of this technique is that it too requires a large number of crystal orientations, because the starting grain orientations have to be selected such that they cover the entire three-dimensional euler angle space.
A fundamental drawback of both the schemes described above is that the twinned region, after suitable lattice reorientation, loses memory of all prior slip activity. That is, the twinned region is essentially treated as a new grain, thereby not allowing one to keep track of the strain hardening information pertinent to a specific grain. In this chapter we present a new crystal plasticity modeling framework that overcomes several of the challenges described above.

### 27.3 Incorporation of Deformation Twinning

The necessary theoretical background, for the situation where the plastic deformation in the crystal is accomplished by slip alone, has already been introduced in Chapter 26 of this book. We now address the extension of the crystal plasticity modeling framework to include deformation twinning as an additional mode of plastic deformation. For the most part, the main aspects of the modeling framework remain the same as described earlier for slip alone in Chapter 26 of this book. We focus here on only those aspects that need modification.

#### 27.3.1 Relaxed Configuration

![Figure 27.4: Interpretation of multiplicative decomposition of the deformation gradient tensor, when deformation twinning is introduced as an additional mode of plastic deformation.](image-url)

Figure 27.4: Interpretation of multiplicative decomposition of the deformation gradient tensor, when deformation twinning is introduced as an additional mode of plastic deformation.
27.3 Incorporation of Deformation Twinning

Figure 27.4 presents a new interpretation of the multiplicative decomposition of the deformation gradient into elastic and plastic components when twinning is included as an additional mode of plastic deformation. For clarity, only one twin system is shown in Figure 27.4; it is, however, possible for multiple twin systems to be activated in the same grain. Furthermore, in Figure 27.4, the twinned region belonging to one twin system is idealized as a continuous block, while in actuality, deformation twins occur as parallel elliptical regions (e.g. Asgari et al. 1997; Salem et al. 2002). It should be noted that further slip is permitted in each of the twinned regions.

The quantities $F$, $F^*$ and $F^p$ in Figure 27.4 represent homogenized values at the grain scale. As such they do not represent the deformation gradients in either the matrix or the twinned regions alone. It is assumed that twinning causes a rotation of the lattice of part(s) of the grain relative to the untwinned regions. It should be noted that the relaxed configuration of Figure 27.4 is an isoclinic configuration for the matrix region, i.e. the lattice orientation of the matrix region in the relaxed configuration is exactly the same as that of the initial crystal.

The relationship between the lattice orientation(s) in the twinned region(s) and the lattice orientation of the matrix is treated here as an additional solution dependent state variable, denoted by $R_{tw/\beta}$. The superscript $\beta$ in $R_{tw/\beta}$ reflects the fact that this rotation is expected to be different for each of the different potential twin systems in the crystal. Also note that we are homogenizing all of the twins of a specific twin system into a single equivalent twin, independent of when they were produced, and this will be given particular attention later in the development of model. As described earlier, it is indeed a property of the twins that when they are formed, $R_{tw/\beta}$ has a simpler description in that it is a 180 degree rotation about the twin habit plane normal. This simple relationship between the matrix and twin lattice orientations is however lost as soon as there is additional slip activity in either the matrix or the newly formed twin (Asgari et al. 1997; Salem et al. 2003). Consequently, it is important to introduce $R_{tw/\beta}$ as an additional solution dependent state variable (quite similar to the way $F^p$ was introduced).

As before, the lattice orientations of the matrix and the twinned regions in the crystal can be computed using $R^*$, the rotation component of $F^*$. In fact, the computation for the matrix part is exactly the same as before (Equation (26.16) in Chapter 26). For the twinned regions, this equation would be modified as

$$\{m_t\}^K = [R^*]_{} [R_{tw/\beta}^{}] [\{m\}]^g = [R^*]_{} [R_{tw/\beta}^{}] [g]_{} [\{m\}]^c \tag{27.1}$$

As described earlier, the special relationship of a 180 degree rotation about the twin habit plane normal for the matrix and the twinned orientations will only be satisfied just when the twin is produced. With further slip in either the matrix or the twin (in reality both will happen), this relationship is lost quickly, and the twin parts and matrix parts rotate differently, but towards the same ideal textures for a given deformation path (Salem et al. 2003). This phenomenon is implicit in Equation (27.1) and the recognition that $R_{tw/\beta}$ is a solution dependent state variable that evolves continuously with imposed strain.

A significant assumption, and perhaps the most significant limitation of the model at the present time, is that it is implied in the idealization of Figure 27.4 that all of the twinned regions of the crystal belonging to a particular twin system would have a single lattice orientation in the current configuration. In actuality, since the twins are produced at different
instances along the deformation path, they will not have the exact same lattice orientation. However, micrographs of twins presented in published literature show that deformation twins belonging to a particular twin system in the crystal occur as clusters of plates parallel to each other in a short range of strain levels (e.g. (Asgari et al. 1997; Salem et al. 2002)). This suggests that the lattice orientations in these twins are somewhat close to each other. More detailed experimental data is needed to assess critically this aspect of the present model.

27.3.2 Elastic Response

Recognizing that the quantities \( F, F^*, F^p, T^* \) and \( E^* \) are all effective (homogenized) quantities in the decomposition described in Section 27.3.1, it can be seen that the elastic response of the crystal (Equation (26.2a) in Chapter 26) has to be modified using an effective elastic stiffness tensor, \( \bar{C} \), as

\[
T^* = \bar{C} E^*
\]  

(27.2)

A simple homogenization model is proposed for the effective elastic stiffness of the grain, and is expressed as

\[
\bar{C} = \left(1 - \sum_\beta f^\beta\right) C^{mt} + \sum_\beta f^\beta C^{tw\beta}
\]  

(27.3)

where \( f^\beta \) denotes the volume fraction of a particular twin system (and is another solution dependent state variable that evolves with deformation), \( C^{mt} \) is the elastic stiffness tensor of the matrix region, and \( C^{tw\beta} \) is the elastic stiffness tensor of the twinned region(s). It should be noted that \( C^{mt} \) and \( C^{tw\beta} \) are related through a coordinate transformation law, expressed as

\[
C^{tw\beta}_{ijkl} = C^{mt}_{pqrs} R^{tw\beta}_{ip} R^{tw\beta}_{jq} R^{tw\beta}_{kr} R^{tw\beta}_{ls}
\]  

(27.4)

where the transformation matrix \( R^{tw\beta} \) was defined earlier in Section 27.3.1.

27.3.3 Plastic Flow Rule

The plastic flow rule (Equation (26.3b) in Chapter 26) is modified to take into account the additional mechanisms of plastic deformation caused by inclusion of deformation twinning.

\[
L^p = \left(1 - \sum_\beta f^\beta\right) \sum_\alpha \hat{\gamma}^\alpha S^\alpha_{v-s} + \sum_\beta f^\beta \hat{\gamma}^{tw\beta} S^{tw\beta}_{v-t} + \sum_\beta f^\beta \left(\sum_\alpha \hat{\gamma}^\alpha S^{tw\beta}_{v-t}ight)
\]  

(27.5)

In Equation (27.5), \( \hat{\gamma}^\alpha \) denotes the shearing rate on a slip system \( \alpha \), \( f^\beta \) denotes the volume fraction of the crystal experiencing the constant twinning shear \( \gamma^{tw\beta} \) associated with the twin system \( \beta \), and \( S \) defines the tensorial direction of the shear (unit magnitude) caused by an active slip or twin system. Note that there are three terms on the right-hand side of Equation (27.5) representing the three different contributions to plastic deformation. The first term
on the right-hand side represents the contribution to plastic deformation from slip in the matrix, the second term depicts the contribution from twinning, and the last term reflects the contribution from slip in the twinned regions. The convention for the different $S$ tensors is as follows. $S^\alpha_{o-sl}$ refers to the slip systems in the matrix, $S^\alpha_{o-tw}$ refers to the potential twin systems in the matrix, and $S^\beta_{tw-sl}$ refers to the slip systems in the twinned regions of the crystal. The first place in the subscript of $S$ refers to the orientation in which the slip or twin system is to be defined (‘o’ indicates the initial orientation of the grain, and ‘tw’ indicates the twinned orientation in the relaxed configuration). The second place in the subscripts of $S$ refers to whether the contribution is from slip or twinning. Note that $S$ is defined as $S = m \otimes n$, where $m$ represents a unit vector in the slip or the twin direction, and $n$ represents a unit vector along the slip or the twin plane normal.

In order to complete the description of the plastic flow rule (Equation (27.5)), specific constitutive functions need to be specified for the shearing rate on the slip system ($\dot{\gamma}^\alpha$) and for the rate of increase of twin volume content in the crystal ($\dot{f}^\beta$). $\dot{\gamma}^\alpha$ and $\dot{f}^\beta$ are expected to be dependent on the resolved shear stress on the slip system and twin systems respectively, and their corresponding slip or twin resistances. Following the previous modeling approaches, simple power law functions can be used for the description of the shearing rate on the slip system (Equation (27.4) in Chapter 26). However, allowing for the possibility of different slip rate sensitivities of different slip systems, it can be modified as

$$\dot{\gamma}^\alpha = \dot{\gamma}_0 \left| \frac{\tau^\alpha}{S^\alpha} \right|^{1/m^\alpha} \text{sgn} (\tau^\alpha)$$

This same function can be suitably modified for the rate of increase in the twinned volume fractions as

$$\dot{f}^\beta = \frac{\dot{\gamma}_0}{\gamma_{tw}} \left( \frac{\tau^\beta}{S^\beta_{tw}} \right)^{1/m^\beta} \text{ for } \tau^\beta > 0$$

$$\dot{f}^\beta = 0 \text{ for } \tau^\beta \leq 0$$

where $\dot{\gamma}_0$ is a reference value of slip rate taken as $0.001 \text{ s}^{-1}$, and $m^\alpha$ and $m^\beta$ are rate-sensitivity parameters for slip and twinning systems, respectively. The rate-sensitivity parameters can be different for different families of slip and twin systems.

Appropriate expressions for the various resolved stresses appearing in Equations (27.5)–(27.7) can be derived by equating plastic power at the macroscale to the plastic power from all slip and twin activity in the crystal. At the macroscale, the power per unit volume in the relaxed configuration can be expressed as

$$\dot{\omega} = (\det F) \mathbf{T} \mathbf{L}$$

Starting with this expression, the plastic power per unit volume can be derived to be

$$\dot{\omega}^p = C^* \mathbf{T}^* \mathbf{L}^p, \quad \text{with} \quad C^* = F^{*T} F^*$$
The plastic power per unit volume from the slip and twin activity in the crystal can be written as

$$\dot{\omega}^p = \left( 1 - \sum_{\beta} f_\beta^3 \right) \sum_{\alpha} \dot{\gamma}^{\alpha \gamma} t^{\alpha} + \sum_{\beta} f_\beta^3 \tau^{\beta \tau^{\beta}} + \sum_{\beta} f_\beta^3 \left( \sum_{\alpha} \dot{\gamma}^{\alpha \gamma} t^{\alpha} \right)$$  (27.10)

where the first term is from slip activity in the matrix, the second term is from twinning (in the matrix), and the third term arises from slip activity in the twinned regions. Setting the macroscale and microscale definitions of plastic power equal to each other (Equations (27.9) and (27.10)) and substituting Equation (27.5), we can derive the following expressions for the various resolved shear stresses:

- for slip in matrix:\n  \[ \tau^{\alpha} = C^* T^* \bullet S_\alpha^{\alpha - sl} \approx T^* \bullet S_\alpha^{\alpha - sl} \]  (27.11)

- for twinning in matrix:\n  \[ \tau^{\beta} = C^* T^* \bullet S_\beta^{\alpha - tw} \approx T^* \bullet S_\beta^{\alpha - tw} \]  (27.12)

- for slip in twins:\n  \[ \tau^{\alpha} = C^* T^* \bullet S_{\alpha - sl}^{\alpha - sl} \approx T^* \bullet S_{\alpha - sl}^{\alpha - sl} \]  (27.13)

Note that \( C^* \) defined by Equation (27.9b) has been approximated to an identity tensor in Equations (27.11)–(27.13), because the elastic stretches in metals are infinitesimal.

### 27.3.4 Evolution of Twin Rotations

![Figure 27.5: Evolution of the twin-matrix orientation relation with further slip activity. In this figure further slip activity is assumed only in the matrix.](image)

As described earlier, we note that the lattice relationship between the matrix and the twinned regions is evolving continuously with strain during a given deformation process. To formulate an appropriate evolution law for this rotational relationship, consider the following hypothetical situation. Figure 27.5 depicts the evolution of the matrix region with the adjoining twinned region in the relaxed configuration in a given time increment at some intermediate stage of the imposed deformation. Let us further assume that the matrix experiences slip, but that the twinned region does not experience any slip in this time increment. Although the slip in the matrix does not produce a lattice rotation (recall that the relaxed configuration is isoclinic for the matrix region), the twin will undergo a rotation as depicted in Figure 27.5. The relative lattice rotation of the twinned region with the matrix will depend on the velocity gradient tensor in the matrix (especially its spin component) and the geometrical relationship between the twin and the matrix regions (i.e. the orientation of the twin-matrix interface).
27.3 Incorporation of Deformation Twining

Here, we employ a simplifying assumption and take relative rotation of the twin lattice with respect to the matrix lattice to be given only by the spin component of the plastic velocity gradient tensor in the matrix. This concept can easily be extended to include the situation where there is slip activity in both the matrix region as well as the twinned regions. Using these ideas, the evolution equation for the relative lattice rotation of the twinned region with respect to the matrix region can be expressed as

\[
W^{\beta} = R^{\alpha \gamma \beta} R^{\gamma \alpha \beta T} = \sum_{\alpha} \frac{1}{2} \gamma^\alpha \left( S^\alpha_{\alpha-s} S^\alpha_{\alpha-s}^T \right) = \sum_{\alpha} \frac{1}{2} \gamma^\alpha \left( S^\alpha_{\tw-s} S^\alpha_{\tw-s}^T \right) \tag{27.14}
\]

It should be recognized that Equation (27.14) could be modified either by rigorously taking into account kinematics of deformation in the matrix and monitoring the orientation of the matrix-twin interface as an additional state variable and/or by extending the constitutive framework presented here to account specifically for dislocation densities on various slip systems; the recent advances in disclination theories might prove useful in this context (e.g. (Seefeldt et al. 2001)).

Figure 27.6: Schematic of the evolution of the twin activity for a given twin system in a time increment at an intermediate stage of the deformation process.

Equation (27.14) does not address the creation of new twins in a given time increment. This has to be treated separately. Figure 27.6 shows a schematic of the suggested process. Here \( R^{\tw(1)} \) represents the lattice rotation tensor of an existing twin (computed from Equation (27.14)), while \( R^{\tw(2)} \) represents the lattice rotation tensor for a new twin of the same twin system formed during the time increment under consideration (essentially defined by the 180 rotation about the matrix-twin plane normal). As all the twins of a given twin system are represented by a single twin in the present framework, there is a need to homogenize the two twins into a single equivalent twin. For this homogenization process, it is convenient to think of a differential rotation \( \Delta R \) in the new twin that represents the difference between the two rotation tensors \( R^{\tw(2)} \) and \( R^{\tw(1)} \), represent it by an axis-angle pair \((n, \theta)\), and to distribute this over the entire homogenized twin by appropriately reducing the rotation angle.
Mathematically, this can be expressed as:

\[
\text{Effective } R^{\text{tw}} = \Delta R^{*} R^{\text{tw}(1)}
\]

with

\[
\Delta R = R^{\text{tw}(2)} \left( R^{\text{tw}(1)} \right)^T = (n, \theta)
\]

\[
\Delta R^{*} = \begin{pmatrix} n, \frac{f^{(2)} \theta}{f^{(1)} + f^{(2)}} \end{pmatrix}
\]

(27.15) \hspace{1cm} (27.16) \hspace{1cm} (27.17)

where \( f^{(1)} \) and \( f^{(2)} \) denote the volume fractions of the two twins shown in Figure 27.6, respectively. Note that this form of homogenization preserves the orthogonal properties of the effective rotation tensor \( R^{\text{tw}} \).

### 27.3.5 Slip-Twin Hardening Functions

As mentioned briefly in Section 27.1.2, extensive data has been reported in recent literature on the consequence of deformation twinning on the strain hardening of both fcc and hcp metals (Asgari et al. 1997; Salem et al. 2002; Kalidindi et al. 2003; Salem et al. 2003). It is clear that the slip-twin interactions are quite complex, and several fundamental questions regarding quantification of these effects remain yet unanswered.

A detailed analysis of the experimental observations in low stacking fault energy fcc polycrystals produced the following insights:

1. Production of deformation twins in fcc crystals leads to strong hardening of the non-coplanar slip systems in that crystal, but has relatively small effect on the co-planar slip systems in that crystal.

2. The creation of deformation twins creates new “grain-boundary” type interfaces, effectively reducing the mean slip length in the crystal. The newly formed twin-matrix interfaces serve as effective sites for dislocation pile-ups and storage. Also, a reduction of the mean slip length necessitates a higher dislocation density to accommodate a given strain increment. This is believed to be the dominant mechanism leading to the higher strain hardening exhibited by the low SFE fcc metals (Asgari et al. 1997).

3. In most of the grains, deformation twinning initially produces only one twin variant, i.e. a set of parallel twin lamellae, usually referred to as primary twinning. The production of a set of parallel twin bands within the crystal significantly increases the stress required to produce another non-coplanar twin variant in the same crystal. At higher strain (or stress) levels, however, two sets of non-coplanar twin variants were found in many grains, and these appeared as twin-intersections in the micrographs (Asgari et al. 1997).

Consistent with these physical ideas, constitutive functions for slip hardening can be formulated as (Kalidindi 2001)

\[
\dot{s}^{\alpha} = h_s \left( 1 - \frac{s^{\alpha}}{s_s} \right) \sum_k \dot{z}^k,
\]

(27.18)
27.3 Incorporation of Deformation Twinning

\[ s'_h = s_{so} + s_{pr} \left( \sum_{k \in \text{Non-coplanar}} f^k \right)^{0.5} \tag{27.19} \]

Equations (27.18) and (27.19) represent a simple saturation-type hardening law for slip, with the saturation value for each slip system being governed by the amount of non-coplanar twin plates that exist in the crystal. This form was motivated by the physical insights obtained from experiments and summarized above. In particular, Equations (27.18) and (27.19) reflect the hypothesis that the twin-matrix interfaces promote dislocation storage and therefore increase the saturation value of slip resistance. Noting that the twin volume fraction is inversely proportional to the mean spacing of twins \( \sum f \propto l_{tw}^{-1}, \) where \( l_{tw} \) is the average spacing of the twins, and that the well known Hall–Petch relation suggests a dependence of flow stress on characteristic length-scale as \( \sigma \propto l^{-0.5}, \) the exponent is Equation (27.19) was therefore presumed to be 0.5. In Equations (27.18) and (27.19), \( h_s, s_{so}, \) and \( s_{pr} \) are material slip hardening parameters. Of these \( h_s \) and \( s_{so} \) represents the initial strain hardening rate on the slip system and the saturation value in absence of deformation twinning, respectively. These can be established using the part of the stress-strain curve before deformation twinning initiates in the microstructure (labeled as Stage A by (Asgari et al. 1997)). \( s_{pr} \) controls the contribution of deformation twinning to the strain hardening response.

The twin hardening functions are formulated, consistent with the physical ideas described above, as

\[ \dot{s} = h_{nc} \left( \sum f \right)^b \sum_{k \in \text{Non-coplanar}} \gamma_{twk} j^k + h_{cp} \left( \sum f \right) \sum_{k \in \text{Co-planar}} \gamma_{twk} j^k \tag{27.20} \]

In the above equation, \( \gamma_{twk} j^k \) represents the equivalent shear rate on twin system \( (k) \) due to deformation twinning, and \( \sum f \) denotes the total twin volume fraction in the crystal. The first term on the right-hand side of Equation (27.20) represents the twin hardening due to creation of the non-coplanar twins, while the second term represents twin hardening due to production of co-planar twins. It is necessary to use different forms for the two components to capture the experimental observations in low stacking energy fcc metals described earlier in Section 27.1.2. The parameters in Equation (27.20) are to be chosen such that the first term dominates in the early stages of deformation twinning and the second term catches up at later stages of twinning. For example, if the parameter \( b \) in Equation (27.20) is chosen to be quite close to zero, then with the onset of deformation twinning in a crystal on one of the twin systems, the non-coplanar twin systems in that crystal would harden much more than the co-planar twin systems. This would dissuade the non-coplanar twin systems from becoming potentially active at this stage of deformation. Asgari and co-workers (Asgari et al. 1997) report that beyond a certain stress level, a second set of twin variants appear in most grains. This feature would also be captured quite well by choosing the value of the parameter \( b \) to be quite close to zero, because this would cause the hardening rate from the first term to increase relatively slowly compared to the second term at later stages of deformation.

In recent work on hcp polycrystals, it was discovered that deformation twinning also produced strain hardening through a mechanism known as Basinski mechanism, in addition to the above described Hall–Petch mechanism. Basinski and co-workers (Basinski et al. 1997)
proposed that glissile dislocations before twinning are converted to sessile configurations as a result of the twinning shear transformation. Consequently, the twinned regions should be harder than the untwinned regions and will thus produce an overall hardening of the material. Because the volume fraction of twins in fcc polycrystals is significantly low, it is believed that this mechanism is not likely to impact significantly the strain hardening in fcc polycrystals. However, in hcp polycrystals, the twin volume fractions are significant (approximately about 40% by volume fraction), and therefore it is expected that this mechanism will play an important role in these metals. Consequently, it is proposed to modify Equation (27.18) as

\[ \dot{\sigma}_\alpha = h_s \left( 1 + B \sum_\beta f^\beta \right) \left( 1 - \frac{\dot{\sigma}_\alpha}{\dot{\sigma}_{\text{sat}}} \right) \sum_k \gamma^k \]  

(27.21)

where \( \sum_\beta f^\beta \) represents the total twin volume fraction, and the saturation value of slip resistance is still given by Equation (27.19). Note also that in hcp polycrystals, all twin systems are non-coplanar with the slip systems and therefore the second term in equation (27.20) drops out automatically.

27.4 Examples

In contrast to the large number of modeling and simulation examples available in current literature for polycrystalline metals that deform by slip alone, there exist very few modeling and simulation results for polycrystals that deform by slip and twinning. This is attributed to the large number of complexities associated with incorporation of deformation twinning in the framework of crystal plasticity models. The framework described in Section 27.3 has only been developed recently, and simulation examples using this framework are just being worked out. In this section, we present briefly some of the examples available in current literature using certain simplified versions of the overall framework presented earlier.

The slip-twin hardening functions presented in Equations (27.19)–(27.21) were used to model stress-strain curves and texture evolution in simple compression, plane strain compression, and simple shear on polycrystalline samples of 70/30 brass. A set of four hundred random grain orientations were used to represent the initial texture in the samples. The strategy adopted was to estimate the values of all the slip and twin hardening parameters (in Equations (27.20)–(27.21)) by curve-fitting the predicted stress-strain response in simple compression to the measured stress-strain response. Once these values were established, the same values were used in subsequent simulations of the other deformation modes. The details of the curve-fitting procedure can be found in (Kalidindi 2001).

Figure 27.7 shows the predicted equivalent stress-equivalent strain curves in simple compression, in simple shear and in the path change from simple shear to simple compression, along with the measurement for simple compression deformation. Note that the simple model presented here captures well two important features of the stress-strain response of low SFE fcc metals: (1) The equivalent stress-equivalent strain response in simple shear is significantly lower than the response in simple compression. (2) Upon a path change from simple shear to simple compression there is an instantaneous jump in the equivalent flow stress, with the flow stress after the path change in reasonable correspondence with the flow stress that would have
27.4 Examples

Figure 27.7: Predicted stress-strain response of 70/30 brass in simple compression, simple shear, and the deformation path change from simple shear to simple compression modes. The measured response in simple compression is also shown.

been obtained in monotonic simple compression deformation at the same equivalent strain level. It was noted earlier that these two features are quite characteristic of the reported measurements in low SFE fcc metals (El-Danaf et al. 2001).

The predicted textures from the simulations described above (with about 20% deformation twins) were found to be almost identical to those predicted by the classical crystal plasticity models with slip alone. As an example, the textures predicted after a true strain of $-1.0$ in simple compression and in plane strain compression with the slip-twinning model described here, are shown in Figures 27.8 and 27.9, respectively, as equal-area projection pole figures. Comparing these to the corresponding measurements (e.g. (El-Danaf et al. 2000)), it is seen that there is a good agreement with the simple compression measurements, while there are significant differences in the case of plane strain compression.

In plane strain compression, the crystal plasticity models with deformation twinning are observed to provide good predictions in early stages of the deformation (El-Danaf et al. 2000; Kalidindi 2001). It is now believed that the major difference between the brass-type texture and the copper-type texture arises due to microscale shear banding and not due to deformation twinning (El-Danaf et al. 2000). Deformation twinning, however, plays an important role in influencing the strain hardening response, and is also likely to play an important role in the initiation of the micro-scale shear bands.

Several research groups are currently pursuing application of crystal plasticity models to hcp polycrystals. Texture predictions, without detailed attention to stress-strain curves and the associated strain hardening regimes, have already been presented in literature (e.g. (Tome et al. 1991; Phillipe et al. 1995)). It is anticipated that in near future some of the recent physical insights derived from experiments will be applied together with above modeling ideas to develop crystal plasticity models for hcp polycrystals that will not only predict accurately the texture evolution but also the concomitant evolution of the anisotropic stress-strain curves.
Figure 27.8: Predicted texture in simple compression ($\varepsilon = -1.0$) from the slip-twinning model shown as equal-area projection pole figures.

Figure 27.9: Predicted texture in plane strain compression ($\varepsilon = -1.0$) from the slip-twinning model shown as equal-area projection pole figures.
Acknowledgements

The author is grateful for his collaborations and many discussions with his colleague, Professor Roger D. Doherty, and his students, Dr. Ayman Salem and Mr. Xianping Wu, and for their many contributions to this work. This work was supported by NSF-DMR0201382.

References


28 The Texture Component Crystal Plasticity Finite Element Method

Franz Roters

28.1 Introduction

In metal forming operations crystallographic texture can be regarded as the main source of elastic and plastic anisotropy. Therefore every simulation aiming at the reproduction of this behavior is concerned with the representation of texture and its development. In the framework of the Finite Element method (FEM), which is commonly used for such simulations, the material constitutive law, i.e. the stress-strain law, has to account for texture and its development.

In the first part of the chapter the recently introduced Texture Component Crystal Plasticity Finite Element Method (TCCPFEM, Zhao et al., 2001; Raabe and Roters, 2004) is reviewed. It is basically a combination of crystal plasticity FEM (CP-FEM, see Chapters 5 and 26) with the texture component method for the representation of statistical textures. The main focus of this part will be on the representation of the random texture fraction and the texture components in the framework of CP-FEM.

The second part of the chapter will focus on different application fields of the method. Starting with rather simple r-value (or Lankford value, that is the ratio of the strain components $\epsilon_{33}$ and $\epsilon_{22}$) prediction more enhanced applications such as deep drawing will be presented. As one example it will be shown how the method can be used for texture optimization in view of desired product properties. Finally an outlook will be given on the potential of the method for further application.

28.2 The Texture Component Method

The basic idea of the texture component method is the approximation of X-ray textures in the form of the orientation distribution function (ODF) by a discrete set of simple distribution functions (texture components) defined in orientation space. It was developed by Lücke et al. (1981, 1986) and later on facilitated and implemented into an easy to use computer program\(^1\) by Helming (1996) (also see Helming et al., 1994).

\(^1\) The software “Multex 1.0” can be downloaded from http://www.texture.de

Edited by Dirk Raabe, Franz Roters, Frederic Barlat, Long-Qing Chen
Copyright © 2004 Wiley-VCH Verlag GmbH & Co. KGaA
ISBN: 3-527-30760-5
28.2.1 Approximation of X-Ray Textures using Texture Components

The mathematical reproduction of the orientation distribution function \( f(g) \) by texture component functions can be expressed by the superposition

\[
f(g) = F + \sum_{c=1}^{C} I_c f_c(g) = \sum_{c=0}^{C} I_c f_c(g) \quad \text{where} \quad I_0 = F, f_0(g) = 1
\]  

where \( g \) is the crystallographic orientation and \( F \) is the volume portion of all randomly oriented crystals (random texture component). \( F \) may be understood as the intensity of the only global component used in the approximation, equivalent to \( f_c(g) = 1 \) for each orientation point in Euler space, \( g \in G \). The intensity \( I_c \) describes the volume fraction of all crystallites belonging to the component \( c \). \( f_c(g) \) describes the orientation density of the component.

Helming (1996) uses two types of model functions for \( f_c(g) \), namely, fiber components and spherical components. The fiber components are characterized by the fiber axis \( f_c \) and the scatter width \( a_c \), while the spherical components are characterized by a preferred orientation \( g_c \) and a scatter width \( b_c \). In both cases a modified Gauss distribution

\[
f_c(g) = \frac{\exp(S_c \cos \omega_c)}{I_0(S_c) - I_1(S_c)}
\]

is used for \( f_c(g) \), where \( I_0 \) and \( I_1 \) are generalized Bessel functions and \( S_c \) is defined as

\[
S_c = \ln 2 \frac{1 - \cos \left( \frac{a_c}{2} \right)}
\]

for the spherical components \( a_c \) has to be substituted by \( b_c \). \( \omega_c \) is the orientation distance \( \omega_c = f(g, f_c) \) or \( \omega_c = f(g, g_c) \), respectively. In case of the fiber components Equation (28.2) is only used for rotation axes that are perpendicular to \( f_c \). If the rotation axis is parallel to \( f_c \) an equal distribution is used. For the spherical components Equation (28.2) is always used, independent of the rotation axis.

28.2.2 Representation of Texture Components in a Crystal Plasticity FEM

A big advantage of the texture component method is the fact that a very comprehensive data set is obtained for the description of statistical textures. In most cases only very few (3 or less) components are needed to obtain a texture approximation, which sufficiently reproduces the materials mechanical anisotropy. Obviously it is not the aim here, to get a perfect simulation of the deformation texture. Nevertheless, the component information has to be passed into a crystal plasticity FE calculation. The following paragraphs describe the procedure how these texture components are mapped to a finite element mesh. It is important at this point, that this procedure should be independent of the number of integration points in the mesh, which is the reason why discrete data sets can not be used in this context.
28.2 The Texture Component Method

28.2.2.1 Spherical Components

As mentioned above the spherical components are fully described by three measures: the preferred orientation \( g^c \), the scatter width \( b^c \) and the intensity or volume fraction \( I^c \). The preferred orientation is described in terms of Euler angles in Bunge notation \((\phi_1, \phi, \phi_2)\), which means that for every spherical texture component five numbers have to be passed into the FE calculation.

Then, in a first step \( g^c \) is assigned to all integration points (Figure 28.1a). In a second step the orientation of every integration point is modified statistically in such fashion, that the entity of all integration points reproduces the distribution function \( f^c \) (Figure 28.1b). In other words the orientation scatter \( b^c \) described initially by a texture component function \( f^c \) is in the finite element mesh represented by a systematically re-oriented set of orientations, each assigned to one integration point, which reproduces the original spherical scatter prescribed by that component. This means, that the scatter, which was originally only given in orientation space, is now represented by a distribution both, in real space and in orientation space, i.e. the initial spherical distribution is transformed into a spherical and lateral distribution.

28.2.2.2 Fiber Components

The fiber components are also fully described by three measures: the fiber axis \( f^c \), the scatter width \( a^c \) and the intensity or volume fraction \( I^c \). As the fiber axis is specified once in the sample coordinate system and once in the crystal coordinate system using spherical coordinates, in the case of the fiber components six numbers have to be passed into the FEM calculation. Then in a first step the crystal has to be rotated in such way that the direction of \( f^c \) specified in the crystal coordinate system coincides with that given in the sample coordinate system. After that a rotation about the fiber axis is performed using an arbitrary rotation angle. Finally a statistical rotation about an axis perpendicular to \( f^c \) is performed in such way that again the distribution function \( f^c \) (Figure 28.1b) is reproduced by the entity of all integration points.
The random texture component, $F$ in Equation (28.1), is mapped to the FE mesh by assigning one or more randomly picked orientation(s) to each integration point. Here it is important to ensure, that the number of integration points is high enough so that the material really behaves isotropic. Figure 28.2 illustrates this for a cup drawn from a circular sheet. Ideally the rim of the cup should be absolutely flat, but as the method uses discrete orientations there will always be some scatter. The influence of the number of randomly assigned orientations on this scatter is shown in Figure 28.3. It can be seen that a clear improvement over the use of one random component is achieved only if four random components are used. As this markedly increases the computing time in most cases the use of one random component is favorable.

After decomposing and representing the initial texture components as a lateral and spherical single orientation distribution in the mesh, the texture component concept is no longer required in the further procedure. This is due to the fact that during the subsequent crystal plasticity finite element simulation each individual orientation originally pertaining to one of the texture components can undergo an *individual* orientation change as in the conventional crystal plasticity methods. This means that the texture component method loses its significance during the simulation. In order to avoid confusion one should, therefore, underline that the texture component method is used to *feed* textures into finite element simulations on a
strict physical and quantitative basis. The components as such, however, are in their original form as compact functions not tracked during the simulation. It must also be noted that the orientation points which were originally obtained from the components do not represent individual grains but portions of an orientation distribution function.

The described re-orientation procedures are formulated as a weighted sampling Monte Carlo integration scheme in Euler space. It is important in this context, that the use of the Taylor assumption locally allows one to map more than one crystallographic orientation on each integration point and to assign to each of them an individual volume fraction $I_c$. This means that the procedure of mapping and rotating single orientations in accord with the initial texture component scatter width is individually conducted for all prescribed components extracted from initial experimental or theoretical data.

### 28.3 The Crystal Plasticity Model

The basic principles of crystal plasticity are introduced in Chapters 5 and 26. The crystal plasticity code used for the TCCP-FEM is based on the fully implicit integration scheme presented in Kalidindi et al. (1992). The only adaptation that is made concerns the treatment if more than one component is present at one integration point. In the original work the volume-averaged stress $\bar{T}$ is calculated as simple average of the Cauchy stresses $T^{(c)}$ of all components $c$. As the texture component method introduces the volume fraction $I_c$ of a component, the volume-averaged stress is calculated as

$$\bar{T} = \sum_{c=1}^{N} I_c T^{(c)}$$

(28.4)

in the TCCP-FEM, where $N$ is the number of components present at the integration point.
28.4 Application of the TCCP-FEM to Forming Simulation

The texture component crystal plasticity is implemented into the commercial FEM software MSC.Marc200x and ABAQUS/Standard by means of user subroutines. This makes it easy to use for almost any kind of forming simulation, also in an applied commercial environment, where direct texture implementation was up to now not common.

28.4.1 R-value Prediction

The so-called r-value is probably one of the most important measures for the in-plane anisotropy. It is typically required as a critical input parameter for analytical anisotropic constitutive material laws like Hill (1948) or Barlat and Lian (1989). The r-value is the ratio of the strain components $\epsilon_{33}$ and $\epsilon_{22}$ and is usually experimentally determined at a technical tensile strain of 20%. The angle between the original rolling direction and the tensile direction is used as index, e.g. $r_0$ refers to a tensile test in the rolling direction and $r_{90}$ refers to a tensile test in the transverse direction.

A tensile test to 20% technical strain can be easily simulated using the TCCP-FEM. As the material orientation is specified by Euler angles the angle between rolling direction and tensile direction can be easily adopted by modifying $\phi_1$. The r-value is then determined as the ratio of $\epsilon_{33}$ and $\epsilon_{22}$ for the center node of the mesh (see Figure 28.4). The simulated dependency of the r-value on the angle between rolling and tensile direction is shown in Figure 28.5 together with some experimental values for the deep drawing steel DC04. It should be mentioned here, that the specified strain rate sensitivities are those of the individual slip systems and need not necessarily coincide with the macroscopic strain rate sensitivity value of the sheet.

28.4.2 Prediction of Earing Behavior

Cup drawing is another standard material test for the characterization of material anisotropy. The cup drawing simulation shown in this section is for AA3104 hot band. Figure 28.6 shows on the left hand side the experimental {111} pole figure of the material. It shows a cube texture...
28.4 Application of the TCCP-FEM to Forming Simulation

Figure 28.5: $R$-value as function of the angle between rolling and tensile direction for two different values for the strain rate sensitivity $m$ of the individual slip systems.

Figure 28.6: Experimental (left) and recalculated (right) $\{111\}$ pole figure of Al hot band.

- Typical for hot rolled aluminum alloys. On the right hand side the pole figure recalculated from the texture component fit is shown. Besides the random portion of the texture only one spherical component was used to fit this rather pronounced cube texture (Table 28.1). Owing to the orthorhombic sample symmetry the single orientation has to be balanced by three additional symmetrically equivalent orientations in order to correctly reproduce the response
28.4.2 Optimization of Earing Behavior

In the early industrial practice texture was regarded as a property of polycrystals which was simply inherited from the preceding processing steps without conducting particular anisotropy optimization. This means that textures were known as an inevitable side-effect of materials processing which was hard to avoid and often accepted as it was. In contrast, modern industrial

<table>
<thead>
<tr>
<th>Euler angles</th>
<th>scatter</th>
<th>intensity</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\varphi_1$ [°]</td>
<td>$\phi$ [°]</td>
<td>$\varphi_2$ [°]</td>
</tr>
<tr>
<td>197.9</td>
<td>6.5</td>
<td>245</td>
</tr>
<tr>
<td>random</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 28.7: Experimental and simulated earing profile.

Table 28.1: Texture components used for the texture fitting.

of the material in the crystal plasticity finite element calculations. The resulting earing profile is shown in Figure 28.7 together with the experimentally measured earing profile. There is a very good agreement of simulation and experiment. Figure 28.8 shows the relative wall thickness distribution for the drawn cup (only one quarter has been simulated due to sample symmetry). It can be seen that the bottom of the cup is thinner while the upper part of its side is thicker than the original sheet. The thickening is most pronounced in the valley of the earing profile (i.e. in the $45^\circ$ position).
28.4 Application of the TCCP-FEM to Forming Simulation

Figure 28.8: Relative wall thickness of the drawn cup (brighter is thicker).

Figure 28.9: Earing is quantified in terms of the ear height and the ear area. The ear height describes the maximum occurring difference between the highest point on the rim and the lowest point on the profile of the drawn cup. The ear area integrates the entire surface between these two extremal values.

process design gradually aims at optimizing microstructures and properties during production, i.e. its goal consists in considering metallurgical mechanisms such as crystal plasticity, recrystallization, grain growth, and phase transformation for the design of well tailored crystallographic textures with respect to certain desired anisotropy properties of the final product.

The most recent phase in the advancement of quantitative texture and anisotropy engineering consists in the introduction of inverse texture simulation methods. Such approaches are designed for the physically based tailoring of optimum textures for final products under consideration of prescribed processing and materials conditions on an inverse basis. This means that variational texture optimization can nowadays be conducted in a way to match some de-
sired final anisotropy and can help to identify beneficial corresponding processing parameters. This amounts to a tenet change in the sense that the process should no longer determine the textures but the desired textures should determine the process.

As shown in the previous section elastic-plastic anisotropy during deep drawing may entail earing. One important consequence of that is – besides the irregular shape of the drawn specimen – an inhomogeneous distribution of the mechanical properties and of the wall thickness due to volume conservation and the kinematically necessary strain rate variation.

The trivial solution for the control and minimization of earing would be the presence of a random crystallographic texture prior to loading. However, such a supposed obvious approach is prevented due to two reasons. First, random starting textures do generally not remain random when the material is plastically deformed. This applies in particular to sheet forming operations. Second, complete spherical and topological randomization of textures is very difficult. Most metallurgical and mechanical processes promote rather reduce orientation distributions. This applies in particular to most face centered cubic (FCC) metals, in particular to those without bulk phase transformation during forming such as aluminum.

Therefore, a more practical approach for reducing shape anisotropy lies in combining the texture components constituting the initial sheet in such a way that the resulting ear profile – accounting also for texture changes during forming – can be minimized owing to the mutual compensation of the shape anisotropy contributions introduced by the individual texture components during forming.

As example a combination of Cube (\(\varphi_1 = 0^\circ, \Phi = 0^\circ, \varphi_2 = 0^\circ\)) and S (\(\varphi_1 = 50^\circ, \Phi = 35^\circ, \varphi_2 = 70^\circ\)) texture components for aluminum, which are typical annealing and rolling components of this material and produce opposite earing, are chosen. Ear profiles were computed for different combinations of the two components. The aim of this simulation series is the prediction of the optimum texture composition for minimum earing. Earing is

**Figure 28.10:** Minimization of ear formation by mixing the S and Cube texture components. a) Ear profiles for different ratios of S and Cube. b) Quantitative analysis of ear height and ear area. The simulation results show that the ear behavior can be minimized by mixing S and Cube texture components at an optimum ratio of about 62.75vol.% / 37.25vol.% (S/Cube).
quantified in terms of the ear height and ear area (Figure 28.9). The ear height describes the maximum occurring difference between the highest and the lowest point on the profile of the drawn cup. The ear area integrates the entire surface between these two extremal values.

Figure 28.10 shows the influence of a mixture of the S and the Cube texture components on the shape after cup drawing. The simulation results present the ear profiles for different ratios of the volume fractions of the two texture components. Figure 28.10a shows that an optimum profile is obtained for a combination of 62.75 vol. % of the S component and 37.25 vol. % of the Cube component in case that the orientational scatter width amounts to 0° prior to elastic-plastic loading (combination of two originally perfect single crystals). Figure 28.10b reveals that this results applies for both, the ear height and the ear area. Figure 28.10b also substantiates that the dependence of earing in terms of height and area on the texture composition reveals a steep change in the vicinity of the earing minimum, i.e. small modifications in the ratio of the volumes of the two texture components entail a strong change in earing. This means that even minor texture changes can lead to a remarkable optimization or degradation of the ear profile. It is important to note in this context that the ideal components with 0° orientational scatter width, which characterize the starting texture develop during the simulation into an array of similar orientations each of which may undergo individual reorientations according to the local boundary conditions. It is the special advantage of the texture component crystal plasticity finite element method to take these individual local reorientation- and strain-paths properly into account.

It is of some interest in this context to observe that the optimum volume ratio between the S orientation and the Cube orientation amounts to about 1.67:1 (62.75 vol. %:37.25 vol. %) rather than to 1:1. The initial assumption, that an optimum ratio between the two texture components might amount to 1:1 was suggested by Figure 28.10a which shows that the ear profiles created by the S or the Cube component alone are opposite indicating the possibility of mutual compensation. The fact that 1.67:1 and not 1:1 is the optimum ratio underlines that the interaction of different texture components is highly non-linear. It is also important to learn from this ratio that the volume fraction of the S component must obviously be much larger than that of the Cube component in order to compensate for anisotropy. This means that the Cube orientation has a much larger effect on the overall anisotropy during deep drawing than the S orientation.

28.5 Outlook

It was shown, that by combining the texture component method with the crystal plasticity finite element method the anisotropic forming behavior of metals can be adequately described. However, even though only small numbers of components are usually needed, the method still is computational quite demanding. Therefore, instead of directly applying it to more complicated forming simulations at the time being it seems to be more reasonable to use it as an virtual materials testing laboratory. As was shown by the first example it can be used to determine input parameters, r-values in this case, for more complex analytical yield surface formulations. These frequently rely on input values, which have to be determined in rather complicated experimental procedures such as multi axial loading, which are difficult to control and time consuming to set up. In this case simple mechanical tests like tension and texture
measurements can be used to evaluate the input parameters for the Texture Component Crystal Plasticity Finite Element Method, which can then be used to simulate the more complex tests.

References


K. Helming (1996), *Texturapproximation durch Modellkomponenten*, Cuvillier Verlag Göttingen, Germany


Microstructural Modeling of Multifunctional Material Properties: The OOF Project

R. Edwin García, Andrew C. E. Reid, Stephen A. Langer, and W. Craig Carter

Recent advances in and applications of the public domain Object Oriented Finite Element software for Materials Science (OOF) are discussed. The OOF software calculates the macroscopic properties from two-dimensional microstructures. It operates directly from microstructural image data to create a computational model that has the same spatial properties as the two-dimensional material microstructure. Recent progress in the code accounts for applied continuum thermal, electrostatic, elastic and chemical potential fields and how they are locally coupled through properties at a microstructural level. We illustrate the application of the software by computing the macroscopic properties of polycrystalline piezoelectrics. We also demonstrate the effect of microstructure on the response of a porous cathode in a rechargeable lithium ion battery and calculate the resulting chemically induced (Vegard) stresses.

We also discuss the impending public release of the OOF2 code, which allows the user to simulate nearly arbitrary sets of applied fields and nearly arbitrary material constitutive equations in microstructures, using an intuitive graphical interface.

29.1 Introduction

Many advances in materials applications come from improvements in microstructures. The optimal microstructure, in turn, depends on how a material’s microstructure interacts with a device during service. Microstructures can be arbitrarily complex and their interaction with applied fields is a challenging materials science problem. In some cases, the lack of predictive methods for particular microstructure’s macroscopic response to applied fields may defeat attempts to optimize a microstructure for a specific application.

OOF is a public domain software program that was created for purposes of analyzing the properties of microstructures. In general, computational modeling should enable rapid screening of candidate materials and microstructures, as well as identification of the most promising candidates for subsequent laboratory study. By simulating material behavior, modeling may guide laboratory tests, accelerate design cycles, and reduce risks of innovation.

For example, electric and magnetic fields may couple to other fields, such as strain, composition and temperature. All field properties are spatially coupled to the microstructure through material constitutive relations. Examples of such couplings are found in piezoelectric, electrostrictive, opto-electric, photo-elastic, magneto-electric, magnetostrictive, piezomagnetic,
and other materials. Still other couplings can be found in colloidal, corrosive, and fuel cell systems, where the Maxwell equations must be solved for particular spatial distributions of electrochemical potential. In conducting materials, while an externally applied magnetic field affects the shape of the solid through the piezomagnetic or magnetostrictive effect, the local electric field is coupled to the compositional field through the electromigration effect. Moreover, concentration and stress fields are coupled through the local stress free strain or Vegard's law. In materials that can conduct charged particles, the spatial distribution of charge will spontaneously adjust until the electrochemical potential (Fermi level, for electrons) is uniform. These redistributed electric fields can affect local charge population and transport properties—such as in diodes and other devices.

Generally, in multifunctional materials the application of one or more fields affects the values of other material properties. The spatial couplings that link the different materials at the microstructural level determine the macroscopic response. Therefore, methods, such as OOF that treat multiple fields and microstructural properties should be valuable tools for materials research.

An accurate simulation of multiply coupled fields in microstructures requires the simultaneous solution for each of the physical laws that pertain to the material's behavior. In OOF, this approach is facilitated by the introduction of microstructural constitutive equations and a data structure that derives from geometrical aspects of a particular microstructure.

To investigate the effects of microstructure on properties, two categorically different approaches may be considered: the fine scale effects can be spatially averaged and the microstructure approximated as a homogeneous material; or the specific spatial configuration of the properties of a particular microstructure can be incorporated into a complex model that treats all known interactions to the level of spatial resolution. The first, or homogenizing, approach will not describe those macroscopic properties that depend strongly on localized interactions but should be an economical method for calculation of those properties that depend on the mean of the microstructural distribution. In this chapter, we describe a method that adopts the second approach, utilizing all available microstructural data. This approach, implemented within the Object Oriented Finite element program (OOF), permits virtual iterations of design optimization through the use of simulations and analyzes multiple microstructural and material properties effects to a priori optimize product development and performance.

Generally, the homogenization approach derives from fundamental work by Rayleigh and Eshelby (Rayleigh (1892); Eshelby (1957); Budiansky (1965); Hill (1965); Hashin (1983)). However, if effects due to peculiarities of a particular microstructural feature (such as spatial correlations of crystallographic orientations, morphological texture, interface or boundary proximity effects, or defects such as pore and crack distributions) are to be calculated, homogenization approaches are very impractical. Homogenization methods that apply to cases where local variations in the microstructural fields are significant or have non-linear couplings (e.g., electrostriction, photoelasticity, electro-optic effects, etc.) would be useful, but are equally impractical. Moreover, when material response depends on improbable events or spatial correlations (e.g., when considering the reliability of piezoelectric devices with a dilute concentration of defects) homogenization approaches will not suffice.

The OOF approach provides a means of assessing the effect of such correlations, nonlinearities, and statistics. It can be used in samples with large spatial correlations or in materials that suffer the inadequacies of small number statistics, where the solution derived from
29.2 Program Overview

OOF is a finite element solver that operates on data structures which come from images. These images can be scanned from a cross-section of a real material, computer simulated, drawn in another software application, or drawn by hand and scanned into a graphics format. The software consists of two parts: PPM2OOF and OOF. PPM2OOF incorporates the microstructural and single-crystal property information through a graphic user interface. Here, the user can specify the properties of the different geometrical features of a material by using simple point-and-click operations or through an automated script. The digital image is a discretized representation to the real material. The microstructure is read by PPM2OOF in ppm (portable pixel map) format. The ppm format was chosen for its simplicity. There are many programs available for converting between the many different image formats; the choice of ppm in OOF’s software is inconsequential.

PPM2OOF includes algorithms for selecting microstructural features and assigning material properties to the pixels that represent them, as well as tools for mesh creation and manipulation—for example, the demography command, selects pixels based on their grayscale value or color. Similarly, the burn command (using an algorithm based on a forest-fire spreading model) selects a set of contiguous pixels given a user-specified rule for whether neighboring pixels of a selected pixel should also be selected.

PPM2OOF has a number of algorithms that modify an image in order to identify the different microstructural features. Examples include tools for non-linear smoothing of an image, image noise reduction while preserving the boundaries of the image. Other available algorithms are gray, maximum contrast, smooth, stencil, and equalize. Furthermore, PPM2OOF can operate on multiple versions of an image simultaneously, so the outputs of external image manipulation programs (e.g., an edge detector) be incorporated into PPM2OOF’s data structure.

In general, image processing, pixel selection, and material assignment require user input and judgment. Choice of PPM2OOF tools and the order in which they are applied will vary from micrograph to micrograph. The main goal of PPM2OOF is to provide an interface to

the homogenization approach would deviate greatly from the actual distribution of fields. The software is intended to improve processing design for devices, particularly when the device scale is similar to that of the microstructure or the relevant design features depend on extremes of microstructural distributions.

Ultimately, material function depends on the relationship between the spatial variations of the thermodynamic fields as well as the anisotropy of the underlying single-crystal properties of each microstructural feature. It is not straightforward to develop criteria to determine which properties are confidently predicted by an effective medium approach and those directly related to localized interactions which must be treated by the OOF-method.

This chapter briefly describes a framework for modeling microstructural effects on complex material microstructures. Recent advances in the modeling of materials with multiple physical and spatial couplings are presented and applications are demonstrated to particular material systems. In particular, we describe applications to polycrystalline piezoelectrics and rechargeable batteries.
rapidly prototype realistic finite element meshes. We refer the reader to the PPM2OOF manual for further details (Langer (2001); Langer et al. (2003)).

Figure 29.1 shows a very simple example in which a computer-generated microstructure is meshed. Here, the circles correspond to a set of hypothetical two-dimensional particles of lithium manganese oxide embedded in a composite matrix of polymeric electrolyte and graphite particles. The user could select individual particles by using the burn tool and assign material properties particle by particle; or, select the specific combination of red, green, and blue and assign properties to particles based on their color. After each phase is identified, a uniform triangular mesh is laid on top of the image and the properties are transferred from the pixels to the mesh. The mesh is improved by adapting the mesh to the underlying material properties. PPM2OOF has several algorithms for property-based mesh adaption. For many of the tools, an energy functional that depends on the local state of the mesh is minimized. This functional has two contributions from each element: homogeneity and shape. The homogeneity energy part reaches a minimum when all of the pixels within the element have the same material properties, and is a maximum for completely inhomogeneous elements. The shape part of the energy functional measures how far from equilateral the shape of a triangle is: it is minimal for an equilateral triangle and maximal for a degenerate triangle with three collinear vertices. The complete expression combines both the homogeneous and shape contributions and is used to move the nodes of the elements (e.g., anneal the mesh), or subdivide the triangular elements into smaller, more homogeneous ones while maintaining nodal compatibility (e.g., the command refine E, etc.).

OOF takes the mesh generated by PPM2OOF and performs finite element experiments. OOF, therefore, is applicable to any image for which properties can be locally specified, such as a material microstructure. The user can easily replace the element properties for a fixed mesh and perform “what-if” experiments on a microstructure and its possible modifications.

OOF has a graphical user interface through which the user can save the calculated results and incorporate them into papers or presentations. It has built-in functions designed to perform simple statistical analyses on all (or subgroups of) elements. It also has methods for data output, which can be used in other programs to perform post-calculation analyses.

From a numerical analysis, the user can go back and modify the input microstructure, materials properties parameters, or boundary conditions. Therefore, the user can rapidly correlate the response of the particular virtual material to the underlying physical mechanisms that control it and modify the processing technique that controls the behavior in the real material.

The user can create different microstructural instances for the same microstructural parameters, thus allowing the user to generate statistics over the different microstructural features by exploring a wide range of microstructural parameters. Therefore, the user can perform reliability calculations and assess the quality of the produced virtual material or device in terms of a simple set of microstructural or processing parameters (Cannillo (2000)).

The publicly available version of OOF contains the physics of the thermal diffusion and linear elasticity equations, which can be coupled via a material dependent thermal expansion

---

2 OOF2, described below, will include more general meshes.

3 The mesh generation process can be automated by combining many of the existing commands into user-defined scripts. A script that works well for one image will (probably) work well for many similar images; therefore, after meshing one micrograph from a set interactively, the rest of the set may be meshed in batch mode.
29.2 Program Overview

Figure 29.1: A sequence of steps illustrating the mesh generation process in PPM2OOF. After assigning material properties to the underlying pixels, a mesh is created and adapted to the microstructure in order to approximate the geometrical features of a real material. (a) initial uniform mesh. (b) refinement of inhomogeneous elements. (c) annealing or moving nodes to improve homogeneity. (d) final mesh, after further refinement and annealing operations.

This version of OOF has been extended (Garcia (2003)) to include more physical properties (chemical diffusion of charged species, chemically induced stresses, piezoelectricity, electrostriction, etc.), non-linear solvers, and mesh refining techniques. The newly introduced physics allow the description of the microstructural fields and macroscopic properties of polycrystalline piezoelectrics and electrochemical materials, such as rechargeable lithium ion
batteries. (Although the modified version is not publicly available, the extensions are being merged into OOF2, currently under development at the time of writing.)

29.3 Modeling of Piezoelectric Microstructures

One example of a multifunctional material is a piezoelectric polycrystal. For such a material, a thermodynamically reversible change in the state of polarization and stress is represented by two equilibrium state equations:

\[ D_i = \varepsilon_{ij} E_j + d_{im} C_{lmjk} \varepsilon_{jk} \]  
\[ \sigma_{ij} = C_{ijkl}^E (\varepsilon_{kl} - d_{mkl} E_m) \]

Here, \( \varepsilon_{ij} \) is the \( ij \)-th component of the total deformation tensor (i.e., the sum of two contributions, one due to elastic deformation in the absence of an electric field and the other due to a stress-free application of an electric field) and \( E_k \) is the \( k \)-th component of the electric field. \( d_{mkl} \) is the piezoelectric tensor, \( C_{ijkl}^E \) is the stiffness tensor at constant electric field, and \( \varepsilon_{ij} \) is the dielectric permittivity at constant strain. Equation (29.2) corresponds to Hooke’s law. The second term of this equation is the contribution to strain due to the converse piezoelectric effect. Equation (29.1) describes the different contributions to polarization in the solid. In some materials and crystal structures, higher order effects may become important under particular boundary conditions.

Equations (29.1) and (29.2) are inserted into the local mechanical and electrical equilibrium equations:

\[ \nabla \cdot \vec{\sigma} = 0 \]  
\[ \nabla \cdot \vec{D} = 0 \]

which describe the continuity of forces across geometrical interfaces, together with the continuity of polarization in the absence of gravity and net electrostatic charges. The second equation is one of Maxwell Equations for static fields.

From this framework, the effects of crystallographic texture and grain size on the macroscopic properties of polycrystalline piezoelectrics can be found by analyzing linear and nonlinear electromechanical responses in OOF. For example, Figure 29.2 illustrates how a specific value of texture induces spatial variations of the electrical and mechanical fields. These fields will eventually lead to depoling or cracking of the material. Furthermore, a statistical analysis over a large population of simulated microstructures permits the analysis of the full range of texture parameters. Figure 29.3 illustrates the case of polycrystalline barium titanate. The error bars show the standard deviation from performing several numerical experiments on different instances of microstructures with the same texture parameters.

The modified OOF program uses adaptive refinement techniques to achieve higher spatial resolution where the fields are rapidly changing and uses a Newton-Raphson solver on the nonlinear electromechanical equations. These algorithms enlarge the variety of physical problems that the OOF software can describe.

From the example calculation shown in this section, one can envision several applications of the electromechanical version of OOF. For example, the effects of particle size, volume
Figure 29.2: Comparison of the distribution of mechanical and electrical energy density fields for two different sets of texture parameters. The top row shows the original microstructure; middle row, 10% unpoled, perfectly texture sample; and bottom row shows energy distribution in an unpoled, untextured piezoelectric. As the polycrystal becomes increasingly disordered and unpoled; the electrical energy density distribution is balanced by the induced mechanical energy.
Figure 29.3: The effect of anisotropy on the macroscopic, normalized piezoelectric response of polycrystalline barium titanate as a function of texture parameters $r$ and $b$. $r$ governs the degree of alignment of the crystalline $\hat{c}$ axes in the different grains, with $r = 1$ completely uncorrelated and $r = 0$ completely aligned. $b$ reflects the probability that a given grain is polarized in the positive $\hat{c}$ direction. When $b = 0$ all grains are polarized towards $-\hat{c}$, $b = 0.5$ indicates random polarization, and $b = 1$ means all grains are polarized along $+\hat{c}$. The microstructures are polycrystals with 191 grains. $d_{31}/d_{31}^{\text{xtal}}$ is the macroscopic piezoelectric constant ($d_{31}$) normalized by its single crystal value. $a$ is predicted experimental value as predicted by Jona et al. Jona et al. (1965)

fraction, distribution, and texture in piezoelectric fibers; the induced electromechanical fields (and their interactions) during the switching of the local polarization state for PZT-based memories; the geometrical optimization of MEM-type piezoelectric actuators and sensors, are just a few. Possible extensions of this framework consist of including the effects of domain wall motion and grain microstructure.

29.4 Modeling of Electrochemical Solids: Rechargeable Lithium Ion Batteries

A rechargeable-lithium ion battery is an electrochemical device designed to provide a specific electron flux (current) and macroscopic voltage. A battery is formed of three parts: anode, cathode, and separator. In the simplest device, the anode stores a finite amount of charge in form of lithium atoms and the cathode accommodates the lithium ions in a host (also active)
The separator is an electrolytic composite abutted by the anode and cathode. A typical rechargeable battery is illustrated in Figure 29.4.

OOF was further modified to study the delivered power and energy density of lithium ion batteries. This version (called the electrochemical version, to distinguish it from the electromechanical version discussed above) uses the large electrochemical potential difference between cathode and anode as the driving force to homogenize the concentration of lithium across the device. The modeled situation corresponds to the case where the anode and cathode are ohmically connected via a resistive load, so the excess electrons will find a transport path to the cathode. This results in a macroscopic electron current and voltage drop from all the simultaneously occurring physical mechanisms.

This framework permits a description of the induced localized fluxes and electrochemical potential gradients resulting from the inherent microstructural features of the battery system. In a high energy density device, lithium ions diffuse through the polymer electrolyte matrix; percolate between the particles of the host material in the front region of the cathode to intercalate into those particles resting on the middle and back regions of the cathode. If the density of particles in the front of the cathode is very high, the number of lithium ions that diffuse to the back will be very low. Some of the active material is essentially unavailable, and the power and energy density of the device are diminished.

All the above physical and microstructural mechanisms were accounted by introducing a modified version of OOF. In this framework, the lithium mass and total charge fluxes are given by the expressions:

\[
\begin{align*}
N^r_{Li} &= -D_{Li} \nabla c_{Li} - M_{Li} c_{Li} \nabla \phi \\
\overrightarrow{J}_q &= -M_{Li} c_{Li} \nabla c_{Li} - \kappa \nabla \phi
\end{align*}
\] (29.5)
The mass flux of lithium ions is $\vec{N}_{Li}$ and the total flux of charge is $\vec{J}_q$. The product $M_{Li}c_{Li}$ is the electromigration mobility for ionized lithium, $\kappa$ is the local electrical conductivity, and $c_{Li}$ is the concentration of lithium. These material parameters can be replaced by their rank-two tensor analogs. The OOF program computes a mass and charge balance in order to account for the instantaneous distribution of the fields. The final set of equations has the form:

$$\frac{\partial c_{Li}}{\partial t} = -\nabla \cdot (\vec{N}_{Li} = \nabla \cdot (D_{Li} \nabla c_{Li}) + \nabla \cdot (M_{Li}c_{Li} \nabla \phi))$$

$$\frac{\partial \rho}{\partial t} = 0 = -\nabla \cdot (\vec{J}_q = \nabla \cdot (\kappa \nabla \phi) + \nabla \cdot (M_{Li}c_{Li} \nabla c_{Li}))$$  \hspace{1cm} (29.6)

An important physical mechanism is the oxidation/reduction process that occurs at the surface of the active material. Here, lithium ions react with electrons and are “inserted” into the cathode particle matrix if the local interfacial potential deviates from its equilibrium value. The constitutive equation that controls the equilibrium surface is known as the Butler-Volmer relation:

$$\vec{J} \cdot \hat{n} = i_o \left\{ e^{\left(\frac{\eta e}{RT}\right)} - e^{-\left(\frac{\eta e}{RT}\right)} \right\} \hspace{1cm} \text{(29.7)}$$

The first term on the right side corresponds to the forward rate of the anodic process which is modulated by an empirical coefficient $\alpha_a$. Similarly, $\alpha_c$ is the empirical parameter associated with the rate of the backward cathodic reaction. The local overpotential is $\eta_e$. $\hat{n}$ is the unit normal vector to the local surface of the cathodic particle.

The functional form of the prefactor $i_o$ is

$$i_o = F k_r (n_T - n_{Li})^{\alpha_a} n_{Li}^{\alpha_c} \hspace{1cm} \text{(29.8)}$$

where $k_r$ is the reaction rate constant, and $n_T$ is the solubility limit of lithium in LiMn$_2$O$_4$ (Doyle (1995)).

Physically, Equations (29.7) and (29.8) describe the rate of lithium exchange at the particle-electrolyte interface. In the absence of diffusion, an overpotential $\eta_e \neq 0$ induces a flux of the involved species. At equilibrium, the forward and backwards rates of reaction are balanced.

In rechargeable lithium ion batteries the intercalation (and de-intercalation) of lithium into the active material results in a local volume change. For large discharge rates, the induced deformation incompatibilities lead to large stresses which, upon cycling, culminate in failure of the device. This is described by introducing a modified Hooke’s law:

$$\sigma_{ij} = C_{ijkl} (\epsilon_{kl}^{T} - \beta_{kl}(c_{Li} - c_o)) \hspace{1cm} \text{(29.9)}$$

where $\beta_{kl}$ is the Vegard tensor (or tensor of chemical expansion) and $c_o$ is the reference (strain-free) concentration. The mechanical fields are solved by substituting Equation (29.9) into Equation (29.3).

Figure 29.5 shows a typical simulation of battery response. The material properties and microstructural parameters were taken from Doyle (Doyle (1995)). The circles correspond to particles of LiMn$_2$O$_4$ which are embedded in a composite matrix of a polymeric electrolyte and carbon black. The voltage plots illustrate that chains of particles shield the active material.
on the back part of the cathode, influencing and thus decreasing the power density of the overall device. Chemically, at the end of the discharge process the surfaces of the particles saturate with lithium (therefore, inducing a drop on the local voltage), even though the core of the particles can accommodate up to ten percent more.

The electric response of this device is summarized in Figure 29.6. The transport limitations are reflected in a rapidly decreasing energy density with increasing extracted power. The predicted battery performance agrees with the trends predicted by several authors (Doyle et al. (1993); Doyle (1995); Hellweg (2000)).

The corresponding discharge stresses are illustrated in Figure 29.7 for $t = 3100$ s. As a consequence of the swelling of the surface of the cathode particles, the outer-most layers enter a compressive state of stress, while the cores of the particles enter a tensile state of stress in order to balance it. The background matrix composite provides no resistance to deformation. The simulation suggests that during charging of the device, the local state of stress will be reversed so the surface of the particles will be in tension and the core in compression. The simulation also suggests that further cycling of the device will induce fatigue failure starting at the surface of the active material.

Figure 29.5: The top row shows the voltage distribution inside the cathode towards the end of the discharge process for a battery with a one hour discharge time. The lower right plot shows the corresponding lithium concentration. The battery is effectively discharged when the surface of the particles is saturated with lithium ions, regardless of whether more lithium can be accommodated within each particle.
The electrochemical version of OOF makes use of a finite difference algorithm to resolve the time evolution of the fields. The user can select the type of finite difference scheme (implicit, explicit, semi-implicit) by changing a simple parameter. The user also has available a Newton-Raphson solver and a limited set of tools to assign properties to the internal interfaces (internal boundaries) and to specify a constant flux (Neumann) or field value (Dirichlet).

Although just one particular battery configuration is illustrated and analyzed here, a number of microstructural effects can be explored. For example, different distributions of particle sizes or mixed phases of active host material in anode and cathode are just a few of several immediate possibilities. Furthermore, electrically and chemically induced phase transitions can be described by explicitly tracking the interface delimiting two (or more) stable phases or using a phase-field approach. Texture effects can be readily assessed. An important possible extension of this work consists of including the detailed interactions between the anode and cathode section, i.e., to study a rocking-chair configuration (Vincent et al. (1997)). For all these instances, understanding of the properties and microstructural effects will yield an improvement of the device.
29.5 The OOFTWO Project: A Preview

Despite the demonstrated utility of the modified versions of OOF discussed above, we have not made those versions available on the web. Contrary to the intent of the original authors, adding new physics to OOF turned out to be fairly difficult. In addition, combining the electromechanical and electrochemical versions into one program would have required an ungainly and unusable user interface, and extending the program to still more physics would compound the problem. However, releasing and maintaining separate elastic, thermoelastic, electromechanical, and electrochemical versions is also impractical.

For these reasons, we are developing a completely new program, OOF2, which addresses the shortcomings of OOF1. The fundamental problem with OOF1, and the one that necessitated the complete rewrite of the code, is that its C++ object classes are not a sufficiently close representation of finite element mathematics or materials physics.

OOF2 should be available on the OOF website by the time this chapter appears in print.

Pluggable Physics

While original versions of OOF1 were hard-coded to solve elastic and thermal conductivity problems and the modified versions were built to solve electromechanical and electrochemical problems, OOF2 has been built without any hard-coded physics at all. Instead, its central engine solves problems in which linear combinations of derivatives of some generalized flux (e.g., stress) can be equated to generalized body forces,

\[ \nabla \cdot J = -f, \]  

(29.10)
and for which this generalized flux is given in terms of the fields of physical interest (e.g.,
displacement), \( \phi \), by a constitutive rule of the general form

\[
J = g(\phi, \nabla \phi),
\]

(29.11)

where \( g \) is an arbitrary function of the fields and their spatial derivatives. The function \( g \) can
be a scalar, vector, tensor, etc., and may include couplings between fields. This represents a
very broad class of problems of physical interest.

For a user of moderate computational sophistication, OOF2 provides a uniform interface
for the addition of user-specific constitutive rules \( g \). These objects, called “properties” in
the OOF2 parlance, must include blocks of code which take as input the indices of degrees of
freedom, corresponding to the physical fields of interest, and provide as output the constitutive
rule’s contribution to the appropriate flux. Various properties can be assembled at run time to
form a “material”. This is in contrast to OOF1, in which the properties of each material are
predetermined except for the numerical values of their parameters. Furthermore, OOF2 users
can define additional fields and fluxes by creating new instances of the field or flux class of
the appropriate dimensionality.

The basic physics embodied in OOF1 have already been implemented in OOF2 (i.e., dis-
placement and temperature fields, stress and heat fluxes, and elasticity and thermal conduc-
tivity constitutive laws), so users will not have to add these capabilities. But user-supplied
routines can not only extend the program’s capabilities to new physics, but also replace the
out-of-the-box constitutive rules with new ones, should such a requirement arise.

Mesh and Element Improvements

OOF1 implements only linear triangular finite elements, whose mathematics is simple enough
that there is no need for classes representing shape functions, master space elements, maps,
and the rest of the finite element tool set. By providing all these classes, OOF2 allows for
the construction of meshes with quadrilateral as well as triangular elements. Higher order
isoparametric, subparametric, and superparametric elements can be added easily (Zienkiewicz
et al. (2000)).

OOF2 includes new versions of the OOF1 mesh modification tools and is capable of re-
fining the mesh by subdivision of these elements in any combination. New tools have been
provided to handle, for example, the merging of adjacent triangles to a single quadrilateral, or
conversely the splitting of quadrilaterals into triangles.

The OOF2 Advantage

Initially, solvers and numerical infrastructure are being developed for the case where \( g \) is
a linear function of the fields or their derivatives and where there is no inertia, dissipation,
or time-dependence. Also, development is concentrating on the case where the equilibrium
equation is given by equating the divergence of a flux to the externally applied forces. This
gives OOF2 the same general constitutive scope as the currently available OOF1 code.

Greater sophistication will follow, drawing on the lessons learned in the applications of
OOF1 described here. The new code has a modular design, allowing for the later addition
of solvers capable of handling, for example, nonlinear or time-dependent constitutive relations. A simple prototype plasticity property has already been coded, addressing a frequently-requested capability, and nonlinear solvers from the PETSc (Balay et al. (2001)) project have been incorporated into the code base in anticipation of constitutive relations such as those demonstrated in the OOF1 context above.

OOF2 retains and expands the advantages of the existing OOF code, in particular the ability to construct finite-element meshes which respect the microstructure of real materials. Tools similar to those in OOF1 are present, allowing meshes to be constructed from micrographs or other imaging data.

This combination of features means that OOF2 is more capable than OOF1 on a variety of fronts and allows investigators to address materials problems not anticipated by the developers, on high-order meshes which are faithful to geometrically complex microstructures using their own custom constitutive rules.

References


A self-consistent matricity model has been developed to simulate the mechanical behavior of composites with two randomly distributed phases of interpenetrating microstructures. The model is an extension of the self-consistent model for matrices with randomly distributed inclusions. In addition to the volume fraction of the phases, the matricity model allows a further parameter of the microstructure, the matricity $M$ of each phase, to be included into the simulation of the mechanical behavior of composites with interpenetrating microstructures. The model is applied to the calculation of stress-strain curves of a 30vol.% ZrO$_2$/NiCr 80 20 composite and of stress-strain curves and strain distribution curves of an Fe/18vol.%Ag-composite as well as to a series of W/Cu composites and its validity and superiority upon previous models is demonstrated. The matricities of the phases influence the stress-strain behavior mainly within the bounds between $M = 0.3$ and $M = 0.7$. Beyond these bounds, there exists only a minor influence of matricity on the stress-strain behavior. Good agreement has been obtained between experiment and calculation with respect to the composites’ mechanical behavior and the matricity model is thus found to represent well metal matrix composites with interpenetrating microstructures. The matricity model can be applied to describe the mechanical behavior of arbitrary microstructures as observed in two phase functionally graded materials, where the volume fraction as well as the matricity of the phases vary between the extreme values of 0 and 1.

### 30.1 Introduction

Coarse isotropic interpenetrating microstructures are often found in powder metallurgically fabricated composites in the regime of 25–75% volume fraction, or result from the infiltration of a porous material with a molten metal of a lower melting point. The arrangement of the phases in most of these technical composites is usually random, resulting in an isotropic overall mechanical behavior of these composites. Besides the volume fraction, such a material requires at least one further parameter to describe the microstructure more closely. In this chapter the selected further parameter is the ‘Matricity’, which was first introduced by Poech (Poech and Ruhr 1993) for two-phase steels and later used by Soppa (Soppa 1995) for the characterization of Ag/Ni composites. This parameter can be measured from a representative micrograph of a microstructure via an image analyzing system and also be included in a finite element model which was developed for this kind of microstructure (Leßle et al. 1997, Schmauder et al. 1997a, b, Leßle et al. 1998). In this overview the model is applied to demonstrate the potential of the matricity parameter in analyzing the influence of phase arrangement.
30.2 Matricity

In the following, composites of two phases $\alpha$ and $\beta$ are considered. Matricity is defined as the fraction of the length of the skeleton lines of one phase $S_\alpha$, and the length of the skeleton lines of the participating phases

$$M_\alpha = \frac{S_\alpha}{S_\alpha + S_\beta} \quad (30.1)$$

By definition, the sum of the matricities of all phases equals to one

$$M_\alpha + M_\beta = 1 \quad (30.2)$$

To obtain the skeleton lines of a certain phase, we must select this phase within an image analyzing system and reduce the detected structure to a typically non-connecting line by maintaining the topology. In Figure 30.1 the matricities have been determined for a ZrO$_2$/NiCr 80 20 cermet which was powder metallurgically fabricated at the University of Dortmund (Willert-Porada et al. 1997). The structure parameters volume fraction $f$ and matricity $M$ have been determined to be $f_{\text{NiCr}} = 0.3$ and $M_{\text{NiCr}} = 0.2$.

![Figure 30.1](image)

**Figure 30.1:** Micrograph of a ZrO$_2$ / NiCr 80 20 composite with 30% volume fraction of ZrO$_2$ (dark phase). Left: greyscale picture, right: binary image with skeleton lines.

30.2.1 Matricity Model

To take both parameters into account for the calculation of the mechanical behavior of the composite, an extension of the self-consistent embedded cell model (Dong and Schmauder...
1996, Schmauder 2002) has been developed. The embedded cell model has been introduced to simulate the mechanical behavior of composites with randomly distributed inclusions, where the volume fraction of the inclusions is the main parameter in the model. To take the matricity as second microstructural parameter into account, the self-consistent embedded cell model has been extended by a second self-consistent embedded cell model (Figure 30.2). In this 'matricity model' we are able to define the matricity of the model in the same manner as the matricity is defined in a real microstructure.

**Figure 30.2:** Matricity-model (schematic) with skeleton lines to adjust the measured parameter 'Matricity' in the model via the factors $W_1$ and $W_2$.

First the single phases are reduced to skeleton lines. The lengths of the skeleton lines of the inclusions (Figure 30.2, left: $\beta$; right: $\alpha$) are zero as the inclusions are spherical and are, therefore, reduced to a point in the process of obtaining the matricity of the phase.

The lengths of the skeleton lines $S_\alpha$ and $S_\beta$ in the matrices are given as the circumference of a circle with a diameter which is calculated as the arithmetic average of the diameter of the embedded cell and the diameter of the inclusion phase (Figure 30.2, left: $S_\alpha$; right: $S_\beta$).

The diameters of the embedded cells are denominated as $W_1$ and $W_2$. The diameters of the inclusion part of the embedded cells depend on the volume fraction of the inclusions and the corresponding factors $W_1$ or $W_2$.

For the 3D case, the diameter of the inclusion $\beta$ (Figure 30.2, left) is derived as a function of $W_1$ and the volume fraction $f_\beta$ of $\beta$ in this cell as

$$D_\beta = W_1(\sqrt[3]{f_\beta})$$  \hspace{1cm} (30.3)

and, analogous for inclusion $\alpha$ (Figure 30.2, right) as

$$D_\alpha = W_2(\sqrt[3]{f_\alpha})$$  \hspace{1cm} (30.4)

Therefore, we derive the skeleton line lengths as

$$S_\alpha = \pi W_1 \frac{(\sqrt[3]{f_\beta} + 1)}{2}$$  \hspace{1cm} (30.5)
and
\[ S_\beta = \pi W_2 \left( \frac{\sqrt[3]{f_\alpha} + 1}{2} \right). \]  

(30.6)

By taking into account that
\[ f_\alpha + f_\beta = 1, \]

(30.7)

the matricity \( M \) can be calculated as a function of the sizes of the embedded cells and the volume fraction of one of the two phases, as the volume fraction of the phases is held constant in both parts of the matricity model

\[ M_\alpha = \frac{S_\alpha}{S_\alpha + S_\beta} = \frac{W_1 \pi \left( \frac{\sqrt[3]{f_\beta} + 1}{2} \right)}{W_2 \pi \left( \frac{\sqrt[3]{f_\beta} + 1}{2} \right) + W_1 \pi \left( \frac{\sqrt[3]{f_\alpha} + 1}{2} \right)} \]  

(30.8)

or

\[ M_\alpha = \frac{W_1 (\sqrt[3]{1 - f_\alpha} + 1)}{W_2 (\sqrt[3]{f_\beta} + 1) + W_1 (\sqrt[3]{1 - f_\alpha} + 1)} \]  

(30.9)

### 30.2.2 Adjusting Matricity in the Model

As can be seen in Figure 30.2, the volume fractions of the phases as well as the diameters \( W_1 \) and \( W_2 \) of the embedded cells are adjustable. To achieve a matricity \( M \) in the model, we first realize the measured volume fraction of the phases in the model and then calculate the diameters \( W_1 \) and \( W_2 \). The corresponding diameters \( W_1 \) and \( W_2 \) are obtained by rearranging Eq. (30.9)

\[ W_1 = W_2 (\sqrt[3]{1 - f_\beta} + 1) \frac{1 - M_\beta}{M_\beta (\sqrt[3]{f_\beta} + 1)}, \quad W_2 = 1.0 \quad \text{for} \quad M_\beta \geq 0.5 \]  

(30.10)

or

\[ W_2 = W_1 (\sqrt[3]{1 - f_\alpha} + 1) \frac{1 - M_\alpha}{M_\alpha (\sqrt[3]{f_\alpha} + 1)}, \quad W_1 = 1.0 \quad \text{for} \quad M_\alpha \geq 0.5. \]  

(30.11)

### 30.2.3 Realisation of the Adjustability of Matricity by Weighting Factors

If the geometrical boundary conditions are modelled at a distance of about five times the radius of the embedded cell, they are of almost no influence on the models’ mechanical behavior. If we take care that the boundary conditions keep remote we can model the embedded cell with the surrounding composite in different manners (Figure 30.3). As the remote boundary conditions have almost no influence on the mechanical behavior of the embedded cell it is assumed that the continuum mechanical stress-strain state in the embedded cell is hardly influenced as
well. Taking this into account, a unit cell for a specific volume fraction can be used for each part of the matricity model. Further we can see from the almost independency from remote boundaries, that it is not necessary to model the matricity as a parameter of the FE-mesh but it is possible to introduce the matricity adjusting weighting factors $W_1$ and $W_2$ only in the evaluation of the results from the inclusion type geometries. As the results have to be determined by an iterative calculation in about 3–5 iterations, the adjusting weighting factors $W_1$ and $W_2$ must be introduced in the evaluation of all iteration steps (Figure 30.4).

### 30.2.4 Calculation of Stress-strain Curves

In principle, stress-strain curves of the two-phase composite are determined from the matricity model in the same iterative manner as it is done for the simple self-consistent embedded cell model. In each increment the components for stress and strain are determined. This is done by a weighted averaging of the stress and strain values over all integration points of both embedded cells. The three-dimensional weighting is done by the ‘integration point volume’ $V_{k0}$ of each corresponding Gaussian integration point, which must be multiplied by $W_1^3$ and $W_2^3$, respectively, to account for the matricity effects as described above. The factors $W_i$ are in the power of three as the length of the skeleton lines depend linearly on $W_i$ but the embedded
cell volumes depend by the power of three (for the 3D case) on $W_i$. With these considerations the stress and strain components can be calculated as

$$
\sigma_{ij} = \frac{\sum \sigma_{ij}^k V_k}{\sum V_k}, \quad i, j = (x, y, z) \quad \text{or} \quad (r, z, \Phi),
$$
(30.12)

$$
\varepsilon_{ij} = \frac{\sum \varepsilon_{ij}^k V_k}{\sum V_k}, \quad i, j = (x, y, z) \quad \text{or} \quad (r, z, \Phi),
$$
(30.13)

or more detailed (only for the stress components)

$$
\sigma_{ij} = \frac{\left(\sum \sigma_{ij}^1 V_{k0} W_1^3\right)_1 + \left(\sum \sigma_{ij}^2 V_{k0} W_2^3\right)_2}{\sum V_{k0} W_1^3 + \sum V_{k0} W_2^3},
$$
(30.14)

where $k$ is the index of summation and 1 or 2 are the part of the matricity model whose embedded cell is weighted by $W_1$ or $W_2$.

From the stress components, the von Mises equivalent stress at each strain increment is calculated as (Eqs. (30.15) and (30.16) are only valid for Cartesian coordinates)

$$
\sigma_v = \sqrt{\sigma_{xx}^2 + \sigma_{yy}^2 + \sigma_{zz}^2 - (\sigma_{xx}\sigma_{yy} + \sigma_{yy}\sigma_{zz} + \sigma_{zz}\sigma_{xx})} + 3(\sigma_{xy}^2 + \sigma_{yz}^2 + \sigma_{zx}^2)
$$
(30.15)

and the equivalent strain is given as

$$
\varepsilon_v = \frac{1}{1 + \mu} \sqrt{\varepsilon_{xx}^2 + \varepsilon_{yy}^2 + \varepsilon_{zz}^2 - (\varepsilon_{xx}\varepsilon_{yy} + \varepsilon_{yy}\varepsilon_{zz} + \varepsilon_{zz}\varepsilon_{xx})} + 3(\gamma_{xy}^2 + \gamma_{yz}^2 + \gamma_{zx}^2),
$$
(30.16)

where $\mu$ is the Poisson’s ratio of the composite.

### 30.2.5 Mechanical Constants

Poisson’s ratio $\mu$ and tangent modulus $T_i$ at zero strain are calculated from the obtained stress and strain components $\sigma_{ij}$ and $\varepsilon_{ij}$. Poisson’s ratio $\mu$ is calculated from the context of elastic constants (Sautter 1995).

$$
\mu = \frac{3K - E}{6K},
$$
(30.17)

where $K$ is the bulk modulus and $E$ is the Young’s modulus. $K$ is defined as (Sautter 1995)

$$
K = \frac{\sigma_H}{3\varepsilon_H} = \frac{(\sigma_{xx} + \sigma_{yy} + \sigma_{zz})}{3(\varepsilon_{xx} + \varepsilon_{yy} + \varepsilon_{zz})},
$$
(30.18)

If we consider a material where the elastic modulus $E$ equals the gradient of the stress-strain curve at zero strain, then we can determine the constants $K$ and $E$ from the calculated stress-strain curve. As the moduli change with changing strains we have to extrapolate the bulk modulus $K$ and the tangent modulus $T$ from values near zero-strain to a value at zero-strain (Figure 30.5).
30.2.6 Yield Stress

The yield stress is calculated as the stress belonging to the cross point of the stress-strain curve and a straight line through the 0.2% strain point with the gradient that equals to the above calculated tangent modulus. The model does not include any damage parameter or failure criterion. This might lead to unrealistic high yield stresses for composites with dominating linear elastic material behavior.

30.3 Results and Discussion

The quality of the matricity model to simulate stress-strain curves and strain distribution frequencies especially for two ductile phases has been demonstrated in (Leßle et al. 1997, Schmauder et al. 1997a, b, Leßle et al. 1998). In this section we compare the matricity model with other models and examine the influence of the matricity parameter on stress-strain curves of different composites and on the influence of residual stresses.

30.3.1 Comparison to Cluster Parameter $r$

The cluster parameter $r_\gamma (\gamma = \alpha, \beta)$ is defined in (Sieg mund et al. 1993) as

$$ r_\gamma = \frac{N_\gamma}{N_\alpha + N_\beta}, \quad (30.19) $$

where $\gamma = \alpha, \beta$ and $N_\gamma$ is the number of clusters of phase $\gamma$.

Figure 30.6 shows some model microstructures which have been calculated with distinguishable assumptions. To compare with the numerical yield stress results based on Eq. (30.19) (Sieg mund et al. 1993) we modelled the microstructures also with the matricity model. To do this, the matricity of the microstructures were derived by phase reduction.
30.3 Results and Discussion

Figure 30.6: Model microstructure of regular hexagons by Siegmund et al. (1993).

Figure 30.7: Determination of Matricity M for the hexagonal model microstructure with $f_B = 0.5$ (Siegmund et al. 1993). Left: original microstructure, right: binary image with skeleton lines.
Figure 30.8: Yield stress of α/β composite. (*, ** = in (Siegmund et al. 1993), calculated from (Bao et al. 1991)).

(c.f. Matricity) which is shown in Figure 30.7. The comparison of the matricity model results with the results in (Siegmund et al. 1993) shows good agreement with real structure calculations for both models (Figure 30.8). The advantage of the matricity model might be that differently sized clusters provide a different contribution to the influence of this cluster to the overall mechanical behavior, whereas the cluster parameter model assumes that all clusters, independent of the cluster sizes, are equally weighted. An obvious advantage of the matricity model is the good agreement of calculated strain distribution frequencies to experimentally obtained results (c.f. Leßle et al. 1997, Schmauder et al. 1997a, b, Leßle et al. 1998).

30.3.1.1 Matricity and Stress-strain Curves in Case of ZrO$_2$ (30%) / NiCr 80 20 (70%)

The influence of the matricity on the overall mechanical behavior can be shown by comparing the calculated stress-strain curves while keeping the volume fraction $f$ of the phases constant and by varying only the matricity $M$ of the phases (this means to vary $W_i$ correspondingly for evaluating the results). A parameter study was made for a ZrO$_2$/NiCr 80 20 composite where the volume fractions $f$ of the phases were kept constant and the matricity $M$ was varied between 0 and 1. This means that the microstructure lies in a range between a pure inclusion microstructure and a pure matrix microstructure for one phase and vice versa for the other phase. In Figure 30.9, stress-strain curves are shown for two different matricities $M_{ZrO_2}$ (= 0.3 and 0.4) with and without residual stresses caused by cooling down the material from
30.3 Results and Discussion

Figure 30.9: Stress-strain curves of ZrO$_2$/NiCr 80 20 composite with $f_{ZrO_2} = 0.3$ and two different Matricities, $M = 0.3$ and $M = 0.4$. ($E_{ZrO_2} = 206$ GPa, $E_{NiCr8020} = 214$ GPa, $\alpha_{ZrO_2} = 10E−6$ l/K, $\alpha_{NiCr8020} = 14E−6$ l/K).

processing to environmental temperature. Figure 30.10 shows the influence of the matricity on the yield stress. The influence of matricity is decisive in the range between $M = 0.3$ and $M = 0.7$. Taking residual stresses into account we can recognize that the influence of matricity on the yield stress diminishes. This can be explained with the modelled isotropic hardening of the ductile NiCr 80 20 phase, which is plastically deformed by residual stresses when acting as a matrix. However, no plastification is obtained when this phase acts as a pure inclusion and, therefore, hardly any hardening of the composite will be expected.

30.3.1.2 Matricity and Stress-strain Curves in Case of Fe(82%) / Ag(18%)

The matricity model was applied to simulate the mechanical behavior of a Fe(82%)/Ag(18%) composite under plane strain (PSn), plane stress (PSs) and axial symmetry (ASY) conditions. The matricity parameter of this composite was determined as $M_{Fe} = 0.63$, Figure 30.11. In Figure 30.12, a comparison is made of uniaxial stress-strain curves with the experimentally derived curve from (Bornert et al. 1994). As expected, the behavior of the plane strain model (PSn) behaves stiffer and the plane stress (PSs) model behaves more compliant as the experimentally observed behavior. Best agreement is obtained with the axially symmetric model (ASY) as it approaches the three-dimensional microstructure best.

In Figure 30.13 and Figure 30.14 the calculated frequency distributions (PSs, $M_{Fe} = 0.6$) are compared with experimental results (Bornert et al. 1994) and information theoretical results from Kreher (Kreher 1995). It is found that the results which are based on the matricity model and the experimental observations of deformation distributions at the material surface are in close agreement.
Figure 30.10: Yield stress of ZrO$_2$/NiCr 80 20 composite with $f_{ZrO_2} = 0.3$ and different Matricities. ($E_{ZrO_2} = 206$ GPa, $E_{NiCr8020} = 214$ GPa, $\alpha_{ZrO_2} = 10E−6 1/K$, $\alpha_{NiCr8020} = 14E−6 1/K$, $\Delta T = −750$ K).

Figure 30.11: Two phase microstructure of a Fe(82%)/Ag(18%) composite with skeleton lines (black = Fe, white = Ag).
30.3 Results and Discussion

Figure 30.12: Stress-strain curves of an Fe(82%)/Ag(18%) composite ($M = \text{Matricity of Fe}$). ($\text{PSn} = \text{plane strain condition}, \text{PSs} = \text{plane stress condition}, \text{ASY} = \text{axial symmetry condition}$).

Figure 30.13: Strain distribution in the silver phase of an Fe(82%)/Ag(18%) composite at 5% total strain ($M = \text{Matricity of Fe}, \text{PSs} = \text{plane stress condition}$).
30.3.1.3 Matricity and Stress-strain Curves in Case of W/Cu

W/Cu compositions in the range of 3%–75% Cu were produced by a sintering and infiltration method (Jedamzik et al. 1997). The corresponding microstructures are presented in Figure 30.15. Besides the volume fraction, the matricity parameter $M$ is also given in the Figure. In Figure 30.15 the matricities have been determined for a graded W/Cu composite. The structure parameters, volume fraction $f$, and matricity $M$, have been determined to be, e.g., $M_{Cu} = 0.08, 0.26, 0.45$ and $0.72$ for $f_{Cu} = 0.03, 0.21, 0.40$ and $0.75$, respectively. It can be seen that although there exists a nearly linear relationship, $f$ and $M$ are distinct parameters (Figure 30.16) and the slope is clearly less than 0.9. $M$ is linearly related to the cluster parameter $r_\alpha$: $M = a r_\alpha + b$, $a = 0.4, b = 0.3$ (Appendix).

The experimental data of Young’s modulus were determined using a resonance method according to ASTM C1259-94. For this purpose quasi-homogeneous rectangular bars of 0.4 mm thickness $t$, 4 mm width $b$ and 22 mm length $L$ were cut from the W/Cu sample using a wire saw. A piezoelectric transducer was attached to one end of the bars and a magnetic pickup to the other end with a thin thread. The frequency $F$ of the fundamental flexure mode of the bars was determined by a computer-controlled frequency generator/lock-in amplifier combination. The dynamic Young’s modulus $E$ of each region was then calculated from the mass $m$ and the dimensions of the bars according to ASTM Designation C 1259-94 (ASTM 1994) as

$$E = 0.9465 \cdot \left(m \cdot \frac{F^2}{b}\right) \left(\frac{L^3}{t^3}\right) \left(1 + 6.585 \cdot \frac{t^2}{L^2}\right).$$  

(30.20)

The elastic behavior of the composite is represented in Figure 30.17 where the Young’s modulus is seen to vary between those of the components at $f = 0$ and $f = 100\%$. It is
30.3 Results and Discussion

Interesting to see that the predictions of the upper and lower bounds from the Tuchinskii model are rather close to each other, while the moduli of W and Cu differ by a factor of 3.16. The experimental values fall into these bounds and the matricity model predictions are in the same range. The experimental data are very well described by the Pompe model (Kreher and Pompe 1989). The thermal expansion coefficient of the composite as observed in the experiments obeys the rule of mixture (between $\alpha_W = 4.7E-6 \cdot 1/K$ and $\alpha_{Cu} = 18.9E-6 \cdot 1/K$) in good approximation and so does the calculation when thermal stresses are taken into account (Figure 30.18). The agreement between the simulation and the Pompe model is very good. Earlier results on metal/ceramic composites have shown that the thermal expansion coefficient $\alpha$ is nearly independent on $M$ (Leßle 1997).

The model predictions for the stress-strain curves of W/Cu (with the corresponding volume fractions and matricities of the real composites) are depicted in Figure 30.19, where a strong variation in the mechanical behavior is seen for $f_{Cu} = 0.4 \div 0.7$. The influence of residual
Figure 30.16: Correlation between volume fraction and Matricity for different W/Cu microstructures.

Figure 30.17: $E$-Modulus for the different W/Cu microstructures. (*) = in (Jedamzik et al. 1997)).

Figure 30.18: Thermal expansion coefficients of the W/Cu microstructures. (*) = in (Jedamzik et al. 1997)).
30.4 Conclusion

The influence of the parameter matricity $M$ besides the parameter volume fraction $f$ on the overall behavior of two phase composites with coarse interpenetrating microstructures is clearly shown. Especially in cases, when the composite consists of a ductile and a linear-elastic phase, a significant influence of the matricity on stress-strain curves has been found. It is also found that the influence of residual stresses on yield stress depends strongly on the volume fraction and the matricity of the ductile phase. The comparison with model structure stresses caused by cooling down the material from processing to environmental temperature is also indicated in this figure. Obviously, no strong influence of the residual stresses is observed. Nevertheless, the residual stress distribution can be predicted reliably in either phase, as shown in (Leßle et al. 1998). Decoupling of the parameters ($f_{Cu} = 50\%$, $M$ = variable) demonstrates that phase arrangement is the parameter which besides volume fraction actually controls the mechanical behavior of the W/Cu composite (Figure 30.20).
calculations shows the high quality of the matricity model. The matricity model must be built once for each volume fraction and the matricity is then considered in the evaluation of the result during the iterative calculation process. The model is, therefore, very simple with respect to FE-meshing, and can be quite coarse, as the stress-strain components were evaluated as arithmetical averages over the elements of both embedded cells.

Appendix

Siegmund et al. introduced the so-called cluster parameter $r_\alpha$, which gives the relative number of clusters $NC(\alpha)$ of phase $\alpha$ in phase $\beta$ in an $\alpha - \beta$-composite $r_\alpha = NC(\alpha)/(NC(\alpha) + NC(\beta))$. In (Siegmund et al. 1993) this parameter was calculated for a number of artificial microstructures consisting of equally sized hexagons. More recently, (Leßle 1996) have measured the matricity parameter for the same microstructures. It was found that $M$ and $r_\alpha$ are linearly correlated. The superiority of $M$ in comparison to $r_\alpha$ is due to a wider range and a simpler determination of the matricity value (Leßle 1996).

Acknowledgements

Financial support by the German Research Society (DFG) under grant Schm 746/12-1 - 12-3 is gratefully acknowledged. Thanks are due to Ulrich Weber for the final preparation of the manuscript.

References


31 Creep Simulation

Wolfgang Blum

31.1 Introduction

It is common practice to subdivide the field of inelastic deformation into time-independent plasticity and time-dependent creep. The two are investigated by different experimental techniques employing quite different testing machines, plasticity in strain control (e.g. so-called tensile test) and creep in load control (creep test). According to the prescribed value of rate \( \dot{l} \) of change of specimen length \( l \), conventional tensile tests are limited in time \( t \); work hardening causes a continuous increase of flow stress \( \sigma \) (Figure 31.1).

In a creep test (Figure 31.1) work hardening occurs during loading up to the prescribed level of flow stress \( \sigma \) and continues at constant \( \sigma \) with strong decrease in creep rate (which due to constancy of stress and elastic strain automatically equals the inelastic strain rate), so that the duration of test may increase beyond experimental limits. In order to avoid that, the test temperature \( T \) is raised so that either fracture or the steady state of creep with dynamic equilibrium of work hardening and dislocation recovery is reached within accessible time.

In consequence of these experimental differences, plasticity and creep are felt to be rather fundamentally different, plasticity displaying rapid work hardening at low homologous temperatures and creep displaying slow inelastic deformation at high homologous temperatures close to steady state. However, this is a misconception. The focus on time and minimum

---

Figure 31.1: Schematic of work hardening at constant strain rate \( \dot{l}/l \) and in creep where stress \( \sigma \) is held constant after loading. The drop in \( \dot{l}/l \) at the transition to stress constancy marks disappearance of elastic strain rate.
creep rate $\dot{\epsilon}_{\text{min}}$ determining the creep life leads people to overly condense the information contained in creep tests to $\dot{\epsilon}_{\text{min}}$, called the creep rate, as the only property being reported, so that its validity can no longer be checked, and to overlook that inelastic strain $\epsilon$, not time $t$, is most closely related to structural evolution. To make things worse, concentration on tensile creep superimposes plastic instability by necking, which may significantly enhance the measured minimum creep rate. In spite of their unclear meaning the $\dot{\epsilon}_{\text{min}}$-data are taken as a basis for identifying mechanisms of creep, which are traditionally believed to be characterized by stress exponents $n$ and activation energies $Q$ of creep rates. This perception of creep as a special high-temperature mechanism of plasticity is ubiquitous and has entered virtually all treatments of plasticity, see e.g. the deformation mechanism maps of Frost and Ashby (1982) containing separate fields devoted to steady state creep.

However, the full view of creep is obtained only, if one accepts the physical fact that creep of crystalline materials usually is just part of crystal plasticity, occurring by crystallographic slip. Thus high temperature creep as well as low temperature work hardening will follow from the same model of crystal plasticity, once this is complete. It is therefore desirable to present the experimental results in ‘plasticity’ and creep, i.e., of strain controlled and stress controlled tests, in analogous forms. This is easily achieved by using $\epsilon$ as structure evolution parameter in both cases. The well known stress-strain curve for deformation at constant rate is thus supplemented by the strain rate-strain curve for creep. Due to the large variation of creep rates it is advisable to display the $\dot{\epsilon}-\epsilon$-curves in semi-logarithmic form. Even before Blum and Ilschner (1967) introduced the log $\dot{\epsilon}$-$\epsilon$-diagram as a standard for displaying creep results (Ilschner, 1973), $\dot{\epsilon}$-$\epsilon$-curves were used by Sherby et al. (1956) to show the difference between normal and inverse transient creep behavior of metals leading to a distinction between two classes of materials (pure metal and alloy class (Garofalo, 1965; Sherby and Burke, 1967)). There is no problem of determining the creep time by error-tolerant integration as $t = \int \dot{\epsilon}^{-1} d\epsilon$.

Using $\epsilon$ as structure evolution parameter is most helpful to check the significance of minimum creep rates. Omitting this simple check may lead to erroneous invention of new mechanisms of creep (Woodford, 1969), like Harper-Dorn creep (see Frost and Ashby (1982)), which tends to fade away under suitable testing conditions (Blum and Maier, 1999; McKnee et al., 2001). The semi-logarithmic $\dot{\epsilon}$-$\epsilon$-plot is also helpful in controlling the validity of the Monkman-Grant law constituting an inverse relationship between $\dot{\epsilon}_{\text{min}}$ and time to fracture; it is easy to show that the Monkman-Grant law holds exactly, if the log $\dot{\epsilon}$-$\epsilon$-curves for different stresses are exactly congruent.

Having recognized that simulating creep means simulating plasticity, the main aim of the present work is to present contemporary plasticity models and apply them to creep.

### 31.2 Empirical Relations

Many authors have been aware that description of creep needs more than a single ‘creep rate’. The modified Garofalo (1965)–equation treats $\dot{\epsilon}$ as sum of primary, secondary and tertiary creep contributions (Kloos, Granacher, and Monsees, 1998):

$$\dot{\epsilon} = \dot{\epsilon}_{\text{primary}} + \dot{\epsilon}_{\text{secondary}} + \dot{\epsilon}_{\text{tertiary}}$$  \hspace{1cm} (31.1)
with $\epsilon_{\text{primary}} = \epsilon_{1\text{max}} \left(1 - \exp\left(-D(t/t_{12})^u\right)\right)$, and $\epsilon_{\text{tertiary}} = C_{23}(t/t_{23})^f$; for completeness, one needs the plastic strain $\epsilon_{\text{load}}$ at the end of loading at $t = 0$. $D$, $u$, $C_{23}$ and $f$ are constants; $t_{12}$, $t_{23}$, $\epsilon_{\text{load}}$, $\epsilon_{1\text{max}}$ and $\dot{\epsilon}_{\text{secondary}} \equiv \dot{\epsilon}_{\text{min}}$ depend on $\sigma$ and $T$. Strain $\epsilon$ is implicit in the expressions for $\dot{\epsilon}_{\text{primary}}$ and $\dot{\epsilon}_{\text{tertiary}}$, as $t_{12}$ and $t_{23}$ are related to $\dot{\epsilon}_{\text{min}}$. Wilshire and coworkers (Evans et al., 1982; Wilshire and Evans, 1985) have drawn attention to the fact that in many technical materials an extended steady state of creep with constant $\dot{\epsilon}$ does not exist, but that the creep rate minimum results from superposition of primary and tertiary creep. Consequently, their $\theta$-projection concept simplifies (31.1) by omitting the secondary creep term:

$$\dot{\epsilon} = \Theta_1 \Theta_2 \exp(-\Theta_2 t) + \Theta_3 \Theta_4 \exp(\Theta_4 t).$$

(31.2)

The $\sigma$- and $T$-dependences of the parameters $\Theta_i$ allow for limited extrapolation of measured creep data. Obviously, Equations 31.1 and 31.2 may be integrated to yield an analytic expression for $\epsilon(t)$ which is useful for simple implementation of creep to finite element codes. A problem arises when $\sigma$ is not constant. As time $t$ does not monitor structure evolution, $\dot{\epsilon}$ must not be calculated as function of $t$ when $\sigma$ changes. Engineering practice has shown that relatively satisfactory results are obtained in the case of variable $\sigma$, if $\dot{\epsilon}$ is treated as function of $\epsilon$ (Granacher et al., 1986; Blum and Granacher, 1991). This underscores the important role of $\epsilon$ in dislocation structure evolution.

### 31.3 Basic Dislocation Processes

#### 31.3.1 Homogeneous Glide Activity

Identification of the basic dislocation processes is an essential prerequisite for modelling crystal plasticity. In each event of crystallographic glide two parts of the crystal above and below a certain area slip and thus become dislocated relative to each other by the Burgers vector $b$. The slipped areas are invisible because crystallographic order is restored after slipping. However, their boundary lines, the dislocation lines, remain visible, because crystallographic order is disturbed here and the crystal lattice is strained. It is important to keep in mind that, by definition, each dislocation segment is part of a closed loop bounding a certain slipped area. The slipped areas tend to expand under suitable stress. Yet their expansion is limited as they interact with other slipped areas through dislocation reactions. After passing a certain average number of dislocations, slipped areas stop expanding and the bounding dislocations are deposited in the crystal (see Nes (1998) for a probabilistic description).

There is only one way to diminish the dislocation content inside a (large) crystal. That is merging of slipped areas of the same $b$ (Weertman, 1960, 1968; Blum, 2002; Eisenlohr, 2004). Along the trace of merging the boundary lines of the slipped areas disappear, i.e. the dislocation length is reduced. The process of merging consists of two steps. The first step is formation of a stable dislocation dipole where the two slipped areas meet. The second step is annihilation of this dipole, requiring motion of the dislocations perpendicular to their slip planes, i.e. climb of edge and cross slip of screw dislocations.

One may object that this description is oversimplified, because it focuses on one slip system only. However, this is not the case. Even in the multiple slip situation where several slip
systems are active each dislocation line segment is part of a closed loop bounding a single slipped area. Thus in a gedanken experiment one can make all dislocations but those of a given Burgers vector disappear. Then only the dislocations with this Burgers vector will remain visible and it is only those which can produce annihilation (or recombination), unless reaction between more than two dislocations are taken into account.

### 31.3.2 Heterogeneous Glide Activity

In crystals deforming by homogeneous activity of crystallographic glide on each given system, dipole formation and annihilation is the only way to free the crystal from dislocation stress fields. However, if glide occurs in a heterogeneous, but elastically compatible fashion, another process comes into play, namely subgrain formation. Subgrains are parts of a grain which are rotated relative to each other by low angles of misorientation $\theta$. If homogeneous strain of a crystal is achieved by homogeneous glide activity, it is evident that no subgrains will form. However, a given homogeneous strain may also be achieved by heterogeneous glide activity followed by rotation. This can be seen from a simple example (Figure 31.2). Consider a two-dimensional crystal. Split the crystal into squares. Then deform the squares by shear on one of two possible slip systems by the angle $\theta$. Thereby the squares will become rhombs. Make the longitudinal axes of the rhombs parallel by rotating them by $\pm \theta/2$. For small $\theta$ the relative difference of the two sides of the rhombs, $1/\cos \theta - 1$, is negligible. Thus the rhombs can be approximated as diamonds. The result of heterogeneous shear on two different slip systems followed by rotation is homogeneous stretching. However, during this operation the crystal has split into diamond-shaped subgrains rotated relative to each other by $\theta$. This example (so-called checkerboard structure) shows that there are more than one possibilities to produce virtually the same change of shape of subgrains within a homogeneously deforming grain by equivalent combinations of crystallographic glide and accompanying rotation, ideally without

![Figure 31.2: Squares in 2D crystal (left), sheared by angle $\theta$ into rhombs by glide on two different slip systems (center), fit after rotation by $\pm \theta/2$. Remaining dislocations are geometrical in nature (subgrain boundary dislocations).](image)
generation of long-range stresses. The dislocations at the borders of subgrains are redundant, i.e., they do not find annihilation partners nearby. At the same time they have gotten rid of their long-range stresses. Such dislocations are termed geometrically necessary. In the present case they are grain boundary dislocations making up low-angle grain boundaries. Long-range stresses would result, if these dislocations would be removed from the crystal.

Subgrain formation is a general phenomenon in plasticity. It is traditionally introduced as a process of recovery where dislocations move into energetically favorable positions. However, this is misleading. As is illustrated in the example of the checkerboard structure, the misorientations result from heterogeneous, compatible glide and not from recovery after glide. Subgrain formation may be statistical or deterministic in nature. The statistical origin has been treated by Pantleon (1998). Kratochvíl (1990) and Kratochvíl and Orlová (1990) have given deterministic reasons. The detailed stability analysis of Sedláček (1999), see also Sedláček et al. (2001, 2002), confirmed that subgrains may form deterministically due to instability of deformation; the dislocations behave like sand on a beach in forming rather regular patterns. The lower rate of work hardening within a subgrain deforming on a reduced number of slip systems or the larger climb contribution in such subgrain lead to easier deformation of the crystal. However, the subgrain boundaries left behind form obstacles to further dislocation motion.

Quantitative analysis showed that migration of subgrain boundaries makes a relative contribution of 5 to 15% to inelastic strain in steady state (Biberger and Blum, 1992). In other words, about 90% of strain occur by homogeneous glide activity and only the rest by heterogeneous glide activity carried by geometrical dislocations. The geometrical dislocations are assembled mostly in subgrain boundaries. The density (length per crystal volume) of subgrain boundary dislocations is \( \rho_b \approx \frac{4\theta}{(wb)} \) where \( w \) is the subgrain size, so that \( 2/w \) is the subgrain boundary area per volume, and \( \theta/b \) is the length per boundary area of one family of parallel dislocations, with the boundary typically consisting of two families. The remaining density \( \rho_f \) of so-called free dislocations in the subgrain interior is the sum of some geometrical dislocations and the major rest carrying homogeneous glide. Data for Al-alloys (Blum, 1991, 1996) show that \( \rho_f \approx \frac{(\sigma/(bG))^2}{\rho} \) at stresses in the order of \( 10^{-4}G \). With the subgrain size varying as \( w \approx 25bG/\sigma \) (Raj and Pharr, 1986; Blum, 1993; Straub and Blum, 1996) the ratio \( \rho_f/\rho_b \approx 0.16\theta G/\sigma \) is about 28 for \( \sigma = 10^{-4}G \) and \( \theta = 1^\circ \). As the strain rate is proportional to the product of density and velocity of dislocations, it follows that the free dislocations move much faster than the boundaries. Combining this result with the fact (following from the stage II work hardening rate), that the slip distance of free dislocations is several times the subgrain size \( w \) (Blum, 1993; Nes, 1998), makes clear that the subgrain boundaries must be permeable to free dislocations during deformation.

31.4 Models

As mentioned before, models of plasticity should in principle also describe creep. Sedláček (2004) has reviewed existing models using one or more parameters to characterize the dislocation structure. The models are similar in describing dislocation generation in terms of a geometrical parameter termed mean free path (see Equation (21.10)). A main difference lies in the treatment of dynamic recovery. It follows directly from Section 31.3.1 that one-parameter
models cannot give a full description of dislocation recovery simply because the density of dislocation dipoles is lacking. Multi-parameter models (Sedláček, 2004) may be appropriate to cover also creep, but this has not yet been verified in detail. In the following we describe two modelling approaches to homogeneous and heterogeneous glide which have been kept as simple as possible in order to provide easy physical insight.

### 31.4.1 Two-parameter Model for Homogeneous Glide

The two dislocation-structure parameters are the densities $\rho_{\text{single}}$ of single dislocations and $\rho_{\text{dip}}$ of dislocations which are in the configuration of stable dipoles and are thus ready to annihilate. The model takes up ideas presented before by Nes (1998) and Roters, Raabe, and Gottstein (2000), without going into the details of storage of immobilized reaction products. Its intention is to see which features of deformation can be explained at this level of simplicity.

The flow diagram 31.3 illustrates the history of dislocations from generation to annihilation. Single dislocations may form stable dipoles or be spontaneously annihilated upon meeting. Dipole dislocations also may be spontaneously annihilated upon meeting singles or alternatively dissolve. The mathematical formulation of these structure evolution processes on the basis of average quantities with differential equations of type (21.10) is rather straightforward (Eisenlohr, 2004; Blum, Eisenlohr, and Breutinger, 2002). The model must be completed by kinetic laws for glide and for dipole annihilation. Glide kinetics (velocity of gliding dislocations) was described using an effective stress approach and different expressions for pure materials and solid solutions. For pure metals thermally activated overcoming of forest dislocations was taken as glide-controlling process. For solid solutions two different laws apply depending on whether viscous motion with a cloud of solutes or jerky motion by thermally activated overcoming of groups of solute atoms is easier. Annihilation kinetics was described in terms of climb of edge dislocations.

The results of modelling the steady state creep behavior of pure Aluminum are shown in Figure 31.4. The steady state creep is normalized in the usual way essentially by the coefficient of self-diffusion; the stress is normalized by the shear modulus $G$ (for better comparison with other materials). The model curve fits the data nicely. It should be noted that the four-power law at low stresses was obtained by assuming (Argon and Moffatt, 1981; Blum and Eisenlohr, 2002) that vacancy emission/absorption in climb of edge dislocations is restricted to jogs whose spacing scales with the average dislocation spacing. Otherwise, a three power law would result. The breakdown of the power law at high stresses is due to spontaneous

![Figure 31.3](image) Dislocations are generated as single dislocations (density $\rho_{\text{sgl}}$). The content of singles is reduced by spontaneous annihilation (left exit) and by formation of stable dipoles (right exit). The density $\rho_{\text{dip}}$ of dipoles is reduced by spontaneous annihilation due to reaction with singles (left exit) and by dipole annihilation (right exit).
Figure 31.4: Steady state creep behavior of pure Aluminum; left: kinetics, right: structure. After Blum and Eisenlohr (2002); see also Blum et al. (2002). Symbols from experiment, lines from model. Coefficient of self diffusion $D$ and $G$ from Frost and Ashby (1982), $k_B$: Boltzmann constant.

... annihilation. This is consistent with the view of a basic difference between low temperature and high temperature creep in the sense that at high temperatures creep is controlled by an independent process of dynamic recovery (like climb), while at low temperatures steady state deformation (identical to flow stress saturation) is glide controlled without an independent recovery process: Once dislocation annihilation is completely due to spontaneous annihilation, recovery in itself becomes glide controlled. The average spacings of single and dipole dislocations lie in the right order of magnitude. The approximately inverse proportionality of spacings to normalized stress results from the fact that thermally activated glide is relatively easy and sensitively dependent on effective stress so that its rate easily adjusts to the rate of dynamic recovery through minor variations of effective stress. It should be noted, though, that the absolute magnitude of the stress exponent of the spacings is in fact slightly less than 1, as necessary to explain the distinct increase of the steady state creep rate with stress. The deviation from inverse proportionality at large stresses $> 10^{-3}G$ results from spontaneous annihilation. While the results for steady state creep are satisfactory, the results for creep transients are not (Blum et al., 2002). We will come back to that in the next section.

The Titanium alloy TiAl6V4 has been modelled by exchanging the velocity law for glide.

---

1 The validity of this explanation has been questioned recently by Nes et al. (2002).
Figure 31.5: Steady state creep behavior of TiAl6V4; left: kinetics, right: structure. $m$: effective stress exponent of dislocation velocity. From Blum, Eisenlohr, and Breutinger (2002).

In addition, the climb law differed somewhat (no jog control) and the parameters of dislocation generation were adjusted differently from pure Al (see Blum et al. (2002) for details). The major features to be seen in Figure 31.5 for TiAl6V4 are the relatively sharp kink in the kinetic relation at $\sigma \approx 2 \cdot 10^{-2}G$ and the change in the densities of dipoles compared to Al. The modelled kink, which fits the experimental result of a distinct increase of stress exponent of steady state creep rate from 3 to > 30, is associated with the change in mode of dislocation motion as jerky glide (without cloud of solutes) becomes easier than viscous glide (with cloud exerting solute drag) with increasing stress. This interpretation is backed by observation of dynamic strain aging effects Breutinger (2004). The lower dipole dislocation densities are supposed to be mainly due to the deceleration of glide by solutes giving dipoles relatively longer times to annihilate. These modelling results are consistent with the classical notion of solid solution hardened metals (alloy class of materials) being glide controlled. The damping of the transient reactions by solute hardening also explains why not only the steady state, but also the transient creep of an Al-Mg alloy was approximated by the model (Blum et al., 2002).

31.4.2 Composite Model for Heterogeneous Glide

The failure of the model of Section 31.4.1 in reproducing the transient creep reaction of pure Al shows that two dislocation parameters are not enough. The same conclusion has been
reached before by others, e.g. Nes (1998); Roters et al. (2000). The two-parameter model predicts that dislocation densities attain their steady state values rather abruptly within a relatively short strain interval. The experimental results, on the other hand, display a relatively long period of transient creep until a pure material with low initial dislocation density reaches its steady state. The major structural feature in this period is the build-up of a subgrain structure. The geometrical dislocations of the subgrain boundaries superimpose to the density of free dislocations which remains relatively constant during most of the transient at constant stress or even decreases (Garofalo, 1965; Blum et al., 1980; Blum, 1991). Thus the increasing heterogeneity of glide generating an increasing amount of geometrical dislocations and related misorientations may well be the reason for the gradual hardening in primary creep. (Alternatively, changes of the fraction of immobile dislocations in the subgrain interior must be invoked to explain the hardening in the primary stage of creep.) Blum (1991, 1993) has associated the late stages of hardening in primary creep with the increase of the volume fraction $f_{\text{sub}}$ of material where the subgrain structure was developed with the steady state spacings of $\rho_{\text{f}}$ and $w$. The increase of $f_{\text{sub}}$ corresponds to a decrease in the volume average of the subgrain size. However, new results suggest that hardening may continue even beyond the point where the spacings of free dislocations and subgrain boundaries have saturated.

In fact, there is a number of reasons why the misoriented structure of subgrains resists deformation and generates back stresses in the subgrain interior. Without going into details, the composite model of plastic deformation accounts for them by treating the subgrain boundaries as hard regions compared to the subgrain interior. The composite framework has been reviewed by Sedláček (2004). Care must be taken in the correct definition of the volume fraction of hard regions. Due to the interaction between boundaries and interior, it is not only the subgrain boundaries themselves where an internal forward stress is acting, but also their vicinity where dislocations have bowed out (Mughrabi, 1987; Sedláček et al., 1994). This removes the apparent problem of vanishing thickness of subgrain boundaries (which ideally are planar grain boundaries) as opposed to thick cell walls.

Nes and coworkers (Nes, 1998; Nes and Marthinsen, 2002) do not use the composite framework in full; they rather restrict attention to the effect which the hard subgrain boundary regions exert on the soft subgrain interior by generating a back stress. This back stress is formulated as an Orowan bowing stress and a justification for this approach has been presented (Nes and Holmedal, 2002). While a bowing stress is consistent with the composite approach (Sedláček and Forest, 2000), the flexibility of build up in the course of deformation of this back stress up to specific levels, which are not simply related to subgrain size, is missing. Nevertheless the Nes model is quite successful in simulating transient creep reactions (Nes and Holmedal, 2002).

Geometrical dislocations do not only act in pure materials, but also in many materials of great practical importance which have undergone thermomechanical processing prior to application, e.g. by forming operations. An example are tempered martensite 9-12%Cr-steels. In this group of alloys which is indispensable for electric power stations where materials have to withstand high temperature, high pressure and corrosive attack for more than three decades, the geometrical dislocations are generated in the course of internal deformation accompanying the martensitic transformation. Subsequent tempering then transforms them into subgrains with sizes in the range of several 100 nm. Precipitation of different hard phases enhances the resistance of the boundaries against migration under stress. In the initial state the subgrains
are smaller than they would be in the steady state at the stresses of application ranging in the order of 100 MPa and below. Therefore, in contrast to dislocation-poor materials, the geometrical dislocation density will decrease during service rather than increase. This is exactly what is observed. The direct consequence of this observation is that the minimum creep rate, occurring at $\epsilon \approx 0.01$ is far from the steady state of creep which is observed only at $\epsilon \approx 0.2$. Analogous consequences hold in other cases of predeformed materials.

The composite model has been applied by Polcik (1999) and Henes et al. (1999) to creep of tempered martensite Cr-steels. Dislocation structure evolution was simulated in a phenomenological way in terms of exponentially decaying deviation from steady state which was quantified by empirical laws based on transmission electron microscopical observations. The kinetic law of the soft regions, (21.20), was formulated with simple proportionality between dislocation velocity and effective stress, consistent with solute drag. The kinetic law (21.21) for the hard region was fitted to match the stress dependence of the minimum creep rate. In the course of creep the volume fraction of hard region decreases continually as the subgrains coarsen while the width of the hard regions, assumed to be stress controlled, remains constant. Qualitatively, one expects a softening of the material during creep. Such softening is in fact found. It is conventionally called tertiary creep and is interpreted as damage accumulation in the course of fracturing in tension. This interpretation is clearly wrong, because the softening occurs also in compression without fracture (Straub et al., 1993). The coarsening of precipitates is not expected to be the dominant cause for softening, because softening is observed also under conditions where the (time-dependent, Götz et al. (2001)) coarsening of precipitates is negligible.

Figure 31.6 shows an experiment yielding rather direct evidence for subgrain strengthening. The $\dot{\epsilon} - \epsilon$-curve of a 10%Cr-steel in the initial state after tempering displays a pronounced creep rate minimum. The subgrain size is 0.41 $\mu$m in this state. Within 2317 h of strain controlled cyclic deformation at 873 K the subgrain size increases to 1.5 $\mu$m. Within this time negligible coarsening of precipitates is expected to occur; rather, fresh precipitation of W-based Laves phase is expected. The coarsened subgrain size is just slightly smaller than the steady state value for the testing conditions of Figure 31.6. From that one expects direct transition to steady state with only little softening. In fact, Figure 31.6 shows that the relative creep rate minimum has nearly disappeared; $\dot{\epsilon}_{\text{min}}$ has increased by a factor of about 20. Creep life has diminished accordingly. The two dashed lines calculated from the composite model are in semi-quantitative agreement with the experimental result. This result appears rather conclusive with regard to subgrain hardening. An alternative explanation in terms of differences in immobile fraction of dislocations inside subgrains does not appear sensible.

While the composite model in the form described above works, the empirical character of dislocation structure evolution used in the model is not satisfactory and needs to be exchanged by evolution laws founded on clear mechanisms of storage of geometrical dislocations and their annihilation by subgrain boundary migration (Nes, 1998; Nes and Marthinsen, 2002).

### 31.5 Concluding Remarks

Simulation of creep is possible with suitable models of plasticity which cover the range of elevated temperatures and relatively low stresses. In order to model recovery of dislocations,
constituting a dominant process in creep, in the case of homogeneous glide activity, the density of stable, annihilating dislocation dipoles needs to be explicitly considered. A simple, one-dimensional, statistical two-parameter model keeping track of single dislocations and dipoles is able to reproduce essential features of steady state creep including the influence of solutes. In alloys, two modes of dislocation glide, with and without cloud of solutes, need to be discerned, consistent with the experimental data on steady state deformation. In addition, hardening by foreign phases, e.g. particle hardening, which is a dominant mechanism of increasing the resistance against creep as well as stress relaxation in technical materials, must be introduced into the model; this is usually done in a semi-empirical fashion with a temperature-dependent particle hardening term to be subtracted from the stress.

Consideration of homogeneous glide is not sufficient, however, as the geometrical dislocations (forming subgrain boundaries), which result from heterogeneous, compatible slip activity even in homogeneous deformation, lead to significant hardening not only in pure materials, but also in many technical alloys. The composite model provides a means to model subgrain hardening.

The neglection of subgrain hardening impairs the success of the two-parameter-model in describing steady state deformation. The complete description of creep and crystal plasticity in general must cover homogeneous dislocation activity (with formation and annihilation of
dislocation dipoles) as well as heterogeneous dislocation activity (with formation and recombination of subgrain boundaries).

**Acknowledgements**

Fruitful discussions with Dipl.-Ing. P. Eisenlohr and Dr.-Ing. R. Sedláček, financial support from several sources, in particular the Deutsche Forschungsgemeinschaft and Schott Lithotec AG, and aid of X. Zeng in extending the model curves of Figure 31.4 are gratefully acknowledged.

**References**


References


References


Computational Fracture Mechanics

Wolfgang Brocks

32.1 Introductory Remarks on Inelastic Material Behaviour

Nonlinear fracture mechanics can be understood as mechanics of fracture for materials with inelastic stress-strain relations, where inelastic behaviour is any kind of irreversible response to thermo-mechanical loading including phenomena like temperature and rate dependence, time dependence, creep and relaxation, Bauschinger effect, cyclic hardening and softening, effects of multi-axial and non-proportional loading. The respective constitutive models include evolution equations for internal variables, which describe material hardening, so that the material behaviour is “stable” (Drucker 1964). The corresponding boundary value problem is elliptic, and FE simulations will yield a unique solution which is convergent with refining the mesh. The physical reality, however, is more complex: materials “soften” due to formation of micro-cracks, initiation, growth and coalescence of voids etc., generally summarised as “damage”. This damage may lead to the initiation and growth of macro-cracks in a structure and to final failure in the end. The “crack tip” as addressed in fracture mechanics is a mathematical idealisation. In reality, a region of material degradation exists in some process zone ahead of a macro-crack, where finally new surfaces will be created. In this process zone, the micro-behaviour becomes important for constitutive modeling, and this will raise questions on materials length scales (Sun and Hönig 1994, Siegmund and Brocks 1998b) which cannot be answered within the limitations of the theory of “simple materials”. The boundary value problem for softening materials may loose ellipticity, and FE simulations will yield mesh dependent results as the element size affects the separation energy (Siegmund and Brocks 1998a).

Three kinds of approaches exist to model damage, material separation and fracture phenomena:

1. No damage evolution is modeled and conventional material models, e.g. elastic-plastic or viscoplastic constitutive equations, are applied, the process zone is assumed as infinitesimally small, and special fracture criteria, e.g. based on $K$, $J$, $C^*$, for crack extension are required;

2. Separation of surfaces is admitted, if some critical value is reached locally, whereas the material outside behaves conventional, the process zone is some surface region, and the fracture criterion is a cohesive law;
3. Softening behaviour is introduced into the constitutive model, e.g. as evolution of damage described by additional internal variables, the process zone is a volume, and the evolution equation for damage yields a fracture criterion.

Classical elastic-plastic fracture mechanics (EPFM) covers a comparably small part of these constitutive theories and phenomena of inelastic deformation. It has been established under the assumptions of the theory of finite plasticity or “deformation theory” of plasticity by Hencky 1924 and the kinematics of small deformations. It has nevertheless become the most important field of fracture research beside linear-elastic fracture mechanics (LEFM) as it allowed for analytical descriptions and solutions and has been successfully applied to describe crack growth initiation in ductile materials like metals at low and moderate temperatures. The plastic deformation behaviour of metals, however, is more realistically described by the theory of incremental plasticity by von Mises, Prandtl and Reuss, which accounts for effects of load history, unloading, and local rearrangement of stresses. As no analytical solutions are possible in this case, numerical methods have won great importance. And with the increasing capacities of computers, the constitutive relations used in structural analyses have become more advanced and sophisticated, including more and more of the above mentioned phenomena. But fracture parameters, and criteria for fracture and crack growth, which are used in practice for engineering assessment methods are still the same as in the early times of EPFM, namely

- The $J$-integral of Cherepanov 1967 and Rice 1968 or its analogon, $C^*$, for creep crack growth (Landes and Begley 1976), and
- Crack tip opening displacement (CTOD), $\delta$, see Burdekin and Stone 1966, Dawes 1985.

Numerical analyses and simulations applying incremental plasticity or more advanced constitutive theories can be used in this context for

- Determination of classical fracture parameters for complex configurations and boundary conditions including thermo-mechanical loading, if no appropriate analytical solutions are available;
- Investigations of the applicability and the limitations of these parameters and of engineering assessment methods in general;
- Application of local criteria of fracture, like those by Beremin 1981, 1983, if the classical parameters fail in giving reliable predictions.

The present contribution will essentially restrict to the application of the von Mises theory of incremental plasticity to cracked specimens and components with particular emphasis on the classical parameter of EPFM, $J$. The numerical modeling of decohesion and separation phenomena by “cohesive elements” will follow. Damage models like those of Gurson 1977 and Rousselier 1987, which have found increasing application, recently, will not be addressed here but in Chapter 45.
32.2 FE Meshes for Structures with Crack-Like Defects

32.2.1 General Aspects and Examples

Cracks and crack-like defects induce high stress and strain gradients which require a fine discretisation resulting in large numbers of elements and degrees of freedom. Nonlinear simulations of components with stress concentrators are therefore expensive with respect to computation time and memory. All possibilities to reduce the number of degrees of freedoms should hence be utilised like

- restricting to two-dimensional models of the structure if physically meaningful,
- coarsening the mesh remote from the defect,
- introducing symmetry conditions,
- introducing symmetry conditions,
- applying singular elements with special shape functions.

Modelling does always mean reduction of complexity and simplification. Models should be as simple as possible and only as complex as unavoidable to cover the interesting effects. For plane specimen geometries the possibility of two-dimensional models should always be considered, at least for pre-analyses of a new problem. Thin specimens and sheet metal components are commonly adequately represented by a plane-stress model, and thick or side-grooved specimens by a plane strain one.

![FE meshes at the crack tip and in the ligament](image)

**Figure 32.1:** FE meshes at the crack tip and in the ligament with (a) collapsed elements for a stationary crack and (b) a regular arrangement of elements for crack-extension simulations.

Figure 32.1 shows the two-dimensional FE meshes in the vicinity of the crack tip of a fracture mechanics specimen. As loading and geometry are symmetric for a mode I configuration, just the upper half is modelled and normal displacements are constrained in the ligament. Collapsed elements (see Section 32.2.2) are applied for the analysis of a stationary crack (Figure 32.1a) and a regular element arrangement for the analysis of an extending crack (Figure 32.1b), respectively. The meshes demonstrate the common strategy used for coarsening the mesh away from the crack. Three-dimensional analyses are necessary, of course, if the geometry is not plane or if 3D effects through the thickness and along the crack front
are studied. Numerous examples of three-dimensional models of structures with surface flaws and material gradients can be found in the literature, e.g. Brocks et al. 1989, 1993, Schmitt et al. 1997.

32.2.2 Singular Elements for Stationary Cracks

Singular elements have been developed for numerical analyses of fracture problems to increase the accuracy of stress calculations and $K$-factors at a time when computer capacities were still rather limited. Barsoum 1977 found that triangular or prismatic isoparametric elements which were produced by collapsing one side or plane and shifting the respective mid-side nodes to a quarter position (double distorted elements), included the $1/\sqrt{r}$-singularity of strains in LEFM as well as the $1/r$-singularity of EPFM for perfectly plastic material. The strain energy however, remains finite for $r \rightarrow 0$ in linear elasticity as well as for HRR-like fields (Hutchinson 1986a,b, Rice and Rosengreen 1968), because the strain energy density has a singularity of the order of $r^{-1}$ in both cases. This is an important attribute for the physical significance of the $J$-integral, see Section 32.3.1.

Numerical studies, see e.g. McMeeking and Rice 1975, Brocks et al. 1985, have shown that triangular or prismatic collapsed 8- or 20-node elements, respectively, with a $1/r$-singularity are well suited for elastic-plastic calculations. Together with a large-strain analysis, commonly performed by an updated Lagrangean formulation (Gadala et al. 1980), crack tip blunting can be simulated, see Figure 32.2a, and principal stresses show their typical shape exhibiting a maximum ahead of the crack tip (Rice and Johnson, 1970), see Figure 32.3.

(a)

(b)

Figure 32.2: Deformed FE meshes at the crack tip showing (a) blunting of a stationary crack and (b) crack extension.

Singular elements have become less important in recent years, mainly because

- $J$ integral calculations by the virtual crack extension method (see Section 32.3.2) yield reliable and accurate results even for rather coarse meshes,

- singular elements can not be applied for crack growth simulations which require a regular arrangement of elements in the ligament as shown in Figure 32.1b; these meshes, however, will not provide sufficiently accurate results of CTOD or stresses at the crack tip for stationary cracks.
### 32.2.3 Regular Element Arrangements for Extending Cracks

If a critical initiation value of some fracture parameter is exceeded, a crack starts to grow. Different from crack growth in elastic materials, where crack initiation always leads to catastrophic failure of the structure, ductile tearing may occur in a stable manner, i.e. under still growing external forces, or at least deformation controlled even beyond maximum load. Crack growth can be simulated by

- node release techniques, controlled by any fracture mechanics parameter as $J$, CTOD, CTOA (see Section 32.3), e.g. Siegele and Schmitt 1983, Brocks et al. 1994, Brocks and Yuan 1989, Gullerud et al. 1999.

- cohesive elements (see Section 32.4), e.g. Needleman 1990, Yuan et al. 1996, Lin et al. 1997, Siegmund et al. 1998,


Figure 32.2b illustrates crack growth by node release which is controlled by a criterion assuming constant crack tip opening angle (CTOA). The simulation has to be performed under prescribed displacements in order to proceed beyond maximum load.
32.3 The $J$-Integral as Characteristic Parameter in Elasto-Plastic Fracture Mechanics

32.3.1 Foundation

Path-independent integrals are used in physics to calculate the intensity of a singularity of a field quantity without knowing the exact shape of this field in the vicinity of the singularity. They are derived from conservation laws. They have been introduced into fracture mechanics by Cherepanov 1967 and Rice 1968. Budiansky and Rice 1973 also showed that this “$J$-integral” is identical with the energy release rate for a plane crack extension. The $J$-integral used in fracture mechanics

$$J = \oint_{\Gamma} [w \, dx_2 - \sigma_{jk} n_k u_{j,1} \, ds], \tag{32.1}$$

where $w$ is the strain energy density and the integration contour runs anti-clockwise around the crack tip, see Figure 32.4a, can be deduced from the equations governing the elasto-static boundary value problem for a material body. Because of its path independence, it can be calculated in the remote field and characterises also the near tip situation, which establishes its role as a fracture parameter. But note that the path independence does only hold if the conditions

1. time independent processes, no body forces, $\sigma_{ij,j}$,
2. small strains, $\varepsilon_{ij} = \frac{1}{2} (u_{i,j} + u_{j,i})$,
3. homogeneous hyper-elastic material, $\sigma_{ij} = \frac{\partial w}{\partial \varepsilon_{ij}}$,
4. plane stress and displacement fields, i.e. no dependence on $x_3$,
5. straight and stress-free crack borders parallel to $x_1$.

are met. Extensions of the $J$-integral to 3D configurations, body forces, surface tractions, thermal loading and multi-phase materials will not be discussed here, see Brocks and Scheider 2003. Analogous considerations as made for the derivation of the $J$-integral, yields the $C^*$-integral for visco-plastic material behaviour (Landes and Begley 1976) if a power (work rate) density, $w Y$, exists so that stresses derive from $\sigma_{ij} = \frac{\partial w}{\partial \varepsilon_{ij}}$. This analogy implies that $C^*$ is path independent under the same conditions which hold for the path independence of $J$.

Beside its identification as energy release rate in hyper-elastic materials, $J$ also plays the role of an intensity factor like $K$ in the case of linear elastic materials. Hutchinson 1968a,b and Rice and Rosengreen 1968 derived the singular stress and strain fields at a crack tip in a power law hardening material, the since called HRR-field. Figure 32.3 shows that the principal stresses for different $J$-values collapse into one $J$-dominated “master curve”, if the abscissa in normalised by $J/\sigma_Y$. Of course, the singular HRR-field cannot describe the stresses in the region of large strains at the blunted crack tip correctly (Rice and Johnson 1970, McMeeking and Rice 1975, Brocks and Olschewski 1986). Again, an analogy to the $J$-dominated singularity in EPFM holds for the $C^*$-integral in visco-plastic materials: stress and strain fields have an HRR-like singularity if secondary creep follows a power law (Riedel and Rice 1980).
32.3.2 The Domain Integral or VCE Method

Calculating a contour integral like Equation (32.1) is quite unfavorable in finite element codes as coordinates and displacements refer to nodal points and stresses and strains to Gaussian integration points. Stress fields are generally discontinuous over element boundaries and extrapolation of stresses to nodes requires additional assumptions. Hence, a domain integral method is commonly used to evaluate contour integrals, see e.g. ABAQUS 2000.

Applying the divergence theorem, the contour integral can be re-formulated as an area integral in two dimensions or a volume integral in three dimensions over a finite domain surrounding the crack front. The method is quite robust in the sense that accurate values are obtained even with quite coarse meshes, because the integral is taken over a domain of elements, so that errors in local solution parameters have less effect. This method was first suggested by Parks 1974, 1977 and further worked out by DeLorenzi 1982a, b.

The $J$-integral is defined in terms of the energy release rate, associated with a fictitious small crack advance, $\Delta a$, see Figure 32.4b,

$$J = \frac{1}{\Delta A} \int_{B_0} [\sigma_{ij} u_{j,k} - w \delta_{ik}] \Delta x_{k,i} \, dS,$$

(32.2)

where $\Delta x_k$ is the shift of the crack front coordinates, $\Delta A$ the correspondent increase in crack area and the integration domain, $B_0$, is the gray area in Figure 32.4b. Because of this physical
32.3 The $J$-Integral as Characteristic Parameter in Elasto-Plastic Fracture Mechanics

32.3.3 Path Dependence of the $J$-Integral in Incremental Plasticity

The severest restriction for $J$ results from the assumed existence of a strain energy density, $w$, as a potential from which stresses can be uniquely derived. This assumption also conceals behind frequently used expressions like “deformation theory of plasticity” (Hutchinson 1968a,b, Rice and Rosengren 1968) or theory of “finite plasticity” (Hencky 1924). But it actually does not describe irreversible plastic deformations as in the “incremental theory of plasticity” of von Mises, Prandtl and Reuss, but “hyperelastic” or non-linear elastic behaviour. It does not only exclude any local unloading processes but also any local re-arranging of stresses, i.e. changing of loading direction in the stress space, resulting from the yield condition. All loading paths in the stress space are supposed to remain “radial” so that the ratios of principal stresses do not change with time. The condition of monotonous global loading of a structure is of course not sufficient to guarantee radial stress paths in non-homogeneous stress fields. Hence, the $J$-integral will become path dependent as soon as plasticity occurs and the contour $\Gamma$ passes the plastic zone (McMeeking 1977).

For small scale and contained yielding, a path independent integral can be computed outside the plastic zone. This means that $\Gamma$ – or the respective evaluation domain – has to be large enough to surround the plastic zone and pass through the elastic region only. In gross plasticity, this is not possible, and some more or less pronounced path-dependence will always occur, so that the evaluation of a “path-independent” integral is a question of numerical accuracy. Because of its relation to the global energy release rate, which is used to evaluate $J$ from fracture mechanics test results (ASTM E 1737), $J$ has to be understood as a “saturated” value reached in the “far-field” remote from the crack tip.

Significant stress re-arrangements occur at a blunting crack tip, see Figure 32.3 and the results by McMeeking and Rice 1975, Brock and Olschewski 1986, and the path dependence

![Figure 32.5: Path dependence of the elastic-plastic $J$ with increasing load-line displacement, $V_{LL}$, calculated for domains of increasing size, #02, ..., #21, in comparison to the $J$ value of ASTM 1737, (a) small-strain analysis and (b) large-strain analysis.](image-url)
increases strongly, see Figure 32.5b. Thus, a small strain analysis is advantageous if only the
$J$-integral and no stresses at the crack tip shall be calculated, as only little path dependence
occurs, see Figure 32.5a. But note that stresses at the crack tip lack physical significance in
this case.

As the work dissipated by plastic deformation always has to be positive, the calculated $J$
values have to increase monotonically with the size of the domain (Yuan and Brocks 1991),
which is confirmed by Figure 32.5b – except contour #21 touching the boundary. The highest
calculated $J$-value with increasing domain size is always the closest to the “real” far-field $J$.
Any different result would indicate an error in the definition of contours or in the evaluation
of $J$.

Moreover, $J$ will keep a finite value in the limit of a vanishingly small contour if and
only if the strain energy density, $w$, has a singularity of the order of $r^{-1}$. This holds in linear
elasticity where stresses and strains have a $1/\sqrt{r}$-singularity and for HRR-like fields. As the
stress singularity at the blunting crack tip vanishes under the assumptions of finite strains and
incremental plasticity, $J$ will not even have a finite value for $r \to 0$ any more. In particular,
the crack-tip $J$ vanishes for ductile crack extension (Brocks and Yuan 1989), which is the
so-called “paradoxon of EPFM” (Rice 1979), that no “sur-plus energy” is provided for crack
extension. This paradoxon, however, is due to the fact that no process zone, where material
damage precedes crack extension, is considered in EPFM.

32.4 The Cohesive Model

32.4.1 Fundamentals

The idea for the cohesive model is based on the consideration that infinite stresses at the crack
tip are not realistic. Models to overcome this drawback have been introduced by Dugdale
1960 and by Barenblatt 1962, who introduced a cohesive zone in the ligament of the crack.
The stresses in the cohesive zone follow a prescribed distribution $\sigma(x)$, which is specific
for a given material but independent of the global loading conditions. Recently developed
models differ from Barenblatt’s model in that they define the traction acting on the ligament
in dependence on the opening and not on the crack tip distance.

For practical applications the model became more interesting when numerical methods,
mostly the finite element method, were applicable to nonlinear problems. Needleman 1987
was the first, who used the model for crack propagation analyses of ductile materials. More
than ten years earlier Hillerborg et al. 1976 applied the cohesive zone model to brittle fracture
using the finite element method for the first time, followed already in the 80’s by Petersson
1981 and Carpinteri 1986 amongst others.

The material separation and thus damage of the structure is classically described by in-
terface elements – no continuum elements are damaged in the cohesive model. Using this
technique, the behaviour of the material is split in two parts, the damage-free continuum with
an arbitrary material law, and the cohesive interfaces between the continuum elements, which
specify only the damage of the material. A different technique to define the cohesive elements
as solid elements, which contain not only the damage of the structure but also its continuum
properties, is the so-called strong discontinuity approach (e.g. Simo et al. 1993; Larsson et al. 1996; Moës et al. 1999). This method is not considered here.

The interface elements open when damage occurs and lose their stiffness at failure so that the continuum elements are disconnected. For this reason the crack can propagate only along the element boundaries. If the crack propagation direction is not known in advance, the mesh generation has to allow for different crack paths.

![Cohesive element: definition of local coordinate system.](image)

Figure 32.6: Cohesive element: definition of local coordinate system.

The separation of the cohesive interfaces is calculated from the displacement jump \( \delta = [u] = u^+ - u^-, \) i.e. the difference of the displacements of the adjacent continuum elements. The separation vector, \( \delta \), is commonly represented in a local coordinate system, Figure 32.6, namely a normal separation, \( \delta_n \), for mode I and two tangential separations, \( \delta_{t1}, \delta_{t2} \). Only 2D problems with with in-plane tangential separation (mode II) will be considered in the following. When the normal or tangential component of the separation reaches a critical value, \( \delta_n^c \) or \( \delta_t^c \), respectively, the continuum elements initially connected by this cohesive element, are disconnected, which means that the material at this point has failed.

A stress (or “traction”) vector, \( \sigma \), with one normal and two tangential components acts at the surfaces of the two connected continuum elements, which depends on the separation \( \delta \). Beside the critical separation, \( \delta_n^c \) and \( \delta_t^c \), the maximum traction components, \( \sigma_{n}^c \) and \( \sigma_t^c \), are used as fracture parameters, also denoted as “cohesive strength”. The curve \( \sigma(\delta) \) is assumed to be a material independent cohesive law. Common to all cohesive laws is that

1. They contain two material parameters, \( \sigma_c \) and \( \delta_c \), per separation mode, and
2. The stresses after failure become zero, \( \sigma(\delta > \delta_c) \equiv 0 \), for both normal and tangential separation.

The integration of the traction over separation, either in normal or in tangential direction, gives the energy dissipated by the cohesive elements, \( \Gamma_c \).

\[
\Gamma_c = \int_0^{\delta_c} \sigma(\delta) \, d\delta, \tag{32.3}
\]

which is used alternatively to \( \delta_c \). The cohesive energy at failure \( \Gamma_c \) equals the \( J \) integral (Rice 1968), Equation (32.1), at initiation of ductile crack extension in mode I under the assumptions made for the validity of the HRR-field.
If tangential and normal separation occur simultaneously, one affects the other. This "mixed mode" case requires an additional hypothesis of the interaction of the failure modes (see Tvergaard and Hutchinson 1993, Scheider 2001b). Further assumptions have to be made for the unloading behaviour of the cohesive elements and the friction in a failed cohesive element under compressive normal stresses. Neither of these problems will be discussed here.

Since the cohesive model is a phenomenological model of the separation process, there is no evidence which form to take for the cohesive law, $\sigma(\delta)$. It has to be assumed independent of a specific material but dependent on the separation mechanism. For ductile materials a polynomial function of third order, first used by Needleman 1987 for pure normal separation, and some years later extended by Tvergaard 1990 for mixed mode loading, is one of the most popular cohesive laws and used by many authors, e.g. Chaboche et al. 1997, see Figure 32.7a. Needleman 1990 applies an exponential cohesive law, which is based on the energy function of atomic bonding as proposed by Rose et al. 1981, see Figure 32.7b. This model is also used by a number of authors, both for ductile (Siegmund and Brocks 1998a) and brittle metals (Xu and Needleman 1994). A more versatile cohesive law was proposed by Scheider 2001a for ductile failure, Figure 32.7c, meeting the requirements of (i) variable initial stiffness, (ii) region of constant traction, (iii) steadily differentiable curve. It is alike a multilinear cohesive law proposed by Tvergaard and Hutchinson 1992, Figure 32.7d, which is not steadily differentiable, however. The latter has also been used, for example, by Roy Chowdhury and Narasimhan 2000. A special case is obtained for $\delta_1 \to 0$, $\delta_2 \to \delta_c$, having a constant stress up
32.4 The Cohesive Model

To final failure of the element, see Yuan and Cornec 1991, Lin et al. 1998. Purely decreasing cohesive laws have been established for brittle materials like concrete and rocks. Hillerborg et al. 1976 proposed a linear decreasing function with infinite initial stiffness, Figure 32.7e, which is also used by Camacho and Ortiz 1996. A bilinear decreasing separation law as shown in Figure 32.7f is in use for concrete (Bazant 1993, Guinea et al. 1994).

The influence of the shape of the cohesive law on the crack propagation has not yet been studied systematically. Some authors report only minor influence (Yuan and Cornec 1991, Tvergaard and Hutchinson 1992), others have found significant effects (Brocks et al. 2003) on the simulated load vs displacement curves of fracture mechanics specimens.

32.4.2 Example: Simulation of Ductile Tearing in a Laser Weld

Cohesive laws allow for a simple, physically plausible and almost geometry independent characterization of ductile tearing resistance by two parameters, namely cohesive strength and separation energy. This works even in cases when classical fracture parameters like \( J \) or CTOD fail as, e.g., for welded structures. The following example (Scheider 2001a) demonstrates the application to a laser welded ferritic steel, St 37. The determination of the cohesive parameters as well as the verification of the model was performed by tests and simulations on micro-specimens,

- Flat tensile bars of pure weld material, 9 mm measuring length, 2 mm width and 0.5 mm thickness;
- Single edge notched bend bars, SE(B), of heterogeneous welded material, 18 mm span, 4.5 mm width and 2.25 mm thickness, having an electro-eroded notch, \( a/W = 0.5 \), in the centre of the fusion zone, see FE mesh in Figure 32.8a;
- Double edge notched tensile bar, DE(T), of heterogeneous welded material, 10 mm measuring length, 4 mm width and 2 mm thickness, having two edge notches, \( a/W = 0.5 \), in the centre of the fusion zone, see FE mesh in Figure 32.8b.

The tensile bars made of pure weld-material broke in uniform elongation at a critical stress of \( \sigma_c = 1000 \) MPa, which was identified as cohesive strength. The bend bars showed a nonlinear deformation behaviour with ductile crack extension, see the load-CMOD curve in Figure 32.9a. The cohesive strength, \( \Gamma_c \), was varied in the simulations until the numerical results fitted the experimental data, and the correspondent value of \( \Gamma_c = 40 \) N/mm was identified as separation energy. Figure 32.9a shows that no significant difference is observed between a 2D plane strain and a 3D calculation. Taking these parameters for a numerical simulation of the mechanical behaviour of the DE(T) reveals that the model gives a good prediction of the experimental load-elongation curve, see Figure 32.9b.

The present example (Scheider 2001a) and other applications in the literature (Siegmund and Brocks 2000, Chabanet et al. 2003) have proven that the cohesive model is a versatile tool for characterising ductile tearing of homogeneous and welded structures.
Figure 32.8: FE meshes of laser welded cracked specimens (Scheider 2001): (a) plane strain model of SE(B), and (b) 3D model of side-grooved DE(T).

Figure 32.9: Load displacement curves of FE-simulations with cohesive model in comparison to test results of laser welded cracked specimens (a) SE(B) and (b) side-grooved DE(T) specimen.

32.5 Summary

The application of elastic-plastic fracture mechanics (EPFM) has been significantly extended by the permanently increasing capabilities of numerical simulations. Classical fracture param-
eters like $J$ or CTOD can be determined even for complicated structures and loading by FE analyses with elasto-plastic constitutive equations. Path dependence of $J$ and $J$-dominance of local stress and strain fields can be studied, thus giving a better understanding of constraint effects and the limits of application of EPFM.

New models acting on a local scale have overcome the limits of global fracture parameters. Whereas $J$ represents the total inelastic work dissipated per crack extension and thus is inherently geometry dependent, the local separation energy, $\Gamma_c$, controls the local material separation and can hence be regarded as a material parameter. The cohesive model with its two parameters, separation energy, $\Gamma_c$, and cohesive strength, $\sigma_c$, provides an advanced characterisation of ductile tearing resistance. However, it is still a phenomenological model.

More physically and micromechanically based descriptions of fracture are supplied by numerous models of damage mechanics, where the local degradation is included as evolution of additional internal variables within the constitutive equations, leading to material softening and failure. Models of ductile damage (Gurson 1977, Rousselier 1987) will be addressed in Chapter 45.

References

References


Int. J. Solids Struct. 11, 601-616.
Shih, C. F., German, M., 1985, Requirements for a one parameter characterization of crack tip fields by the HRR-singularity. Int. J. of Fract. 29, 73-84.
Siegmund, Th., Brocks, W., 1998a, Tensile decohesion by local failure criteria, Technische Mechanik 18, 261-270.
References


33 Rheology of Concentrated Suspensions: A Lattice Model

Yves Brechet, Michel Perez, Zoltan Neda, Jean Charles Barbe, and Luc Salvo

A lattice model to describe the dynamics of cluster formation in suspensions is applied to predict the rheology of concentrated suspensions. This model is applied to the modelling of the apparent viscosity of semisolids as function of shear rate, solid volume fraction and grain size.

33.1 Introduction

Suspensions play a key role in many industrial fields, from the food industry to the cosmetic industry, or to the industry of solvents. In the present chapter, we will concentrate our attention to the case of semi-solids which appear during alloy solidification. In the processing from the liquid state, the limiting factors are mould filling and hot tearing (Campbell, 1997; Lovatt et al., 1999). Both phenomena depend on the rheology of the semisolid, on its viscosity, on its ability to sustain stresses. In the present chapter, we focus our attention on the question of viscosity of a semi solid suspension, which is a key ingredient in the understanding of mould filling defects, and casting porosities associated with limitations of liquid flow.

The study of the mechanical behaviour of mixtures has a long history: the two limiting situations, namely the dilute suspension in a viscous liquid, and the flow of a liquid in a porous solid, have been considered by A. Einstein, (Einstein, 1906) and H. Darcy (Darcy, 1856). The results of their investigations pertain to the classics of the literature: Darcy’s law relates the velocity of a fluid in a porous medium when a pressure difference is imposed (Darcy, 1856), Einstein’s law gives the correction to the viscosity due to the extra dissipation in the fluid flow (Einstein, 1906) introduced by the solid particles. The consequences of fluid flow during solidification on the properties of the casting are reviewed in the classic textbooks by M. Flemings (Flemings, 1974) and J. Szekely (Szekely, 1979).

Beyond their prominent role in the generation of casting defects, the rheological properties of the semisolid are also of interest in metal forming operations such as Thixoforming. The rheofluidification effect which is observed for certain microstructures – i.e. the decrease of the apparent viscosity with an increased shear rate- is particularly interesting.

The rheology of concentrated suspensions and of complex fluids has been studied extensively, in particular in relation with biological fluids, and has led to the concept of “effective volume fraction” which mimics the existence of clusters of particles trapping the fluid. Similarly the study of semisolids at the end of solidification has led to the concept of “cohesive fraction” which illustrates the formation of a “solid dendritic tree” and the transition from
a behaviour alike a suspension, to a behaviour alike a porous medium. These two central concepts will be the backbone of the present chapter.

We will focus on the modelling of the mechanical behaviour of the semisolid, keeping in mind the necessity to introduce the relevant microstructural features and length scales necessary for incorporating a physical interpretation in a phenomenological approach. For a purely phenomenological approach, the reader is referred to Quemada (Quemada, 1985, 1998)

The physical picture which motivates the present approach is the following: the viscosity of a suspension depends on the volume effectively influenced by the particles. For concentrated suspensions, this effective volume has to take into account the liquid trapped inside more or less dense particle clusters. These clusters form by collision between floating particles, and are also destroyed when shocks energetic enough occur. This complex cluster dynamics is the key step in modelling shear rate dependant viscosity. We have chosen a simple cellular automaton approach to account for the collective dynamics of clusters, and when the cluster size and structure under given shearing conditions is known, a semi empirical rule allows to model macroscopic viscosity. The present chapter provides thus an example of modeling at the continuous scale which requires, for a physical understanding of the phenomenon, incorporation of discrete simulation at the mesoscale.

### 33.2 Behaviour of Suspensions: The Generation of Clusters

Figure 33.1 gives the physical reason for the commonly observed “rheofluidisation” at high strain rates. In concentrated solutions, particles tend to form clusters which entrap the liquid so that the effective volume fraction of clusters is larger than the nominal one and accordingly the viscosity is higher.

![Figure 33.1: The three regimes of cluster formation as function of the shear rate and their consequence on the apparent viscosity.](image)
33.2 Behaviour of Suspensions: The Generation of Clusters

At high shear rate the clusters are destabilized by the fluid flow and the viscosity $\eta$ is accurately described by the law of Krieger and Dougherty (Krieger and Dougherty, 1959) which reduces to Einstein formula for vanishingly small volume fractions $\Phi$:

$$\eta = \eta_0 \left( 1 - \frac{\Phi}{\Phi_M} \right)^{-2.5 \Phi_M}$$

(33.1)

At lower shear rates cluster trapping some liquid can remain, whereas at even lower shear rates, when the volume fraction is large enough, very “fluffy” clusters leading to a percolation cluster are expected. We are interested in the present contribution, in the regimes where well separated clusters can be identified, i.e. when percolation is not expected. The dynamic formation of a cluster population, resulting from the competition between particle aggregation and cluster breaking by shear forces from the liquid is expected to depend both on the volume fraction of particles and on the shear rate. It will allow to evaluate as function of these parameters the expression of the effective volume fraction $\Phi_{\text{eff}}$. This effective volume fraction will then replace the nominal one $\Phi$ in Krieger formula to provide with an expression of the viscosity as a function of the shear rate and the nominal volume fraction.

We will consider situations where the evolution of the solid particles will be dominated by the formations of clusters, i.e. the coarsening processes driven by the reduction of surface tension will be considerable slower. A simple estimation of the validity of this hypothesis can be obtained as follows: the typical time for coarsening which is diffusion controlled is $a^2/D$ where $a$ is of the order of the particle radius and $D$ is the diffusion in the liquid. This time scale is of the order of $1 \text{ s}$. The time scale for cluster formation is given by the inverse of the collision frequency, which is of the order of $\Phi \dot{\gamma}$ where $\dot{\gamma}$ is the shear rate. This time scale is of the order of $10^{-3}$ in the range of shear rates of interest. We can therefore safely neglect coarsening and focus on the aggregation dynamics.

Two spheres of radii $a$ colliding in a shear flow will remain in contact during a time of the order of

$$\tau_c = \frac{5}{2\dot{\gamma}}$$

(33.2)

During this time spent in contact, a neck will form, driven by the reduction of surface energy. The theory of neck formation during liquid phase sintering provides an expression for the neck radius $x$ after a time equal to the contact time (Adler et al., 1994).

$$x \propto \left( \frac{a^2}{\dot{\gamma}} \right)^{1/5}$$

(33.3)

This neck will require an energy proportional to its area to be broken. The proportionality constant, in a situation where ductile failure occurs, is expected to be of the order of $10^3$ times the surface energy. This energy will be provided by the kinetic energy of colliding clusters.

This simple description already provides a direct explanation for Figure 33.1: when the shear rate increases, the contact time is lower, the neck are smaller and easier to break, and the kinetic energy available to break them is larger: it can be qualitatively understood that the cluster formation will be less efficient at higher shear rates, and accordingly the apparent viscosity will be lower.
The previous equations have served as a basis for a lattice simulation for the cluster formation. For the details, the reader is referred to the original paper (Perez et al., 2000). We will just give here the principle of the simulation and the main results.

In order to develop a simulation with appropriate numerical efficiency, and since we are interested in a steady cluster distribution, it seems appropriate to assume an ergodic situation: the size distribution of a population of clusters is equivalent to a size distribution of a given cluster along its life. We will therefore consider only one representative cluster and let it evolve. The system is described on a 3D square lattice, where a two state description is chosen: 0 for a liquid site and 1 for a solid site. The time step is given by the frequency of collision. During a collision between two clusters, the probability of aggregation (lower than one) account for the need for approaching with a favorable crystallographic orientation. The probability of breaking a cluster during a collision will be taken as a linear function of the ratio of the kinetic energy to the energy necessary to break the neck. The fracture will take place randomly inside the cluster. Both the gyration radius of the cluster and the amount of liquid trapped inside are measured and allow to estimate the effective volume fraction. Two typical clusters for two different shear rates are shown in Figure 33.2.

![Clusters for two different shear rates.](image)

The size distribution of a given cluster along its life time can be accurately described by a log normal distribution. Without being a proof of the ergodic hypothesis, the fact that a well defined distribution shown in Figure 33.3a, can be obtained by normalization to the average radius for different shear rates is a good sign.

Figure 33.3b shows a clear decrease of the average size of the cluster with increasing shear rate, as expected qualitatively. In addition, a lower shear rate leads to less compact clusters.

The simulated size distribution can be compared with one of the very few available 3D data obtained via serial slicing of solidified specimens (Wan and Sahm, 1990).
33.3 Conclusions

As a consequence, the effective volume fraction decreases with increasing shear rate, as shown in Figure 33.5. As could be expected since the cluster formation depends on the available surface, the difference between the effective and the nominal volume fraction is more pronounced for smaller grains.

The resulting decrease in viscosity with increasing shear rate is shown Figure 33.6, with a comparison with experimental data.

**33.3 Conclusions**

One can conclude that the present model, in spite of its simplicity, describes correctly the rheological behaviour resulting from cluster formation when the elementary particles can be safely regarded as spherical. The implementation of a dendritic structure as the elementary building block is the next challenge to be dealt with.
References

Figure 33.5: Effective volume fraction normalized by the nominal volume fraction, as a function of the shear rate.

Figure 33.6: Comparison between simulation results and experimental viscosity measurements (Poulain et al., 1999) of an Al-6.5wt% Si alloy, as a function of shear rate and solid fraction. The shaded area representing simulation results are obtained for a range of particle radii between 50 and 100 µm.

Acknowledgements

The present work has greatly benefitted from numerous discussions with Dr. M. Suery.

References


Part III

Application to Engineering Materials Processes
34 Solidification Processes: From Dendrites to Design

Jonathan A. Dantzig

Nearly all metal production begins with melting and casting. In this article, I discuss the development of microstructure during solidification, and the processes which occur over the range of associated length scales. The emphasis will be on modeling techniques which permit resolution of the multiple scales. I then describe methods for using inverse design and optimization in conjunction with the models to improve product and process design.

34.1 Introduction

Most metal production begins with a melting and casting operation. In some processes, such as foundry and investment casting, the cast part is essentially produced at net shape. In other processes, such as continuous casting, the cast product is subjected to substantial thermomechanical processes, such as rolling or forging, to reach the final shape. In either case, the cast microstructure is of primary importance in determining the properties and performance of the product.

The microstructure forms from the interaction of heat transfer with solute transport during solidification. Hot liquid metal is poured into a cold mold, and the heat transfer from the melt to the mold determines the solidification rate. The important length scale for heat transfer is the dimensions of the casting. External dimensions, and differences in section size produce various solidification times, typically ranging from a few seconds to as long as hours in a very large ingot.

As the solidification front moves, the solute is locally redistributed from the (usually) homogeneous state in the melt to a segregated state in the solid. The segregation arises from the solute redistribution required to satisfy thermodynamic equilibrium in the solid. The data necessary to understand this difference is incorporated in the equilibrium phase diagram for the alloy. To understand the evolution of structure, let us consider a characteristic time \( t_c \) for the solidification process, determined by the heat transfer rate.

The characteristic length scale for thermal diffusion is then given by \( \delta_T \sim \sqrt{\alpha t_c} \), where \( \alpha \) is the thermal diffusivity. (Cf. Dantzig and Tucker (2001).) Using a similar analysis, the characteristic length scale for solute diffusion is \( \delta_C \sim \sqrt{Dt_c} \), where \( D \) is the chemical diffusivity. The ratio of these two length scales is then \( \delta_C/\delta_T \sim \sqrt{D/\alpha} \), which for a typical metallic system is 0.01 or smaller. This disparity in length scales produces the microstructure.

There are numerous other important physical phenomena during solidification as well, including kinetics of interface attachment and nucleation of critical nuclei. The hierarchy of length and time scales is summarized in Figure 34.1. Also shown in the figure is a series of...
of boxes, representing the approximate length and time scales which are included in certain types of analyses. The size of the box is dictated by the computational speed, memory and disk storage capacity of current computers. Notice that the length and temporal scales cover a range of about three orders of magnitude. Thus, the possible phenomena which may be included in a given model are limited.

![Figure 34.1: Hierarchy of length and time scales in solidification.](image)

In this article, we focus in the first part on solidification microstructure, and that will require us to extend the box in order to include the important physical phenomena. We accomplish this through the use of adaptive finite element analysis grids, which enables resolution of the relevant length scales. The above microstructural scaling analysis assumed that the dominant transport process is diffusion of either heat or solute. However, when fluid flow is present, transport is typically dominated by advection. We further explore this phenomenon in the next section, where we discuss modeling of dendritic growth in the presence of fluid flow.

In the following sections, two different levels of solidification modeling are described. In the next section, I present models of dendritic growth using the phase-field method. The emphasis is on the resolution of the disparate length scales using adaptive finite element analysis techniques. In the following section, I describe methods for solving inverse problems, with an application to casting design.

### 34.2 Dendritic Microstructures

Dendrites are the most common solidification microstructure in metals and alloys. The basic process of solidification can be understood by considering a spherical solid nucleus of a pure crystal in an infinite undercooled melt. Mullins and Sekerka (1964) showed that the spherical geometry is unstable to morphological perturbations once the nucleus grows to a still-small
size. When the interface becomes unstable, crystalline anisotropy also becomes important, and dendritic structures begin to appear. Eventually, the dendritic seed will have extended arms, which grow at a steady speed $V$ which depends on the undercooling $\Delta T$. A thermal boundary layer appears ahead of the dendrite tip, with extend $\delta = \alpha/V$.

The theory of dendritic growth is presented in detail in Provatas et al. (1998) and Karma and Rappel (1995), among others, so the discussion here is limited to that which is necessary for the current study. We model dendritic growth using a phase-field model (Langer, 1987), in which the sharp interface between the liquid and solid is replaced by a diffuse interface of controlled width $W$. Karma and Rappel (1995) showed that the phase-field model converges to the sharp interface model when the ratio of the interface width to the diffusion boundary layer thickness $W\nu/\alpha$ is small.

A converged computation requires that the grid spacing $\Delta x$ resolve the diffuse interface, so that $\Delta x \sim W$. Further, the computational domain length $L_B$ must be large compared to $\alpha/V$. These two criteria combine to require that $L_B/\Delta x \sim 10^3$. Thus, if we were to use a uniform grid, we would require $10^6$ grid points in 2-D, and $10^9$ grid points in 3-D.

We can finesse this length scale problem by recognizing that we need the highest resolution only near the interface, and the far-field diffusion can be resolved on a much finer grid. Provatas et al. (1998, 1999) developed an 2-D adaptive finite element code based on this notion, using quadtree data structures and gradient-based error estimators to perform the grid refinement. The reader is directed to the original works for the details. Example results are shown in Figure 34.2. An important finding of this work is that at high undercooling, each arm of the dendrite behaves as a single, isolated branch, whereas at low undercooling, the thermal fields of adjacent arms interact to affect each other during growth.

![Figure 34.2: Computed dendritic structures at high undercooling (left) and low undercooling (right). The color refers to temperature, and it is clear that the low undercooling case has the thermal filed of both arms interacting, whereas the arms are independent at higher undercooling.](image)

Jeong et al. (2001, 2003) extended this work into three dimensions in order to examine the role of fluid flow. The computational complexity of these problems is considerably higher, and they also developed a parallel implementation of the algorithm. As an example, consider
a similar problem to that in 2-D, but this time there is also an imposed fluid flow directed
toward the growing seed.

Figure 34.3 shows a sample result of a dendrite grown from a small spherical seed with a
superimposed fluid flow. The left hand figure shows the finite element grid with the elements
assigned to each of 32 processors filled with a different color. The figure demonstrates the
power of the adaptive grid algorithm, where the mesh density is distributed to resolve the
interface, and the surrounding domain can be made quite large with relatively little additional
computational cost. The right hand figure gives another rendering of the same dendrite with
the fluid flow indicated by the particle tracers. Notice the enhanced growth and instability
on the leading edge, and suppressed growth on the trailing edge, indicating the importance
of advective transport in microstructure development. The reader is referred to the original
works for further details of the calculations.

Figure 34.3:Finite element grid (left) at a late stage of the computation of the evolution of a
dendritic seed in a fluid flow. Color indicates which of 32 processors a particular element is
assigned to. The right hand side shows the corresponding dendrite and particle paths.

The work presented in this section indicates the detailed structural information which can
be computed using advanced computational methods. It should be noted that this calculation
is close to the limit of currently viable computation. However, we have computed only one
dendrite, whereas there are millions of such dendrites in a cast product. Like snowflakes, each
of these crystals will be different. We can use models such as this one to determine growth
laws to be used in more macroscopic computations.

34.3 Inverse Problems and Optimal Design

The macroscale product resides at the other end of the length scale spectrum from the indi-
vidual dendrites of the last section. External geometry and boundary conditions determine
the heat flow, and thus solidification characteristics of the product. These data would then
feed back to a microscopic model to describe the evolution of the dendrite. But there are also important issues at the macroscopic scale.

Once the product geometry and boundary conditions are set, the heat flow problem to determine the temperature history is straightforward. However, the result may not produce a cast product with desirable properties. Local heat transfer conditions will determine the microstructure, as discussed earlier, but some conditions may lead to defects or other undesirable properties in the part. In this section, a formal procedure is presented for using simulations in concert with numerical optimization techniques to improve the product design.

In order to do this, we introduce a vector of design variables $b$ which represent parameters of the simulation which have the potential to be changed to improve the product. The design variables may include certain part dimensions, processing conditions, material selection, etc. Virtually any parameter which enters the simulation can be used as a design variable in this framework. We also introduce a measure of product quality $G$ which may depend on the solution (e.g., the temperature field) which we write in abstract form as $u$ and on $b$, i.e. $G = G(u, b)$. The objective function $G$ is chosen such that an extremum value of $G$ corresponds to the optimal solution. Thus, we may quantify the process of design improvement as the minimization (or maximization) of $G$ with respect to the design $b$. It is also possible to include in a straightforward way constraint functions of the form $F(u, b) \geq 0$. We limit the discussion here, however, for brevity and clarity of presentation.

The optimization procedure begins with an initial design, and then we progress through a series of alternative designs in search of the optimum. We recognize, however, that each design we choose to evaluate represents a potentially expensive finite element simulation of solidification of a complex geometry. Thus, we must limit the number of designs we evaluate in order to traverse the design space efficiently. Efficient optimization algorithms require derivatives in design space, called design sensitivities. We can compute these formally as

$$\frac{dG}{db} = \frac{\partial G}{\partial b} + \frac{\partial G}{\partial u} \frac{\partial u}{\partial b}$$

(34.1)

The last term in Equation (34.1) is called the response sensitivity, and it represents the change in the field solution when the design variables change.

Note that the evaluation of the response sensitivities by finite difference perturbation of the design variables is both inaccurate (due to roundoff errors) and inefficient, because multiple analyses must be run. We now describe an efficient procedure for evaluating $\partial u/\partial b$. For any given design, we may write the analysis in residual form

$$R(u, b) = 0$$

(34.2)

The problems of interest are nonlinear, and so we solve them using Newton-Raphson iteration:

$$\frac{\partial R}{\partial u} \frac{\Delta u}{\partial u} = -R(u^i, b)$$

(34.3)

where the subscripts refer to the iteration count, and we implement the update scheme

$$u^{i+1} = u^i + \Delta u$$

(34.4)

This scheme is then iterated to convergence.
Once the above “forward solution” is obtained, we proceed to compute the response sensitivity. Begin by direct differentiation of Equation (34.2) with respect to the design

$$\frac{dR}{db} = 0 = \frac{\partial R}{\partial b} + \frac{\partial R}{\partial u} \cdot \frac{\partial u}{\partial b}$$  

(34.5)

Rearrangement of Equation (34.5) provides a form for determination of the response sensitivities

$$\left( \frac{\partial R}{\partial u} \right) \cdot \frac{\partial u}{\partial b} = -\frac{\partial R}{\partial b}$$  

(34.6)

The efficiency of this approach comes when we notice that the same tangent matrix $\partial R/\partial u$ appears in both the sensitivity and forward problems. Thus, we may take advantage of the work done to decompose the tangent matrix in the forward problem, and obtain the sensitivities by computing a series of new right hand side vectors (see Equation (34.6)) and performing back-substitution. This process costs a small fraction of that of the forward solution, typically 10–15% per design variable. This process can be extended in a straightforward way to transient problems (Morthland et al., 1995).

As an example, consider the solidification of a steel hammer casting. This work has been described elsewhere (Morthland et al., 1995), so only a brief sketch is given here. A large steel hammer is cast with a riser on top to provide feed metal for solidification shrinkage. In the original design, the riser was not large enough to provide adequate protection of the casting. The failure of the initial design is illustrated in Figure 34.4 (left), where the outline of the casting is shown along with a gray area corresponding to the area of remaining liquid after a certain time. Note that there are two separated regions, which will lead to porosity defects in the final product.

We then parameterized the riser dimensions and made them the design variables. The objective function which we minimized was the volume of the riser, and we also imposed a constraint on the temperature field in order to enforce directional solidification of the casting into the riser. See reference Morthland et al. (1995) for further details. The forward problem was solved using FiDAP (Engelman, 1991), which we modified to perform the sensitivity computations described above. We then combined the code with numerical optimization software DOT (Vanderplaats, 1984) to perform the optimization automatically. A total of 24 designs was evaluated to produce the optimal design shown in Figure 34.4 (right).

### 34.4 Conclusion

Solidification problems can produce a wide array of interesting pattern selection problems. At the microstructure scale, heat transfer, crystallography and mass transfer combine to produce dendrites. These structures are amenable to analysis using the phase-field method.

At the more macroscopic scale, the moving interface sweeps out patterns which determine the boundary conditions which produce the microstructure. Under some conditions, these patterns may be undesirable. Numerical optimization in conjunction with large scale finite element analysis provides a powerful tool for manipulating the product design to remove such defects.
Figure 34.4: Comparison of optimal and original designs. The gray area represents remaining liquid in the casting.

Acknowledgement

This article represents a collection of work performed over a number of years. Sponsors have included NASA, the National Science Foundation, Ford, and Deere and Co. I have also had numerous collaborators whose contributions to this work are very significant. These include N. Goldenfeld, N. Provatas and J.-H. Jeong for the work on dendritic growth, and D. Tortorelli, P. Byrne and T. Morthland for the work on optimal design.

References


References


This chapter describes micromechanical and macroscopic material models for granular and porous materials and applications to powder technological processing steps. The emphasis lies on powder compaction and sintering, while for the other steps of the process chain (powder fabrication and conditioning, pouring of the loose powder in the die, post-sintering mechanical densification, service behavior) appropriate models are named and key references are given. The constitutive models are implemented in finite-element codes and applied to industrial problems. Case studies are presented.

35.1 Introduction

Many materials are produced by powder metallurgical processes. The reason may be that the material is difficult to melt and cast, as for example ceramic materials, hard metals, cermets or refractory metals. Another reason for using powder technology may be that it is economically attractive to make complex shaped structural parts by powder compaction and sintering; this is the case for sintered steel parts, which are used in large quantities, for example, in combustion engines and gear boxes. For other products, such as magnets or multilayer electronic devices, there is no practical alternative to powder technological production processes. Also pharmaceuticals are often available in powder form, which is then compressed to produce tablets. Further, some polymers, such as PTFE, are processed by pressing and sintering. Finally, powder metallurgy is applied to obtain high-grade materials with well-controlled, uniform microstructure, e.g. for gas turbine components.

The present article is grouped around two of the most important processing steps in powder technology, namely powder die compaction and sintering. The preceding steps in the process chain (powder production and conditioning, die filling), and post-sintering steps are only briefly described. Besides die compaction there is a great variety of alternative shaping techniques, by which a powder compact with the desired shape can be produced. A few of the more important examples are slip casting (for sanitaryware ceramics), tape casting and lamination (for functional ceramics), cold isostatic pressing (for tableware and engineering ceramics), powder injection molding (for complex shapes) and extrusion (for ceramic insulators).

Uniaxial powder compaction in a die usually does not yield a homogeneous density distribution in the pressed part (the so-called green body). Rather, the green density is more or less inhomogeneous depending on the part geometry, the tool design and the friction between powder and die wall. As a result, the part undergoes shape distortions during sintering, as
regions with different green density shrink at different rates and by different amounts. It is not unusual that cracks develop during pressing and sintering. Sinter distortions may also be a consequence of gravitational forces or friction on the support during firing. Since, however, the requirements for close geometrical tolerances increase continuously, an extra processing step, such as hard machining or sizing may become necessary, if the distortions are too large.

The aim of the modeling activities described in this article is to provide simulation tools, which allow to identify problems, such as distortions and cracking, already in the design phase, before large investments into tools and other equipment have been made and much time and money has been spent in trial and error cycles. Most of the simulation tools make use of the finite element method in combination with appropriate material laws, which may either be phenomenological continuum models (like the Drucker-Prager-Cap model used for powder die compaction) or micromechanics based models (like the models used for sintering).

35.2 Powder Production

Atomization of a molten metal for metal powder production as the first step in the production route usually is achieved by means of two-fluid atomization with inert gas. Modeling and description of the melt disintegration process is of great importance and needs primarily information about the gas flow field in the nozzle vicinity and the influence of nozzle configuration and operation parameters on the resulting spray. Several simulation methods, mostly based on computational fluid dynamics, have been developed, e.g. by Grant (1995), Lavernia and Wu (1997), Fritsching et al. (1997), Pryds et al. (2002), Delplanque et al. (2002).

35.3 Die Filling

Filling the powder into the die by a moving shoe is a very complex process, in which inhomogeneous filling densities as well as totally unfilled areas can occur, especially if the cycle times are shortened more and more in order to increase the throughput. In practice, the initial packing density depends on the kinematics of the die filling process. There is a vast literature on granular flows, ranging from studies of hopper flows (Waters and Drescher 2000) to avalanches (Hutter 1996). An extensive overview about modeling of granular materials is given by Herrmann and Luding (1998). In the area of powder technology first filling simulations based on discrete elements have emerged only recently. Gillia et al. (2002) and Wu et al. (2002) have simulated the filling of a stepped die (Figure 35.1). The authors use a discrete element method considering irregularly shaped particles, gravity and the effect of the escape of air from the cavity. They obtain qualitative agreement with model experiments.

35.4 Powder Compaction

Continuum plasticity models (see also Chapter 6) have often been used to describe the material behavior during powder die compaction. These phenomenological models, which were originally developed in soil mechanics, are characterized by a yield criterion, a hardening function
and a flow rule. Several specific formulations are used for modeling powder compaction (e.g., Cocks 2001, Kraft 2003).

Besides the phenomenological models micromechanical models were also developed (Fleck et al. 1992, Fleck 1995, Storakers et al. 1999). Starting from the interaction forces between powder particles these authors derive macroscopic stress-strain relations for granular materials. Evolving anisotropy as a consequence of non-isotropic compaction is taken into account. These models are conceptually convincing, and they were successful in explaining the results of triaxial tests on metal powders with nearly spherical particles (Akisanya et al. 1997). For many commercial powders, however, the assumptions underlying the micromechanical models are apparently violated to an extent that the phenomenological models are more successful than the micromechanical models (Rottmann et al. 2001).

### 35.4.1 The Drucker-Prager-Cap Model and Finite Element Implementations

Die compaction of many powders has been modeled successfully with the Drucker-Prager-Cap model, whose yield surface consists of the Drucker-Prager failure line $F_s$ (Drucker and Prager 1952) and the elliptic cap $F_c$ (DiMaggio and Sandler 1971), which provide a combined convex yield surface in the plane of the first and second stress invariants as shown in Figure 35.2 and characterized by the equations:

$$F_c = \sqrt{(p - p_a)^2 + (Rq)^2} - R(d + p_a \tan \beta) = 0$$

$$F_s = q - p \tan \beta - d = 0$$

where $p =$ hydrostatic pressure (i.e. negative mean stress), $q =$ von Mises equivalent stress. The parameters $R =$ cap eccentricity, $d =$ cohesive strength, $\beta =$ cohesion angle are constant in the original version of the model, and $p_a$ is a hardening function depending on the density. In Figure 35.2 two extensions of the original model are also included: the tension cutoff $T$,}

---

*Figure 35.1: Simulation of the filling of a stepped die (left: discrete element simulation, right: velocity field) (from Gillia et al. 2002).*
which characterizes the tension strength of the powder compact, and the von Mises yield strength $\sigma_y$ of the dense material. Inside the yield surface the powder behaves elastically. If the stress state reaches the yield surface, the powder deforms plastically. The density increases, if the stress state is on the cap, whereas it decreases (dilatation), when the stress state reaches the failure line. Dilatation implies softening, so that strain localization and cracking may occur. In some general purpose finite element packages this model is already implemented.

Figure 35.2: Modified Drucker-Prager-Cap model in the $p - q$ plane ($p =$ hydrostatic pressure, $q =$ von Mises stress) with tension and von Mises cutoffs.

To describe the powder behavior more realistically – especially with respect to crack formation during pressing, unloading or ejection – the Drucker-Prager-Cap model was modified by Coube and Riedel (2000). It is plausible that not only the hardening function $p_a$, but also the cohesion parameters $d$, $\beta$ and $T$ should depend on the density. In the following relations, the density $\rho$ and the volumetric plastic strain are alternatively used. They are related by

$$\varepsilon_{\text{vol}}^p = \ln(\rho/\rho_0)$$

(35.3)

where $\rho_0$ is the fill density. As is common in soil mechanics, the volumetric strain is defined positive during compaction. The hardening relation, the cap eccentricity and the cohesion parameters are described by the following empirical expressions:

$$\varepsilon_{\text{vol}}^p = W(1 - \exp(-c_1 p_a))$$

(35.4)

$$R = R_1 + R_2 \exp(R_3 \rho)$$

(35.5)

$$d = d_1 \exp(d_2 \varepsilon_{\text{vol}}^p)$$

(35.6)

$$\tan \beta = b_1 - b_2 \varepsilon_{\text{vol}}^p$$

(35.7)

The parameters $W$, $c_1$, $c_2$, $R_1$, $R_2$, $R_3$, $d_1$, $d_2$, $b_1$ and $b_2$ are determined by experiments. Numerical values for an iron-base powder were given by Coube and Riedel (2000), and values for an alumina powder are reported in the next section. Instead of Equation (35.5) alternative
functional forms have also been proposed to describe the observed dependencies (PM Mod- 
net 1999). This model was implemented as a user subroutine in ABAQUS/Explicit. In this 
implementation, the flow rule is taken to be associated on the cap, while it can be chosen to 
be non-associated or associated on the failure line.

The model in the above form was applied to numerous complex three-dimensional geomet- 
ries, and it was found to predict density distributions well. However, the model turned out 
to be too stable against strain localization and cracking. Therefore, a second version of the 
model was proposed by Coube and Riedel (2000), in which the cohesion parameters \( d \) and \( \beta \) 
are treated as state variables obeying evolution equations, rather than as unique functions of 
the density. This model describes cracking more realistically and it was also implemented in 
ABAQUS/Explicit.

At the present time many groups use dynamic explicit codes for compaction simulations, 
especially for complex three-dimensional parts due to several advantages compared to im-
licit codes. For example, explicit codes appear to be less sensitive to large element distortions and can handle contact and friction more efficiently. Often the well-known commer-
cial finite-element codes are used, sometimes with own material models implemented, but 
some authors have developed their own FE codes like PRECAD (Baccino and Moret 2000) or 
CRADA (Ewsuk et al. 2001).

35.4.2 Experiments to Determine the Drucker-Prager-Cap Parameters

To determine the model parameters, tests on relatively simple samples can be used. The hard-
ening variable \( p_a \) can be obtained from uniaxial compaction tests on cylindrical samples. The height-to-diameter ratio of the pressed cylinders should be small in order to minimize the effect of wall friction. Otherwise, the effect of friction must be taken into account in the evaluation of the pressing experiment (Roure et al. 1996). More expensive triaxial tests are necessary to determine the shape of the cap (Pavier and Dorémus 1999, Cocks 2001). Yet, the exact value of the cap eccentricity has an only minor influence on the final density distribution in many practical cases. In special cases, however, for example when a considerably profiled upper punch intrudes into the powder, the shape of the cap can influence the predicted density distribution significantly, so that a more reliable value of the eccentricity parameter \( R \) is needed. The failure line can be determined by measuring the green strength in at least two test configurations with different stress triaxiality, such as the compression test, the Brazilian disc test, and the four-point-bending test. For a general description of characterizing green compacts see, e.g., Darvell (1990), and for a discussion of possible problems with the Brazilian disc test see Fahad (1996).

Figures 35.3 to 35.5 show results for a 96% pure alumina ceramic powder with a fill den-
sity of 1.02 g/cm\(^3\) from CeramTec AG. Figure 35.3 plots the hardening variable versus the plastic volumetric strain. To determine the cohesion parameters \( d \) and \( \beta \), uniaxial compres-
sion, Brazilian disc and 4PB tests were carried out on cylindrical specimens pre-compacted to 
different densities. The ratio \( q/p \) of deviatoric and hydrostatic stress is 3 for the compression 
test, \((3/2)\sqrt{3} \) in the center of the Brazilian disc specimen and \(-3\) for the 4BP test. The test 
results were plotted in the \( p-q \) plane, and points from the three test configurations with equal 
density were connected by a straight line, the Drucker-Prager failure line (Figure 35.4). The 
slope and the intercept with the \( q \) axis give the cohesion parameters. The results for \( d \) are
shown in Figure 35.5. The cohesion angle is practically constant. The cohesion parameters then have the values $d_1 = 0.00877 \text{ MPa}$, $d_2 = 6.33035$, $b_1 = 2.2747$, $b_2 = 0.09038$. The cap eccentricity is taken to be constant, $R = 0.5$, and the hardening parameters are obtained from a die-compaction test: $W = 0.84042$, $c_1 = 0.61894 \text{ MPa}^{-c_2}$, $c_2 = 0.57469$.

![Figure 35.3: Hardening variable $p_b$ versus plastic volumetric strain for an alumina powder. (For the relation of $p_b$ to $p_a$, see Figure 35.2).](image)

![Figure 35.4: Construction of failure lines for different densities for an alumina powder.](image)

Finally, friction coefficients between powder and die are needed for the simulation, especially for slender parts, in which the density distribution may depend significantly on friction (PM Modnet 1999, PM Modnet 2002). Several experimental methods are available (Roure et
35.4 Powder Compaction

Figure 35.5: Cohesive strength $d$ versus plastic volumetric strain for an alumina powder.

al. 1996, PM Modnet 2002). Usually, Coulomb type friction laws are used, yet other relations like the Tresca law, which relates the shear frictional stress with the shear yield strength of the compact, were also applied. Despite the fact that the friction coefficient for most materials decreases significantly with increasing powder density and depends also on other effects like sliding velocity or contact pressure (PM Modnet 1999) the need for very sophisticated friction models is limited (Wikman et al. 1997). Thus, a simple density dependent expression for the friction coefficient is normally sufficient. Even a constant value can often be used without a considerable loss of accuracy.

35.4.3 Example

As an example, Figure 35.6 shows the predicted green density distribution for a disc seal made of $\text{Al}_2\text{O}_3$. For details of this example see Kraft and Riedel (2002). For practical applications it is important that the geometry of the part can be imported from a CAD system without too much manual labor, and that automatic mesh generation is available. In this example, the mesh of the powder contains about 56,000 tetrahedral elements not considering the rigid elements used for the tooling. The part is pressed with a multi-level tool. The three penetrating upper punches densify the powder below them to a high density, while the density is low at the rim of the indentations. In this simulation, the movement of the tool components was prescribed to simulate the real process. It is then possible to investigate the effect of other pressing schedules on the density distribution.

In a few cases the calculated density distributions were compared with experiments. Excellent agreement was found for several simplified synchronizer hubs made of steel powder (PM Modnet 1999, PM Modnet 2002) and for a crown gear wheel (Koch 2002). Kraft et al. (2003) show that the simulated density distribution in another alumina disc seal is consistent with computer tomography measurements.
35.5 Sintering

Nearly all production processes in powder technology include a sintering step at high temperature. The driving force for particle coalescence and for the reduction of porosity results from the surface energy of the particles. High temperatures are needed to allow the transport of matter by solid-state diffusion to form the necks between the particles and to fill the pores. Examples for materials consolidated by solid-state sintering are molybdenum, ZrO$_2$, AlN, Al$_2$O$_3$ and SiC (for detailed tables see German 1996) On the other hand, some materials are preferably sintered by liquid-phase sintering, e.g. hard metals, most sintered steels, Si$_3$N$_4$, and some varieties of Al$_2$O$_3$ and SiC. These systems contain sintering aids which form a melt at moderate temperatures. To be effective, the melt must wet the solid particles and it must be able to dissolve the solid providing an easy diffusion path for it. Liquid-phase sintering is preferred in practice, since it allows lower process temperatures and it is indispensable in cases like Si$_3$N$_4$, which becomes chemically unstable, before solid state diffusion takes place at a sufficient rate.

Sintering is nearly inevitably accompanied by grain coarsening, a process which is usually undesired since coarsening has an adverse effect on toughness and strength. Apart from this
practical argument coarsening must be an integral part of a mechanism-based sintering model, as the grain size influences the sintering rate strongly.

Another general aspect is that the porosity changes its morphology during sintering. In the early and intermediate sintering stages the pore space forms a three-dimensionally interconnected network the particles being connected by isolated contact areas, called necks. In a certain density range (around 90% relative density) pore channels are pinched off leading to a transition from open to closed porosity in the final sintering stage.

35.5.1 Models for Solid-State Sintering

Sintering theory developed continuously from the late forties onwards. The most successful approaches to understand the mechanisms of neck growth and shrinkage have been the two-sphere model due to Frenkel (1945) and Kuczynski (1949) for the early sintering stages (particle bonding), the cylindrical pore channel model due to Coble (1961) and the extrapolated two-particle model due to Kingery and Berg (1955) for the intermediate shrinkage state, and the spherical pore model due to MacKenzie and Shuttleworth (1949) for the late sintering stages. More sophisticated models were developed taking superpositions of various sintering mechanisms into account (Ashby 1974, Arzt 1982, Exner and Arzt 1996).


In the following, the sintering model developed by Riedel and Blug (2001) (see also Kraft and Riedel 2003a) is briefly described. It summarizes results from many of the papers quoted above, and it combines results on sintering with models for grain growth in porous solids. The model distinguishes three sintering stages. The first stage is characterized by grain rearrangement and by the formation of necks between particles. These initial processes have been modeled in an empirical way so far. It is assumed that re-arrangement contributes to the densification rate at low densities, whereas its contribution vanishes when the density approaches that of a random dense packing, i.e. 63% relative density for a packing of equal spheres.

The second sintering stage is still characterized by open porosity, but the necks are now well developed and the particle surfaces having constant curvature. This means that the surface is in equilibrium. Figure 35.7 shows a few such equilibrium configurations for a particle belonging to a periodic, body-centered cubic arrangement of particles. The pore surface is displayed with the computation mesh, whereas the contact areas to neighboring grains are shown without the mesh. The dihedral angle $\psi$ is defined by the local equilibrium between the specific grain boundary energy $\gamma_b$ and the surface energy $\gamma_s$ according to

$$\cos \psi = \frac{\gamma_b}{2\gamma_s}.$$  \hspace{1cm} (35.8)
These results from Svoboda et al. (1994) can nowadays be reproduced conveniently by a program called Surface Evolver available from Brakke (2003).

From a sequence of such equilibrium configurations one obtains the evolution of the surface curvature and of the size of the contact areas. Densification occurs by grain boundary diffusion of matter out of the contact areas. Solution of the equation for stress-directed grain boundary diffusion gives the relation between the interparticle force and the rate at which they approach one another. Using a Taylor-Bishop-Hill-type argument, one can translate this local relation into a macroscopic relation between stress and strain rate as in Equation (35.9) below. The solution for grain boundary diffusion alone is modified to allow volume diffusion as a parallel diffusion path following an argument of Ashby (1974). Surface diffusion is also taken into account. It acts as a process in series to grain boundary diffusion.

At a certain relative density (about 90% depending on the dihedral angle) the open pore channels close, and the transition to closed, isolated pores takes place. The grain boundary diffusion problem for that case was solved by Riedel et al. (1994a). An interpolation between open and closed porosity was proposed by Riedel and Blug (2001).

On the macro-level the solutions of the diffusion equations on the micro-level lead to the following linear viscous constitutive equation with bulk viscosity, $K_v$, shear viscosity, $G_v$, and sintering stress, $\sigma_s$.

$$
\dot{\varepsilon}_{ij} = \frac{\sigma'_{ij}}{2G_v} + \delta_{ij} \frac{\sigma_m - \sigma_s + \Delta p}{3K_v}
$$

(35.9)

where the prime denotes the deviator, $\sigma_m$ is the mean (or hydrostatic) stress, $\Delta p$ is a gas overpressure which may develop in closed pores, and $\delta_{ij}$ is the Kronecker symbol. The parameters $G_v$, $K_v$, and $\sigma_s$ depend on temperature, density, and on the history dependent grain size. Phenomenological models are based on an analogous form with empirical functions for $G_v$, $K_v$, and $\sigma_s$ (Olevsky 1998).

The densification rate is given by the trace of the strain rate tensor

$$
\dot{\rho} = -\rho \dot{\varepsilon}_{kk} = \rho \frac{\sigma_s - \sigma_m - \Delta p}{K_v}
$$

(35.10)
where $\rho$ is the relative density. The densification rate during pressureless sintering ($\sigma_m = 0$) decreases with the fourth power of the grain size, since $\sigma_s \propto 1/\text{grain size}$ and $K_v \propto (\text{grain size})^3$. The grain size plays the role of an internal variable. Its evolution equation is based on Hillert’s theory of grain coarsening (Hillert 1965) modified by a factor accounting for the presence of pores. The pores are either dragged with the migrating grain boundaries or they detach from them, if the boundaries move too fast.

The described model contains a number of parameters. The most important ones are the grain boundary diffusion coefficient, the grain boundary mobility and the bulk diffusion coefficient (each with activation energy and pre-exponential factor). All parameters can be determined from sinter forging tests, in which the specimen is exposed to a uniaxial load during sintering, and axial and radial strains are measured. The test is also called load dilatometry. Riedel and Blug (2001) give the results of a test series on a SiC ceramic and adjust the model parameters. It is found that the model is able to describe the coupled sintering and grain coarsening kinetics under various mechanical stresses rather accurately. Kraft and Riedel (2003a) have applied this model to simulate sintering (and compaction) of a face seal made of SiC. The model was also applied to molybdenum successfully (Sun et al. 1997, Plankensteiner et al. 1999). An example for the adjustment of the model parameters is shown in the following section.

Depending on the preceding shaping process, sintering rates may be different in different directions. After uniaxial die compaction, an only moderate anisotropy is observed, while tape casting usually leads to a pronounced anisotropy. An anisotropic sintering model was presented by Markondeya Raj et al. (2002).

Sintering and grain coarsening have been modeled also with Monte-Carlo simulations (Tikare et al. 2001). Other computer based multi-particle simulations have been applied in the past and will become more powerful as the performance of molecular dynamics codes increases.

### 35.5.2 Liquid-Phase Sintering

The starting point for liquid-phase sintering theories has been the early model due to Kingery (1959) in spite of its geometric simplifications and lacking qualitative agreement with experimental observations. Reviews on sintering models for liquid phase sintering are available (German 1985, 1996, Exner and Arzt 1996).

The liquid phase sintering (LPS) model developed by Svoboda et al. (1996) includes: (i) primary rearrangement of solid particles when the liquid phase is formed, (ii) grain shape accommodation by contact flattening via the solution/precipitation mechanism, and (iii) filling of large pores and grain coarsening in the final stages of sintering. All of these can act simultaneously, although rearrangement dominates initially, followed by contact flattening, and finally coarsening and large pore filling.

The model is formulated in terms of the state variables relative solid density, $D_s$, flattening strain, $\delta$, grain radius, $R$ and the liquid fraction, $g$, which is considered to be a prescribed function of temperature, and hence a prescribed function of time for a known temperature history. $D$ is the total relative density and is related to the solid density through $D = D_s/(1 - g)$. Parameters of the model are the reaction rate constant of the solution/precipitation process, the diffusion coefficient, the viscosity of the liquid, two interface energies, and three parameters
describing the pore size distribution function. Adjustments of the model parameters to Si₃N₄ and to a hard metal are given by Svoboda et al. (1996) and McHugh and Riedel (1997).

35.5.3 Parameters of the Liquid-Phase Sintering Model for an Alumina Ceramic

The model parameters can be obtained again from sinter forging (or load dilatometry) tests. In the following, the determination of the parameters is shown for a 94% purity Al₂O₃ powder. This ceramic sinters via liquid phase sintering, although the amount of liquid phase is only small. Figures 8 to 10 show the evolution of the grain size, of the density and of the axial and radial strains. The model parameters were adjusted using a PC-FORTRAN program, which integrates the evolution equations of the model. The parameters were varied by hand, until the simulated behavior agrees with the measured one. Apparently, the quality of the fit is excellent for the strains.

The resulting parameters of the LPS-model for the Al₂O₃ powder have the following values (for the meaning of the quantities see Svoboda et al. 1996 and McHugh and Riedel 1997). Initial grain radius \( R_0 = 1.05 \mu \text{m} \); melting of the sintering aids: start temperature \( T_m = 1430 \text{ K} \), melting rate above \( T_m \dot{g} = 0.000595 \text{ K}^{-1} \), maximum volume fraction of the melt \( g_{max} = 0.1 \); maximum solid density by rearrangement \( D_1 = 0.63 \); inverse viscosity of the melt: pre-exponential factor \( \frac{b \delta}{\eta}0 = 1 \cdot 10^{-5} \text{ m}^3/\text{Ns} \), activation energy \( Q_\eta = 425 \text{ kJ/mol} \); reaction rate of solution/re-precipitation: pre-exponential factor \( L_0 = 3.9 \cdot 10^{-4}/\text{Ns} \), activation energy \( Q_L = 400 \text{ kJ/mol} \), diffusion through liquid phase \( \delta D_l = 5.2 \cdot 10^{-38} \text{ m}^2/\text{s} \), \( Q_D = 425 \text{ kJ/mol} \); solid/liquid interface energy \( \gamma_{sl} = 0.25 \text{ J/m}^2 \), liquid surface energy \( \gamma_l = 0.25 \text{ J/m}^2 \), volume fraction of large pores \( f_0 = 0.1 \), parameters for size distribution of large pores \( B = 0.12, n = 1 \); ratio shear-to-bulk viscosity \( G_v/K_v = 0.27 \) (theoretical value).
35.5 Sintering

35.5.4 Finite-Element Implementations and Applications

The described sintering models have been implemented in the finite-element code ABAQUS® (/Standard and /Explicit) in order to be able to calculate stresses and distortions in real components. In simple cases when distortions result primarily from inhomogeneities of the green density, the less sophisticated model of Riedel and Sun (1992) suffices to predict the distortions. If, however, distortions are caused by temperature gradients and external forces like gravity and friction, the advanced model of Riedel and Blug (2001) is needed. Such a coupled thermo-mechanical analysis was applied to the sintering of large molybdenum cylinders with density dependent properties and gravitational forces (Sun et al. 1997, Plankensteiner et al. 1999). McHugh and Riedel (1997) describe an implementation of the liquid-phase sintering model together with an application to the sintering of a hard metal cutting tool.

Figure 35.9: Evolution of the density during sintering.

Figure 35.10: Evolution of the axial and radial strains during sintering.
As an example of a sintering simulation Figure 35.11 shows the disc seal whose density distribution was shown in Figure 35.6. The upper part of the disc, since it has a relatively low density after pressing, shrinks by a greater amount than the lower part during sintering. This leads to the pronounced distortion shown in Figure 35.11. In the next step one would try to improve the density distribution and to minimize the distortion by changing the pressing schedule, i.e. the sequence of the punch movements. For the present disc it was not possible to obtain an acceptable result in this way. Then one would either change the tool concept completely or one would compensate the distortions by shaping the punches. This procedure was not carried out for this part, but it was demonstrated for another disc (Kraft et al. 2003).

![Figure 35.11: Sinter distortions of the disc seal resulting from density inhomogeneities of the green body.](image)

In several cases predicted sinter distortions were compared with measured ones, and the agreement is generally found to be good. A few examples for hard metal cutting tools (Sun and Riedel 1995, Riedel 1997, McHugh and Riedel 1997) and for a ceramic disc seal (Kraft et al. 2003) were published. Kraft et al. (2002) describe the distortion of tableware ceramics during sintering due to gravitational forces. The calculated distortions agree reasonably well with the measured ones.

### 35.6 Sizing and Post-Sintering Mechanical Densification

Sintered steels are often subjected to a sizing operation after sintering, which means that they are pressed in a die with the aim to improve the geometrical accuracy of the parts. Usually sintered steels contain porosities between typically 7 and 15%, which substantially impairs the ductility and fatigue strength of these materials. Hence in some cases it may be rewarding to exploit the sizing operation to densify the material specifically at critical locations to improve the fatigue performance and to increase the ductility. For different reasons sintered refractory metals also contain pores, which must be removed by swaging, forging or rolling.

Appropriate constitutive models for the simulation of post-sintering densification are those of Gologanu et al. (1994) and Kailasam et al. (2000). These models take the shape changes of the pores during forming operations into account, and thus describe the closure of the pores more realistically than, for example, the Gurson (1977) model, which considers the pores to remain spherical. Applications to forming operations on porous molybdenum are described by Parteder et al. (2002) and Plankensteiner et al. (2002).
35.7 Fatigue

To close the process chain one should be able to simulate the service behavior as a function of the prior processing conditions. Therefore, Kraft and Riedel (2003b) implemented a fatigue lifetime model into ABAQUS, which is based on the nucleation and growth of microcracks. The parameters of the fatigue model as well as the elastic-plastic properties depend on the porosity. As a demonstration example the authors track a powder metallurgical gear wheel through its processing and service history. The final result is a contour plot of the number of cycles to failure.

35.8 Conclusions

Simulation tools for the most important powder technological processing steps are now available. Pressing and sintering is routinely simulated even for complex shaped parts. Calculated and measured density distributions and sinter distortions are generally reasonably consistent. Cracking during powder compaction, unloading and ejection is an important practical problem. Attempts to model crack formation have been partly successful, but the predictive capabilities are not yet fully convincing. Besides the issues of improving the constitutive models, numerical problems are a limiting factor for the competitiveness of simulations. This is especially true when contact and friction play an important role, as in the case of die compaction. Implicit codes then often fail to converge. Explicit codes are less sensitive, but a complex problem usually still needs several attempts before a solution is obtained. Sintering is not critical, unless contact and friction on the support are to be modeled.

For other processing steps there are fewer practical examples, but simulation methods are being developed with the aim to model the complete process chain including the behavior in service.

References


References

Rottmann, G., Coube, O., Riedel, H., 2001. Comparison Between Triaxial Results and Models Prediction with Special


Introduction

A material of given chemistry can have very different properties. This is due to the fact that the properties of metallic materials depend on the spatial distribution of elements and crystal defects, i.e. its microstructure rather than on overall chemistry. This microstructure changes significantly during processing (Figure 36.1) and, therefore, minor changes in processing can result in conspicuous changes of material properties, from scrap to high performance alloys.

Consequently, the state variables of materials properties are the microstructural state variables and prediction of material performance requires knowledge of the entire processing history and the respective microstructure evolution. While this defines materials modeling on a microstructural, i.e. mesoscopic or even atomistic scale, material properties are defined and evaluated on a macroscopic engineering level. Materials’ modeling requires, therefore, interfacing the atomistic (physics) and macroscopic (engineering) world, i.e. the mesoscopic.
and macroscopic scale (Figure 36.2). This can be accomplished by a multiscale approach, where the integral behavior on a mesoscopic level is reflected by an average property on a macroscopic level.

![Figure 36.2: Interaction of mesoscopic microstructural models to macroscopic FEM (C = solute contents, L = particle spacing, M = Taylor factor, k = flow stress, ρ = dislocation density, ODF = orientation distribution function).](image)

Conventional materials' modeling on a macroscopic scale, therefore, tacitly assumes homogeneous processing conditions, i.e. constant behavior throughout the workpiece. This assumption is practically always violated during processing of a material due to local gradients of processing conditions and material constitution. Such gradients can be accounted for by spatial and temporal discretization, which requires, though, to calculate the material properties locally and to establish compatibility at interfaces. A long established and widely accepted engineering tool for this purpose is the finite element method (FEM). Traditionally the mechanical behavior of a finite element is based on an isotropic continuum. This is far from the truth, however, and advanced FE modeling requires actual materials properties to be taken into account for the behavior of an element.

Hence, more realistic materials modeling on engineering scale requires a link between mesoscopic microstructural modeling and FEM. This can be achieved either by using a very fine FE mesh to account for local variations of materials properties, e.g. accounting for local crystallographic slip, grain orientation and phase distribution. Such approach is partially realized in crystal plasticity FE (CPFEM) models that usually require long computing times and are restricted to relatively small volumes of a material (see Chapter 20, Becker and Panchedeeswaran 1995, Dawson et al. 1994, Dawson 2000, Hosford et al. 1993, Kalidindi et al. 1992, Kocks et al. 1998, Raabe and Becker 2000, Raabe 2001, Zhao et al. 2001). A computationally more efficient way is the use of an FE mesh size such that an element represents a homogeneous local average behavior of the material (Aretz et al. 2000, Dawson and Beaudoin 1998, Gottstein and Sebald 2000, Hiwatashi et al. 1998, Li et al. 2003). In the following we will demonstrate the latter approach by three examples, namely (a) through-process modeling of aluminum alloy AA2024 from solidification through homogenization and hot rolling,
(b) through process texture modeling of aluminum alloy AA5182 during industrial multistep hot rolling and cold rolling to annealing and finally (c) through-thickness texture evolution during warm rolling of an IF steel.

36.1 Through-process Modeling of Aluminum Alloy AA2024 from Solidification through Homogenization and Hot Rolling

The mould filling process of the ingot casting of AA2024 was simulated with the commercial FDM (finite differences method) software MAGMASOFT to obtain the initial temperature field, which serves as input for the solidification simulation. A two phase solidification model (Ludwig 2002) was developed based on the commercial CVM (control volume method) software FLUENT to simulate the macro structure considering both grain movement and convection. For the simulation of the microstructure a model (Pustal 2003) was developed and coupled to the FEM code CASTS (Neises 2001) and the thermodynamic library of Thermo-Calc (Sundmann 1985) to perform a multi-scale calculation. As input values the initial temperature and concentration fields had to be defined as well as a set of thermodynamic data of the alloy apart from its thermo-physical properties. The results of the simulations were distributions of the grain size and the phase fractions (Figure 36.3) as well as segregation profiles. They were forwarded to the subsequent simulation of the homogenization process of the cast ingot, where they served as essential input.

As a result of high cooling rates during the casting process the as-cast condition is far from thermodynamic equilibrium. During the homogenization of the as-cast ingot the existing phases may change their composition and their amount. The development of the number and size distribution of the precipitated phases was calculated by the statistical model ClaNG
Integration of Physically Based Materials Concepts

(classical nucleation and growth, Schneider et al. 2002), based on the classical nucleation and ripening theory. To calculate the chemical driving forces and equilibrium concentrations the commercial Gibbs energy minimizer ChemApp (GTT Technologies, Herzogenrath, Germany) and thermodynamic databases (COST 507) were embedded. After the above described FEM microstructure simulation of casting the ClaNG model could be run for each element of the FE mesh separately to account for inhomogeneous microstructures in the ingot. Due to the homogeneity in the present example only one ClaNG calculation was performed for the whole ingot. The results of this simulation were number and size distributions of second phase particles (Figure 36.4) and the chemical composition of the matrix phase. They were essential prerequisites for the subsequent simulation of hot rolling, because they have a lasting influence on the flow stress and its evolution.

For the simulation of the hot rolling of the homogenized AA2024 ingot the following interactive simulation scheme (Neumann et al. 2003, Crumbach et al. 2004) was applied, reflecting how microstructure evolution models can be implemented directly into FEM:

In the FE code LARSTRAN-SHAPE the flow stress evolution of each element was characterized by a model (3IVM model) based on dislocation densities as internal state variables (Roters et al. 2000, Goerdeler and Gottstein 2001). Before starting the simulation the model was tuned to and its predictions assessed for a field of flow curves of a binary alloy Al-4%Cu (Figure 36.5). With the data delivered by the ClaNG simulation of the homogenization of AA2024 the 3IVM model was capable of predicting flow curves of the ternary AA2024 alloy (Figure 36.5) by incorporating the effects of different solute contents and particle size distributions. In the simulation of the actual forming process (hot rolling with 40% thickness reduction at about 480°C) the FE model provided necessary input data for the 3IVM model – strain, strain rate and temperature locally in each element. But besides strain rate, temperature, solutes and particles the flow stress is additionally altered by the texture evolution during the
forming process. To account for this, the flow stress model 3IVM was linked to an FC-Taylor model, which also was integrated into the FEM and which calculated the texture evolution in each element according to the local displacement gradient tensor. With the respective local Taylor factor the 3IVM model calculated the local flow stress in each element. This interaction of FEM, 3IVM and FC-Taylor model was repeated for each incremental step of the whole forming simulation. Results of this simulation were rolling forces, torques and distributions of different quantities throughout the workpiece during the forming process (Figure 36.6), including local dislocation densities and local texture.

Figure 36.5: Flow curves of AA2024 at 400°C and different strain rates – predicted with 3IVM on basis of data on binary AlCu4.

Figure 36.6: Flow stress $k_f$, temperature $\theta$ and strain $\varepsilon_{VM}$ during hot rolling of AA2024 – predicted with LARSTRAN, interactively coupled to 3IVM and FC-Taylor model.
The texture evolution predicted by the FC-Taylor model for the FEM was sufficiently accurate to capture the influence on the flow stress. To gain more detailed information on the deformation texture evolution, and to simulate the development of the texture during a subsequent recrystallization annealing, more sophisticated texture models were applied subsequent to the FEM simulation. The deformation texture evolution during hot rolling was recalculated with the GIA (grain interaction) model (see Chapter 15, Crumbach et al. 2001a, Crumbach et al. 2001b). The temporal evolution of the displacement gradient tensor, temperature and flow stress was input to the GIA model. This evolution was tracked during the FE simulation for two specific elements – one located in midthickness ($S = 0$) and one located at the surface ($S = 1$). The resulting hot rolling textures compared well to the experimental results (Figure 36.5), and demonstrated the markedly different textures in midthickness and surface.

Figure 36.7: AA2024 experimental and GIA-modeled hot rolling textures in mid-thickness and surface of the sheet.

To simulate the recrystallization textures orientation dependent stored deformation energies and nucleation spectra were derived from the GIA model and a model that predicts in-grain orientation gradients (Crumbach et al. 2002, Crumbach et al. 2004). This data was input to the StaRT (statistical recrystallization texture) model (Sebald and Gottstein 2002, Goerdeler et al. 2002a). The model was run with the conditions that almost all nuclei stem from orientation gradients generated during deformation. A comparison of the modeled recrystallization textures with experiment (Figure 36.8) demonstrates good agreement and shows that the observed randomization of the textures ($f(g)_{max} \leq 2.3$) is well captured by the simulations.
36.2 Through-process Texture Modeling of Aluminum Alloy AA5182 during Industrial Multistep hot Rolling, Cold Rolling, and Annealing

The same simulation scheme as in example (a) was applied here, except that the initial conditions prior to the deformation were not simulated but measured. With a measured starting texture, solute content and particle size distributions the previously outlined interactive scheme of FEM (LARSTRAN-SHAPE), 3IVM model and FC-Taylor model was applied to industrial four step tandem hot rolling of AA5182 (Goerdeler et al. 2002a, Goerdeler et al. 2002b). The detailed rolling texture evolution was simulated for a midthickness element of the sheet. It was assumed that in-between the four passes no recrystallization took place, but that after the last pass the material recrystallized completely. For nucleation conditions of the recrystallization texture simulation the same analysis with the GIA model was performed as described above, i.e., most of the nuclei were assumed to stem from orientation gradients. In addition, many nuclei from cube bands, some random nuclei caused by PSN and shear bands and few nuclei from grain boundaries (Crumbach et al. 2002, Crumbach et al. 2004) were taken into account. The simulated hot band texture (StaRT model) and the measured texture of the industrial material are shown in Figure 36.9.

Figure 36.8: AA2024 experimental and modeled annealing textures in mid-thickness and surface of the sheet.
To trade the influence of deviations from the experimental textures on the simulation of further processing of the material, subsequent cold rolling texture development was simulated for the sheet center plane, where virtually ideal plane strain deformation can be assumed, which renders FEM simulation unnecessarily. The GIA texture model was applied incorporating a flow stress evolution predicted by the 3IVM model and the simulated hot band texture as input texture (Figure 36.9). Finally, the texture evolution during annealing of this cold band was addressed. Again, GIA provided the nucleation spectra and stored energies for the recrystallization texture model StaRT. Owing to the cold rolled initial microstructure most nuclei were assumed to have random orientations due to the high shear band density and nucleation at particles. In addition, nuclei from orientation gradients, from cube bands and from grain boundaries were taken into account as well. The simulated and measured texture evolution during this whole process chain is given in Figure 36.9 and reveals excellent agreement of the terminal texture. Despite all success, it is noted, however, that a serious deficiency of the simulation of recrystallization textures is still the intelligent quantitative selection of the individual nucleation mechanisms. This will be an important topic of texture research.

Figure 36.9: Through process texture simulation of AA5182 from tandem hot rolling to cold rolling and coil annealing.
36.3 Through-thickness Texture Evolution during Hot Rolling of an IF-Steel

For this exercise a similar but somewhat simpler simulation framework as the above described interactive coupling of FEM (LARSTRAN-SHAPE), 3IVM model and FC-Taylor was applied. Instead of the dislocation density model for the flow stress prediction an empirical Voce type hardening law was used in the FEM. For each deformation increment the textural changes in each element were accounted for in terms of the respective Taylor factor computed with an FC-Taylor model integrated into the FEM (Beckers et al. 1998). It was the aim of the simulation to match the experimentally observed through thickness texture gradient in an IF-steel sheet hot rolled to 40% thickness reduction.

![Experimental and simulated through thickness texture gradients of IF-steel sheet, hot rolled to a thickness reduction of 40% — Comparison of uncoupled FC-Taylor model taking the strain path from stand alone FEM results (different friction coefficients), and of interactive FEM ⇔ FC-Taylor simulation.](image)

This texture gradient, caused by the deformation geometry and rolling friction, ranged from a typical α-fiber rolling texture in mid-thickness ($S = 0$) to a pronounced shear texture at the sheet surface ($S = 1$, Figure 36.10). Typically, a layer of almost random texture extends between these extremes at thickness level $S = 0.4$, and the maximum of the shear texture was located somewhat below the sheet surface at $S = 0.8$.

Besides the experimental data, Figure 36.10 also shows a series of simulation results, in which the FC-Taylor model was not interactively coupled to the FEM. Rather only the displacement gradient tensor evolution provided by a stand alone FE calculation was input for
texture simulation, which neglected the influence of the textural changes on the flow stress. The intention of these simulations was to determine the hardly measurable Coulomb friction coefficient needed for the FE calculation. For a friction coefficient $\mu = 0.35$ best accordance of simulated and measured textures was found and this value was used for the interactive FEM⇔FC-Taylor simulation. Apparently, the influence of the textural changes on the flow stress is needed to predict all essential features of the through thickness texture gradient, namely the almost random texture at $S = 0.4$, and the shear texture maximum at $S = 0.8$ (Figure 36.10).

### 36.4 Conclusions

The interactive integration of mesoscopic physics based models covering the relevant microstructural changes with regards to microchemistry, work hardening, texture and recrystallization into a macroscopic process model like FEM provides a flexible system with high predictive power for very different materials and processes. A concept like this allows for predictions of materials behavior far off common processing routes because it is based on physical properties rather than on empirical process data.

However, there are still serious deficiencies in some of the microstructural models that limit the predictive power of such approach. In particular the nucleation of recrystallization is far from being understood, and quantitative prediction of recrystallization texture and microstructure requires additional information on nucleation density and kinetics up to now. Recent developments indicate, though, that more fundamental nucleation models will become available in the near future.

More complex processes or materials will require even more refined approaches, like multiphase microstructure evolution, e.g. stress induced phase transformations in tool steels. Nevertheless, macroscopic FE codes that provide spatially resolved parameters like temperature, stress, strain rate etc., in combination with embedded intelligent microstructure models have the potential to address also more complex microstructural systems.

### Acknowledgements

The authors gratefully acknowledge financial support by the Collaborative Research Center “SFB 370 – Integral Modeling of Materials”, Deutsche Forschungsgemeinschaft, and by the European Union through the VIR\[FAB\] project. They thank all coworkers in SFB 370 and the VIR\[*\] projects for excellent collaboration and fruitful discussions.

### References


References


37 Integrated Through-Process Modelling, by the Example of Al-Rolling

Kai F. Karhausen

37.1 Introduction

Modelling of the microstructural development during thermo-mechanical processing of aluminum alloys has been an area of research for a number of years. Considerable modelling knowledge exists and a substantial amount of models and software has been developed, covering a range of aspects of the process chain and a variety of microstructural mechanisms during thermal and/or mechanical loading of a number of aluminum alloys.

All these isolated solutions provide insight into important aspects of Al processing but generally do not allow for a quantitative prediction of properties after a complex industrial production schedule. The development of a true through process model (TPM) requires an interdisciplinary effort. Especially in aluminum production the final product properties are affected already at very early stages in the production chain, e.g. the content of elements in solid solution is dominated by the homogenizing treatment after casting and may directly influence the final sheet properties. Therefore, it is essential to be able to trace a number of appropriate state variables through the process chain, or at least through a number of relevant process stages in order to acquire quantitative predictions of product properties. In this context the linking of micro scale to macro scale models is required to feed realistic process information into the microstructural modules. Furthermore, the state of the material after one operation affects the material behavior in subsequent operations. Therefore, also the inter-linking of micro scale models is required to be able to feed realistic material histories into the microstructural modules. Once the microstructural development can be traced via variables of state, material property information can be derived from these via separate sub-models at any stage of the process chain.

Already at the very early stage of developing a concept for a TPM, usually based on already available models, it needs to be defined clearly what results are realistically to be expected from it. Encyclopedias describe a model as being an "Auxiliary presentation, occasionally also one idealized; simplifying description of physical processes in order to grasp the characteristics esteemed substantially". By applying this simple definition to the construction of a TPM we can conclude that this task does not primarily involve the generation of physical models which attempt to describe nature to utmost detail. In an early stage one has to define the substantial characteristics that should be covered by the model and exclude mechanisms that do not play a dominant role for the overall performance, or the target properties to be predicted. This choice may not be easy, since it involves the isolated quantification of the effects
of concurrently acting mechanisms. Often also well conducted experiments do not allow an isolation of the interacting mechanisms and sometimes only the linking of models itself can provide such information.

The “Through Process” aspect of the TPM places further focus on inter-disciplinary effort. Model development has to depart from the generation of ever more advanced and detailed models in specific fields towards the inter-linking of already existing models, accepting their known limitations at this stage. Only after the combined model is operative, it may become advisable to go back to the sub-models again and to improve the accuracy of factors that have been identified as being critical for the overall model performance.

37.2 Features of the Al Production Chain for Rolled Products

The conventional production chain for aluminum sheet processing sets out on DC casting of rolling ingots of drop weights up to 30 t. The most dominant metallurgical features of the cast ingot are the cell size as well as type, size and distribution of constituent particles and eutectic precipitates. Furthermore, the highly heterogeneous cooling rates impose variations in microstructure throughout the ingot. Due to grain refinement, a close to random texture is usually observed. Subsequently follows an ingot homogenization treatment, during which internal stresses are relieved, elements in supersaturated solid solution are precipitated, microsegregations are levelled to some extent by long range diffusion and primary phases are transformed. When the alloy is held at high temperature for a sufficient time, precipitates coarsen. It has been shown e.g. for alloy AA3003 (Doko et al. 1988), that the mean dispersoid size can be varied between 0.03 and 0.17 µm with a corresponding variation in dispersoid density between 740 and 11 µm$^{-3}$ for low and high temperature annealing respectively. As the dispersoid size strongly affects recrystallization during and after hot rolling, it needs to be controlled tightly. After homogenization the slab is usually cooled down to hot rolling temperature. This leads to a further reduction of elements in solid solution and a growth of coarser secondary particles. Hot rolling is performed at the highest possible temperature to lower flow stress in order to reduce loads and to achieve high reductions. An upper temperature limit is imposed by the need for controlling recrystallization and precipitation as well as by the occurrence of pick-up. During conventional rolling, the coarse as cast structure changes to a deformed and recrystallized structure, usually with considerably finer grain size. Coarse primary precipitates are crushed and distributed more evenly. Decreasing temperatures during rolling and an increase in vacancy concentration in the lattice due to deformation lead to further precipitation of secondary phases.

With respect to microstructural modelling, hot rolling can be seen as a succession of deformation and annealing steps. The most dominant metallurgical features are recrystallization and texture modifications which are interrelated and influenced by ongoing precipitation. In commercially pure alloys such as AA1050 elements with low solubility, like Fe and Si, interact with the grain boundaries and efficiently obstruct recrystallization when present as finely dispersed precipitates. The presence of Mn and Cr causes the same effect unless the elements are already bound in coarse particles. Within the Al-Mg-Mn alloy system, successive recryst-
tallization covers a broad range in the scope of industrial conditions and thus can be influenced to a large degree by the choice of rolling parameters.

Grain size and especially texture development are determined by recrystallization, which in turn govern the properties of products delivered in hot rolled condition as well as those of subsequently cold rolled sheets. Cold rolled products which receive further sheet forming operations, require a distinctive texture. A cube texture is desired in the hot rolled condition in order to reduce earing upon the drawing operations e.g. of a can manufacturer. The formation and evolution of cube components has received much attention of researchers and industry (Weiland and Hirsch 1991, Daaland et al. 1992, Maurice and Driver 1993). During hot rolling with little inter-stand recrystallization, the cube nuclei are formed at the expense of rolling texture components, and they can grow upon subsequent annealing. A coarse size and large inter-particle spacing are desired to enhance recrystallization during coiling. On the other hand, this may cause inter-stand recrystallization leading to the formation of large cube bands which deform unfavorably in subsequent passes. To restrain inter-stand recrystallization, the rolling temperature can be lowered, or the rolling speed can be increased. In this situation another restriction is introduced by the occurrence of particle-stimulated nucleation (PSN) (Engler et al. 1996). In addition, the importance of size and distribution of coarse constituent particles (Marshall 1996) emphasizes the need for understanding the development of microstructure throughout the whole process chain.

Hot rolling is the dominant process in the production chain of aluminum strips and sheets inducing microstructural heterogeneities in the end product. It is obvious, that discontinuous rolling on a reversing mill causes heterogeneities along the rolling direction. The delay time between head and tail continuously increases, so that for some passes the slab may partially recrystallize if no precautions are taken. Furthermore, especially in the roughing passes, the deformation is characterized by large gradients in strain rate and temperature in strip thickness. Heterogeneities in width may become severe, when operating at low speeds and if the mill is not equipped with selective cooling systems.

In the consecutive cold rolling operations, dislocation climb and cross slip are very much restricted. The material undergoes considerable strain hardening where fracture usually occurs before a steady state is reached. A defined strain is applied during cold rolling to achieve the desired product strength by work hardening, and to produce a plane sheet with well-defined and sometimes structured surfaces. An additional change in microstructure is caused by further crushing and alignment of constituent particles.

The strength of non-heat-treatable Al-Mg (Mn, Cu) alloys of the AA3xxx or AA5xxx groups is based on solid solution hardening and, to a lesser degree, on dispersoids. For these alloys, strain hardening during cold rolling may increase the strength to above 400 MPa. However, the strong interaction of Mg with dislocations may lead to problems with peak stress and serrated flow curves (Portevin - Le Châtelier effect) and resulting Lüders line formation. Measures can be taken to avoid this effect, but they generally result in a loss in strength. For this reason, age-hardenable alloys (AA2xxx, AA6xxx) are often utilized for panel applications, such as car bodies, where homogeneous surfaces are dominant product characteristics. The strength of these alloys is based on precipitation hardening, which is obtained after a solution treatment and subsequent ageing. In any case, the microstructural and textural characteristics of the hot rolled band determine the achievable product quality after cold rolling and heat treatments to a large extent (Hirsch 1997).
Texture determines the anisotropy of various sheet properties that are of crucial importance to most subsequent processing operations. For non-heat-treatable alloys, the cold rolling passes and annealing conditions need to be designed in order to meet the requirements on both strength and texture (Hirsch et al. 1995). Partially recrystallized hot rolled bands will show the typical rolling texture (β-fibre) that leads to ±45° earing. The dominant cube texture components, as present in the fully recrystallized hot rolled band, will result in 0/90° earing, while the RD-rotated cube texture component produces additional ears at 0/180°. Six-ear or eight-ear cup profiles reflect the intensity of different texture components. Subsequent cold rolling and eventual intermediate annealing procedures, which are adjusted to the condition of the hot rolled band, can be performed in such a way, that anisotropy is balanced, while at the same time work-hardening attains the desired strength level.

The layout and optimization of the above described interacting systems can only be performed if the microstructure can be quantified at all processing stages. Thus, predictive models are essential for this task and are subject of major research efforts in the aluminum industry.

### 37.3 TP Modelling of the Al Process Chain for Rolled Products

For the example of the first three stages of aluminum production a modular modelling system is shown schematically in Figure 37.1. It could be continued analogously for all following processing stages down to the final sheet properties. The interface between the models used in different production stages consists of a transfer of previous processing conditions, the field variables, and suitable variables of state. These information need to be provided in a space resolved manner. In addition an input of not modelled factors from measurements is fed into each sub-model. For a fully physically based model this block would be superfluous, but present models rely to some extent on empirical adjustments.

It needs to be emphasized that also material testing can be treated as a processing stage in this context since e.g. a tensile test can be modelled like any other forming operation. Material testing can be appended to each processing stage for validation purposes.

The microstructural development needs to be traced via suitable variables of state and properties are derived via separate sub-models at any stage of the process chain. The definition of the state variables which are passed over as interface information is of extraordinary importance. They rely to some extent on existing models and determine the need for further development of single sub-models. As a set of base variables of state the texture, the dislocation density defining material strength, the dislocation density defining the driving force for softening mechanisms, the grain size and the state of precipitation/solute level can be assumed. Of course both mentioned dislocation densities are in practice one state variable, they are mentioned separately for illustration purposes only. In Figure 37.2 the dependencies of these variables on the processes, the dependency of the processes on the state variables and the inter-dependencies of the state variables is shown. Obviously a number of additional dependencies does exist, but this illustration could be regarded as a definition of the minimum requirements for a TPM. With such coupled model system it should be possible to trace basic material characteristics and to derive property information on any processing stage, e.g. flow.
stress during forming and yield strength on material testing with significantly higher accuracy than today’s state of the art.

Finally, one challenge of TP-modelling is the level of integration of the different sub-models. In some cases an off-line coupling is sufficient, e.g. if recrystallization during annealing is modelled, the recrystallization itself has no impact on the thermal model and can be employed as a post-processing routine. On the other hand, a flow stress model interacts with the forming model and requires a full integration. Also do various micro-models interact where simultaneous reactions occur, e.g. the state of precipitation/solute level and nucleation of recrystallization are strongly inter-linked.

Ideally, if all models and corresponding sub-models are based on a mutual set of defined state variables, alloy development on the computer will be possible. Tracing a defined number of physically relevant parameters and variables of state will provide the means of determining the material condition through the different process stages and finally to derive the resulting material properties.

### 37.4 Application of Through Process Modelling

For process simulation of multi-pass rolling, thermomechanical codes are the core of a TPM (see Figure 37.3). Complementary models providing accurate boundary conditions such as
friction and heat transfer need to be incorporated here. Plastomechanical codes rely on information on the material behavior provided by means of constitutive models describing ongoing work hardening and softening, dependent on suitable variables of state of the material. Since purely physically based constitutive models are not available for complex alloy systems, generally they need to be determined semi-empirically. First a set of flow curves is measured for different combinations of temperature and strain rate e.g. by uniaxial compression, plane strain compression or channel die compression. Utmost care has to be taken to eliminate unwanted influences of the testing procedure such as friction on tools, heterogeneous strain fields and adiabatic heating during the test (Karhausen et al. 2002). The measured curves also need to be compensated for the prevailing texture in the samples. This is done by an external computation of texture development in the test and by correcting the flow curves with the computed polycrystal Taylor factor. The governing state variable for work hardening is the dislocation density. When using dislocation based constitutive models some constants such as activation energies for bulk and core diffusion, cross slip and climb need to be adopted to the specific alloy. This is done by non-linear regression techniques such that the level of computed dislocation density allows for an accurate description of the measured flow stress including work hardening and dynamic recovery. The magnitude of dislocation density after deformation serves as driving force for appended models of statical recrystallization. The accuracy of the combined models can be checked by metallographical analysis of the state of recrystallization in the compressed samples after a defined holding period at deformation temperature.

For final application to the industrial process, both work hardening and softening models are then implemented into the thermomechanical code and in addition the texture model needs to provide the influence of crystal orientation with respect to the mode of deformation.
37.4.1 Tracing of Dislocation Density

The constitutive law supplies the process models with the flow stress of a material which is
determined by the current state of the material in terms of microchemistry, dislocation density,
graın size and texture, according to Figure 37.2. A number of constitutive models, based on
internal state variables, have been proposed (Estrin et al. 1998, Roters et al. 2000, Nes 1998).
In the applications presented here the 4IVM (four internal variable model) is used, which is
described in detail in (Karhausen and Roters 2002).

Dislocation density based flow stress models are usually composed of two basic equations:

\[
\begin{align*}
\text{kinetic equation of state:} & \quad \sigma = f(\dot{\varepsilon}, \vartheta, S) \\
\text{evolution law for the state variable:} & \quad \dot{S} = g(\dot{\varepsilon}, \vartheta, S) \\
& \quad S = \text{Variable(s) of state}
\end{align*}
\] (37.1)

The 4IVM is a composite model which distinguishes cell walls and cell interiors of a
dislocation cell structure and uses the Orowan equation as kinetic equation of state for both
regions. The calculation of the external stress is largely analogous to the procedure as de-
d\varepsiloncribed in (Roters et al. 2000) with the exception that other state variables are chosen. In a
first step the shear stress in the respective areas is calculated as

\[
\tau_x = k_B \vartheta \left( c_1 \sqrt{\rho_{ex} + \rho_{sx}} + \frac{c_2}{L_{eff}} \right) \\
\text{asinh} \left[ c_3 \gamma_0 \left( \sqrt{\rho_{ex} + \rho_{sx}} + \frac{1}{L_{eff}} \right) \exp \left( \frac{Q_{glide}}{R \vartheta} \right) \right] + c_4 \sqrt{\rho_{ex} + \rho_{sx}}
\] (37.2)

where the subscript \( x \) denotes either \( w \) for the cell walls or \( i \) for the cell interiors. \( c_1 - c_4 \) are
constants, \( L_{eff} \) is the effective spacing of precipitates and all other symbols have their usual
meaning. Finally the externally required stress is calculated as weighted sum over the areas.

$$\sigma = M \left( f_w \tau_w + (1 - f_w) \tau_i \right)$$  \hspace{1cm} (37.3)

where $M$ is the polycrystal Taylor factor and $f_w$ is the volume fraction of the cell walls.

The model considers four different dislocation densities: immobile edge dislocations in the cell walls ($\rho_{ew}$), immobile screw dislocations in the cell walls ($\rho_{sw}$), immobile edge dislocations in the cell interiors ($\rho_{ei}$) and immobile screw dislocations in the cell interiors ($\rho_{si}$). The mobile dislocation density is treated as being constant. In contrast to some other flow stress models (Estrin et al. 1998, Nes 1998) additional microstructural features such as grain size, subgrain size, or misorientation between cells are not explicitly treated within the framework of the model. For each of the above mentioned dislocation densities an evolution law is formulated.

The evolution laws for the edge dislocations consist of a generation term and two annihilation terms which consider dislocation climb being controlled by bulk diffusion as well as core diffusion. The corresponding term for core diffusion is almost the same as for bulk diffusion only that it is proportional to the cube of the dislocation density (Prinz and Argon 1984).

$$\dot{\rho}_{ew} = \left( c_5 + \frac{c_6}{\tau_{ext}} \right) \dot{\gamma} - \left[ c_7 \exp \left( -\frac{Q_{bulk}}{k_B \theta} \right) + c_8 \exp \left( -\frac{Q_{core}}{k_B \theta} \right) \right] \frac{\tau_w}{k_B \theta} \rho_{ew}^2$$  \hspace{1cm} (37.4)

$$\dot{\rho}_{ei} = c_5 \dot{\gamma} - \left[ c_7 \exp \left( -\frac{Q_{bulk}}{k_B \theta} \right) + c_8 \exp \left( -\frac{Q_{core}}{k_B \theta} \right) \right] \frac{\tau_i}{k_B \theta} \rho_{ei}^2$$

The evolution laws for the screw dislocations are also composed of generation and annihilation terms. While the generation terms are essentially the same as for the edge dislocations the terms for annihilation are different, since the annihilation of screw dislocations is controlled by cross slip and not by climb.

$$\dot{\rho}_{sw} = \left( c_9 + \frac{c_{10}}{\tau_{ext}} \right) \dot{\gamma} - \frac{c_{11} \tau_w}{k_B \theta} \exp \left( -\frac{Q_{cross}}{k_B \theta} \right) \rho_{sw}^2$$  \hspace{1cm} (37.5)

$$\dot{\rho}_{si} = c_9 \dot{\gamma} - \frac{c_{11} \tau_i}{k_B \theta} \exp \left( -\frac{Q_{cross}}{k_B \theta} \right) \rho_{si}^2$$

In summary the model consists of five coupled differential equations which have to be solved in each time step of a simulation.

In order to compute the development of dislocation densities during a holding period after deformation the generation terms in the evolution laws are omitted, assuming that recovery kinetics in the static regime do not differ from the dynamic kinetics. In this way, static recovery is implemented and a subsequent forming operation will set on with reduced dislocation densities. Furthermore, the onset and progress of recrystallization is determined by the dislocation density as driving force while simultaneously the total dislocation density is diminished by static recovery and recrystallization. The kinetics of recrystallization are formulated as a function of the total dislocation density $\rho_{tot}$ which is given by the accumulation of the four
computed components, weighted by the volume fraction of dislocations in the cell wall $f_w$. 

$$\rho_{tot} = (\rho_{ew} + \rho_{sw}) \cdot f_w + (\rho_{ei} + \rho_{si}) \cdot (1 - f_w)$$  \hspace{1cm} (37.6)

The model assumes site saturation where the number of nuclei $n$ is proportional to some power of the total dislocation density at the end of the preceding forming step 

$$n = a \cdot \rho_{tot}(t_p = 0)^b$$  \hspace{1cm} (37.7)

where $t_p$ is the time in-between deformation cycles and $a$ and $b$ are constants. The grain boundary motion is modelled as thermally activated and with the grain boundary velocity being proportional to some power of the actual value of $\rho_{tot}$ 

$$v = c \cdot \rho_{tot}(t_p)^d \cdot \exp\left(-\frac{Q_{GB}}{R\vartheta}\right)$$  \hspace{1cm} (37.8)

where $Q_{GB}$ is the activation energy for grain boundary motion and $c$ and $d$ are constants. The newly recrystallized volume fraction $\Delta X$ in a time increment $\Delta t = t_2 - t_1$ can then be calculated as 

$$\Delta X = (1 - X) \cdot f \cdot \rho_{tot}(t = 0)^b \cdot \left\{ r(t_1) + c \cdot \rho_{tot}(t_1)^d \cdot \exp\left(-\frac{Q_{GB}}{R\vartheta}\right) \Delta t \right\}$$  \hspace{1cm} (37.9)

where $r$ is the radius of the nuclei and $f$ a constant.

In case of partial recrystallization in-between passes, fractions of the volume will have experienced drastic reductions of dislocation density due to the growth of recrystallization nuclei into the work-hardened material, while others still follow recovery kinetics. Thus, recrystallized and work-hardened fractions have to be treated separately in the model. As a result, the complete constitutive model requires a rather complex sub-structure management. Considering that during multiple-stand deformation recrystallized volume fractions may be generated at different process stages, all generations of sub-structures will have to be followed separately through the process chain. Furthermore, the application within plasto-mechanical codes requires the simultaneous tracing of sub-structures for sometimes large numbers of local volume elements. The above dislocation based work-hardening and softening routines are implemented in a procedure, which takes care of the sub-structure and element management as described in more detail in (Karhausen and Kopp 1992, Karhausen 1995).

To illustrate the behavior of the integrated model, alloy AA3104 was rolled at 460°C with a reduction from 30 to 14 mm at a velocity of 0.2 m/s. It must be noted that under conventional laboratory conditions extreme gradients in temperature and strain rate occur through thickness due to the low roll velocity, small roll radii and low work roll temperatures which do not reach a steady state as in industrial practice.

By the example of this hot rolling pass the development of computed total dislocation density is illustrated in Figure 37.4. Initially strong work hardening takes place in the pass while counteracting dynamic recovery and decreasing strain rate lead to a steady state or even drop in dislocation density towards the exit of the roll gap. From there on static recovery
37.4 Application of Through Process Modelling

Figure 37.4: Computed dislocation density of a laboratory hot rolling pass in different thickness positions \( s = \frac{y}{h_0}; \ s = 0: \) center, \( s = 1: \) surface.

prevails, which leads to a further decrease of dislocation density. Flow stress is determined by the dislocation density, which also presents the driving force for static recrystallization during a holding period. Both are shown in Figure 37.5.
The spread of dislocation density across thickness causes a rather wide spread in flow stress from 60 MPa in the center to 95 MPa on the surface. Due to these gradients, the plastomechanical code computes a spread in strain rate across thickness as well as gradients in adiabatic heating. These in turn influence the further development of dislocation density. Therefore, all these variables are heavily inter-linked. The plotted flow stress after the pass corresponds to a value that would be measured at the given temperature but at close to zero strain rate.

Recrystallization kinetics and dislocation density are strongly related. The effects are visualized by a further calculation where two passes of 50% reduction are simulated with an inter-stand time of 10s. Furthermore two variations of initial temperature are computed. The temperature paths (Figure 37.6a) document the quenching effect in laboratory rolling, when the aluminum surface contacts the cold work roll at rather low speed, leading to through thickness gradients of 107°C and 67°C for the first pass. Although this gradient is levelled out quickly after the pass, the consequence for microstructural development is evident in recrystallization kinetics (Figure 37.6b) and dislocation density (Figure 37.6c). In comparison to the slab center, static recrystallization kinetics are accelerated on the surface resulting from the deformation at lower temperature in combination with increased shear components in the strain tensor. Both factors increase the stored energy towards the pass exit, which is clearly visible in the increased level of computed dislocation density. During the inter-stand time recovery and recrystallization take place which decreases dislocation density. It can be observed that the initially high dislocation density on the surface is rapidly reduced by recrystallization and falls below the level of the center position before the next pass. The center line of the initially colder material shows a reduction purely caused by recovery, while on the surface partial recrystallization is responsible for an accelerated reduction in dislocation density.

In the application to industrial processes, with a number of rolling passes and holding periods in succession, the knowledge of the state of dislocation density is vital to trace material history and its influence on consecutive processes. An application is given in Figure 37.7 by the example of an industrial 4-stand tandem mill. It is evident, that inter-pass recrystallization strongly affects the initial conditions of a consecutive pass and thus the complete mill behavior as well as final properties are influenced.

A further application is given by the simulation of an annealing process after industrial cold rolling. Alloy AA5182 is annealed at 245°C. In Figure 37.8a the simulation results for an isothermal metal temperature are shown. Under these conditions, recrystallization commences rapidly and is completed after 3.5 hrs. Also the dislocation density, translated into yield strength, reaches the fully softened condition then and remains unchanged thereafter. In industrial production the material is commonly batch annealed on coil. When the furnace temperature is held isothermally at 245°C the metal temperature only gradually increases to the desired value, largely dependent on coil size, coil build up and furnace efficiency. Furthermore, different locations in the coil may experience varying thermal histories. The result of the simulation of such process is shown in Figure 37.8b for the hottest and coldest location in a coil. In this case 15 hrs are in fact required to achieve complete recrystallization. The slope of yield strength clearly reflects the initial recovery period followed by a further drop after recrystallization is initiated.
Figure 37.6: Computation of a two-pass laboratory rolling schedule. a) temperature, b) recrystallized volume fraction, c) dislocation density, \((s = y/h_0; \ s = 0: \text{center}, \ s = 1: \text{surface})\).
37.4.2 Tracing of Texture

Tracing of texture through a process chain requires deformation texture models to account for the change of crystal orientation due to mechanical loading as well as recrystallization texture models to account for changes of crystal orientation in inter-pass times.
37.4 Application of Through Process Modelling

Many attempts of modelling texture evolution during rolling with the help of Taylor type models are found in literature (e.g. Aernoudt et al. 1980, Hirsch and Lücke, 1988, Kocks et al. 1998). However, they fail in reproducing a sufficiently accurate rolling texture since conventional 1-point approaches do not take grain interaction into consideration. Thus, two grains with the same orientation will deform identically, whereas in reality they may behave differently due to interactions with surrounding grains. Therefore, rolling texture evolution is weakened and Taylor type models usually predict a too rapid texture evolution. More advanced n-point formulations, like the grain interaction (GIA) (Crumbach et al. 2001) or the LAMEL model (Van Houtte et al. 2002) yield more realistic rolling textures, yet the textures still develop too quickly. As recently shown by (Engler 2002), the deformation texture prediction can be further enhanced by allowing the individual components of the strain tensor to fluctuate randomly around the nominal plane strain state to account for local deviations.

In a multi-step production chain, a complete tracing of texture also has to include the development during recrystallization periods. Time discrete statistical-analytical approaches have been successfully applied (Engler 1999, Sebald and Gottstein 2001, Goerdeler et al. 2002), based on nucleation and nucleus growth. The nucleation models calculate orientation dependent numbers of nuclei within deformed grains with a given orientation. Thus, the orientation and misorientation distributions of the nuclei are determined. Thereafter, the growth rate of a nucleus is given by the grain boundary mobility and the local driving force. The grain boundary mobility depends on the misorientation between the growing grain and the consumed deformed grain. The driving force for static recrystallization is the difference between the stored energy of the deformed matrix and the nucleus. Information on the stored energy in the different matrix components are taken from the above mentioned deformation texture model.

A complete tracing of texture is desired since the grain orientation distribution does not only influence the final product by causing anisotropy in properties, at some stages also production processes are directly affected. This is especially true for final sheet forming operations but may also play a role at hot rolling already. Figure 37.9a shows the computed development of the Taylor coefficient during compression for two textures occurring in practice in the transfer slab before finish rolling. One is dominated by cube components and the other one is dominated by rolling texture components. The effect on flow stress is given in Figure 37.9b for a AA4xxx alloy. A difference of up to 15% is observed, which e.g. linearly influences mill load and thus related factors such as strip temperature, strip profile and flatness or work roll wear.

In order to illustrate the impact of texture evolution on final properties, the results of a complete simulation of hot and cold rolling for can body stock material (alloy AA3104) is shown in Figure 37.10 (Engler 2003). After hot rolling in four passes on a tandem mill, the material is self annealed on the coil and then cold rolled with a further reduction of 87% to meet the strength requirements. Thus, the sheets comprise a pronounced rolling texture which will, in a subsequent drawing operation, cause distinct earing in ±45° to the rolling direction. This 45° earing can be minimized by enforcing a sharp cube texture in the hot strip, since the quite slow rotation of cube oriented grains in cold rolling and the resultant mixed texture efficiently balances the earing profile. To illustrate the impact of hot strip on the final sheet properties, the texture development during hot rolling was simulated without inter-stand recrystallization and with complete inter-stand recrystallization. The simulation without
Integrated Through-Process Modelling, by the Example of Al-Rolling

Figure 37.9: a) Taylor coefficient during hot compression, b) Effect of texture on flow stress.

inter-stand recrystallization shows a cold rolling texture quite similar to the measured texture at final gauge. The simulation assuming full inter-stand recrystallization during hot rolling shows a different, weaker rolling texture with less retained cube orientation and a somewhat different intensity distribution along the rolling texture fibre. Obviously, such differences in rolling texture will impact the resulting earing profile of drawn can bodies.

37.4.3 Tracing of Microchemistry

The content of alloying elements, and especially the distribution of elements in solid solution or in precipitates and dispersoids is another example where TPM is a vital tool in order to quantify process behavior and mechanical product properties. The microchemistry of a given alloy is largely dominated by homogenization, but may also significantly be altered during
hot working or intermediate annealing. A sensitivity study (Löchte and Gottstein 2003) may already reveal which processing steps control final properties most efficiently. An alloy of high Mg content, such as AA5182 for can end or automotive applications, derives yield strengths of above 400 MPa largely from work hardening but it is also prone of loosing part of this strength during multi-pass rolling due to recovery at increased process temperatures. An alloy such as AA3103, e.g. for can body application, with lower Mg but increased Mn content possesses a lower overall strength in the range of 250 MPa but with a higher percentage effect of solid solution and dispersoid strengthening. This alloy is less sensitive to rolling conditions but homogenization needs to be controlled accurately.

The effect of alloy composition is documented by the simulation of a cold rolling schedule of AA5182 in Figure 37.11. The yield strength development in six cold rolling passes is computed, starting from hot band. In each pass yield strength increases due to work hardening, followed by a reduction due to recovery during coil cooling. The simulations show the strength variation when material of the extreme Mg and Mn content within the AA5182 specification is rolled. Obviously such variations take a recursive effect on the rolling processes (Figure 37.12). For the high alloy content the specific roll forces are significantly increased. Furthermore, since the higher Mg level adds to work hardening, the exit temperature is considerably increased by adiabatic heating. In order to reduce the risk of mill fires, high exit temperatures often cannot be tolerated and may have the practical consequence that mill speed needs to be reduced and productivity is brought down. Moreover, elevated exit temperatures lead to coil cooling periods at increased temperature levels, which promote recovery such that the initial gain in strength may be partially lost again. On the other hand, the increased Mn content counteracts this additional softening with the effect that the final yield strength in the product differs by 63 MPa for these variants. This system of interacting mechanisms is highly non-linear and illustrates the need for advanced TPM for process planning.
37.5 Conclusions

The key to a successful TPM is the early definition of suitable variables of state which need to be traced through the processing route. Three variables of state, namely dislocation density, texture and microchemistry have been discussed in detail by applications to laboratory and industrial processes. The interactions of the different state variables and the interaction of material and process are shown by example calculations.

The generation of a TPM demands from research to move from the creation of ever more advanced and detailed models of specific mechanisms towards the inter-linking of already existing models, accepting their known limitations at the present stage. Unfortunately, this task is often not seen as a field of research in academia, but is often considered a mere programming activity. However, through process modelling presents a range of challenges. Obviously, efforts have to be made in establishing the communication of sub-models which have been developed independently and which may come from different areas of research. The joint models have to run numerically stable, they have to converge and they have to run in acceptable computing times. Experts in a field have to develop some understanding of neighboring fields in order to be able to define interfaces and in order to evaluate which information needs to be exchanged. Equally important, it has to be decided at which level, in which frequency and in which spatial distribution these information need to be provided. This usually involves a number of sensitivity studies. And finally, the combined models compute complex interactions which are sometimes not easy to evaluate and to verify. Only at this later stage it may become advisable to go back to the sub-model level and to improve the accuracy of parameters that have shown to be critical for the overall model performance.
References


Daaland, O. et.al., Proc. 3rd Int. Conf. on Aluminium Alloys, Trondheim, Norway, 1992, p. 297.


38 Property Control in Production of Aluminum Sheet by Use of Simulation

Jürgen Hirsch, Kai F. Karhausen, and Olaf Engler

38.1 Introduction

Sheet production and processing operations are using simulation and modeling tools in an increasing range of applications, including the design and construction of modern rolling lines or the process control for load, flatness and profile. In recent years also the material itself and its properties have been integrated into rolling simulation models, since besides tight geometrical tolerances material producers are obliged to deliver material also within well specified property limits. Furthermore, high quality sheet production in any case mandates reliable quantitative material data, that may also be used as input to simulation methods for the process layout of subsequent customer operations (e.g. forming of automotive parts, prediction of crash behavior, etc.).

The new customer demands can only be met by a deep understanding of all aspects of processing in all stages during production and of material and microstructure features down to a microscopic scale. This entails a detailed knowledge about the response of the microstructure evolution during the specific processing steps, i.e. during deformation and heat treatment of the different alloys. In many cases the variation of final material properties can be traced back to very early stages in the production line. As an example, the pre-heating of the as-cast ingot already has significant influence on the microstructure evolution during subsequent thermo-mechanical processing like hot/cold rolling or extrusion and thus on final properties for many aluminum alloys.

The principles of microstructure evolution and the resulting properties during the industrial production of aluminum sheet are described and discussed for typical non-heat treatable aluminum-magnesium alloys, which are typically used in automotive applications and packaging. The metallurgical effects that occur at the various steps of the processing chain depend on alloy composition and process parameters (i.e. temperature, strain, strain rate and time). For the case of non-heat treatable Al-Mg-Mn alloys these effects may be quantitatively described with the help of models that incorporate strain hardening, dynamic and static recovery, recrystallization as well as solution/precipitation reactions occurring before, during, in-between and/or after pre-annealing, hot and cold rolling passes and various annealing treatments.

These effects need to be controlled in practice to meet the required high quality standards since they determine the main mechanical properties like strength, formability and anisotropy of the semi-finished sheet and thus also the material behavior of the final product. Microstructure simulation models can be generated from sound physical descriptions aided by special
laboratory experiments that provide the specific responses of material and microstructure on the related processes. They are integrated into thermomechanical simulations and used to predict and investigate the metallurgical processes involved during production of semi-finished products (sheets, extrusions, etc.) and during final processing. They can be used in modern industrial R&D to investigate metallurgical effects and analyze critical parameters of the various processes. They can be applied in practice for process control, design of equipment and material optimization strategies for high quality aluminum sheet.

### 38.2 Optimization Strategies in Sheet Processing and Material Quality

The first steps in sheet process development have been of purely empirical nature. The production methods of flat sheets by rolling were experimentally developed in the plants and – after testing of at best a few key material properties – the resulting sheet was then used as is. Any knowledge on processing and material properties was then developed step-by-step and obtained merely by trial and error. This method requires numerous plant trials under varying conditions, selecting the best combination of process parameters found. This procedure has to be repeated for each individual production equipment.

The traditional way of process optimization has the following major shortcomings:

- **Plant trials are expensive and time consuming.** They block production capacities which is most unpopular for high volume production facilities.
- **The results tend to be ambiguous due to the limited amount of trials.** The amount of data often does not allow a statistically relevant evaluation and sound conclusions.
- **Some decisive parameters in the process are not easily accessible to direct measurement.**
- **The results are limited to the specific equipment used and the specific alloys tested.** It is difficult to extrapolate to different or new equipment and new alloy variations.

Thus a more intelligent way is required, based on the sound knowledge and quantitative description of the thermomechanical process parameters and related metallurgical effects involved. With their integration into full simulation programs they are capable of resolving also more complex interactions (Hirsch et al. 1994; Wagner et al. 1998; Karhausen and Kopp 2002). For this purpose it is necessary to systematically analyze and understand all important material parameters under the different thermomechanical conditions that control the microstructure, affecting the resulting material properties. These effects need to be quantified – in a first approach still empirically, but, whenever possible, new algorithms based on physical principles are introduced and translated into suitable material models applied in practice, including potential on-line applications.

These models can be integrated into industrial process models which determine the local process conditions in a material and so can be used to simulate the corresponding process chain on a computer in all possible – or even imaginary – variations with the material properties as a result. Once established and verified, only a limited number of verification plant trials
needs to be performed in order to assess the performance of a specific equipment. Such an intelligent control for modern hot and cold rolling lines enhances performance and flexibility, since it allows a large variety of alloys to be processed individually and almost simultaneously according to their specific behavior under varying process conditions even in complex interdependent production schedules.

38.3 Processing and Microstructure Features of Aluminum Sheet

The conventional processing route for the production of Al-alloy sheet is depicted in Figure 38.1 together with the corresponding typical microstructures of non-heat treatable Al-Mg-Mn alloys. Production starts with a DC (Direct Chill) casting operation. As the overall efficiency of the processing route is to a large degree determined by batch size, it is usually aimed for the largest ingots that can be processed in the hot mill (e.g. in excess of 30 tons at AluNorf/Germany). The resulting ingots are characterized by a coarse cell size and constituent particles of eutectic precipitates of Al-Mn-Fe-Si phases. Furthermore, the local variations in the DC casting process impose certain variations in microstructure throughout the slab. Due to grain refinement, a close to random texture is usually observed. The ingots are scalped on their (later) rolling surfaces in order to remove surface blemishes. In preparation for the hot rolling, the ingots are then preheated to a temperature between $480^\circ$C and $600^\circ$C. Depending on the temperatures achieved internal stresses are relieved and elements in supersaturated solid solution (e.g. Mn, Fe, Si, Cr, . . .) are precipitated. Also micro-segregation is reduced to some extent by diffusion processes and primary phases are somewhat rounded in their form and sometimes transformed in their constitution. When the alloy is held at high temperature for a sufficient time, precipitates may coarsen. As the dispersoid size strongly affects final material properties, it needs to be controlled by optimum heat treatment conditions.

From homogenization temperature the slab is usually allowed to cool down to the hot rolling temperature. This may lead to a further reduction of elements in solid solution and growth of dispersed particles. The microstructure state of the slab before hot rolling can significantly influence the subsequent microstructure evolution by affecting recrystallization mechanisms and kinetics and thus also final properties in a number of alloys and product groups.

Conventional hot rolling is performed either on a single-stand reversing mill or on a roughing reversing mill followed by a multi (2-4) stand tandem finishing mill. Temperature needs to be adjusted in a way that high reductions can be achieved and recrystallization and precipitation will occur under controlled conditions. During conventional hot rolling, the coarse as-cast structure is transformed into a deformed or recrystallized structure of considerably finer grain size, mostly with a pronounced cube texture. The hard coarse constituent phases are crushed and aligned and evenly distributed – together with the fine dispersoids formed during homogenization. Decreasing temperatures during rolling and severe plastic deformation may lead to further precipitation reactions of phases from the over-saturated aluminum matrix.

The final cold rolling flattens the grains, strain-hardens the material (i.e. forms a typical dislocation/sub-grain structure), generates the typical fcc rolling texture and further crushes
and aligns the particles. Intermediate annealing treatments may be necessary between the rolling passes to reduce strain hardening especially of highly alloyed sheets. Besides alloy composition, the microstructure elements of grains (size, substructure and texture) and particles (constituents and dispersoids) are the principal factors controlling material performance in use and during most of the following forming operations.

### 38.4 Thermomechanical Simulation of Rolling Processes

Thermomechanical simulation of rolling processes for the automatic control of the pass performance is a common feature in modern, high speed hot and cold rolling equipment. An essential feature is the correct description of material flow stress which – besides friction and heat transfer – determines rolling forces and torque. Correct information on material strength under various forming situations is furthermore essential to compute correct temperatures which change due to adiabatic heating in the passes. Thus errors in strength tend to accumulate during the process simulation of several consecutive passes. Since the value of flow stress essentially represents an integral value over a number of intrinsic material properties, it is a key parameter connecting plastomechanical and material models, requiring full coupling.

Flow stress is classically determined in basic material tests such as uniaxial compression. A set of measured data is shown in Figure 38.2b. This testing method has proven to be adequate since it allows homogeneous deformation, if lubrication accomplishes full sliding between workpiece and tools. Under such conditions of constant and homogeneous strain rates
and temperatures a constitutive model can be derived by relating flow stress to the nominal compression conditions. The homogeneous temperature rise in the sample due to adiabatic heating can be corrected by iterative procedures (Karhausen 1995). Metallurgical analysis of the deformed samples reveals homogeneous microstructures and thus recrystallization kinetics can be quantified and empirical constitutive models can be adjusted accurately.

In the absence of dynamic recrystallization, the flow stress of aluminum alloys reaches a steady state value at elevated temperatures due to concurrent work hardening and dynamic recovery. However, at combinations of high strain rates and low temperatures flow stress does not saturate within the strain range of upsetting. For this reason a generalized Vöce equation, as proposed by Tomé et al. (1984), is most suitable for aluminum alloys:

\[
\sigma = \sigma_0(Z) + (\sigma_1(Z) + \Phi_1(Z) \cdot \dot{\varepsilon}) \cdot \left(1 - \exp\left(-\frac{\Phi_0(Z) \cdot \dot{\varepsilon}}{\sigma_1(Z)}\right)\right) \\
\text{with } Z = \dot{\varepsilon} \cdot \exp\left(\frac{Q}{R \cdot T}\right) \tag{38.1}
\]

\(Z\) (Zener Hollomon parameter). \(\sigma_0\) and \((\sigma_0, \sigma_1)\) symbolize the flow stress at the onset of straining and at the onset of constant \(d\sigma/d\varepsilon\) respectively, and \(\Phi_0\) and \(\Phi_1\) are the corresponding strain hardening coefficients. Equation (38.1) can be used over the whole range of industrial hot and cold forming conditions. It allows to describe a saturation of flow stress at hot forming conditions as well as a steady work hardening with strain especially at lower temperatures as it occurs for Mg-containing alloys.
Figure 38.3: (a) Microstructure of a cylindrical sample (quarter section) inhomogeneously hot compressed at 1 s\(^{-1}\) at 350°C, average strain of 0.8, partially recrystallized (after 10 s) and (b) corresponding finite element method simulation (local strain variation ranging between 0 and 1.54).

For many materials, close to full sliding conditions can be achieved in uniaxial compression (Kopp et al. 1993), by applying suitable lubrication on so-called Rastegaev specimens (cylindrical specimen with a defined groove milled into the face of the cylinder). This, however, is not always the case for Al-alloys at temperatures above 300°C. Figure 38.3a shows the microstructure of a Rastegaev sample of AA3104 after inhomogeneous deformation due to severe surface friction. Strong gradients in static recrystallization are clearly visible, indicating an inhomogeneous state of stored energy immediately after deformation.

Sticking of the surface of Al-alloy samples to the tool cannot always be prevented. Due to the good formability and low relative strength of aluminum the tool surface imprints into the sheet which on deformation prevents sliding of asperities and crushed oxides. Direct metal contact occurs in which free cohesive forces cause strong adhesive binding to the steel tool
surface. However, when friction is known, such inhomogeneous tests can be of some advantage. Without any lubrication complete sticking prevails that can be simulated accordingly (Figure 38.3b). The quality of the prediction can be assessed from the barreling of the compressed specimen and also from the pattern of recrystallization (Figure 38.3a), which both agree well with the strain distribution determined by an finite element method simulation of the (high friction) compression deformation.

In order to derive suitable constitutive equations from inhomogeneous compression tests, the development of the local strain distribution during compression needs to be analyzed. It requires a recursive method to provide the data, which again uses the constitutive equation as input. Such inverse computational methods have been successfully applied to this problem (Karhausen et al. 1999). A multi-dimensional optimization of FEM calculations is performed in which the parameters of the generalized Voce equation (38.1) are adjusted by combinatorial algorithms in an iterative procedure until the finite element method code provides accurate force-displacement curves. Since the specimens are axially symmetric and sticking friction prevails, the evolving sample geometry is forced into a pre-described barrel shape. Therefore, the solution space with respect to an optimum constitutive equation is restricted and reasonable convergence of the solving routine is achieved.

38.5 Microstructure Evolution During hot Rolling

Besides flow stress data compression tests can also provide good insight into the microstructure response under well-defined test conditions (including holding times). Effects of local deformation, recovery and recrystallization can be analyzed by metallographic and other suitable methods (SEM, EBSD, ... ) by inspecting the compressed specimens. The evaluation of the above-mentioned inhomogeneous deformation patterns by iterative finite element method calculation does not only yield an improved flow stress evaluation, but additionally the corresponding variation in recrystallization kinetics and resulting microstructures can be analyzed within one sample. Holding the sample for a defined time at temperature after compression leads to local variations in the amount of recrystallization and resulting grain size that correspond to the local variations in strain, temperature and strain rate history.

In this way the grain structure evolution can be analyzed in uniaxially compressed samples (Figure 38.4) and the results can be transformed to hot rolled samples of comparable strain history by a von Mises strain transformation. However, the texture evolution, which also needs to be considered for a complete determination of flow stress data is not accurately simulated by these samples due to the different strain geometry. While in compression a $\langle 110 \rangle$ fiber texture is formed that recrystallizes into a $\langle 111 \rangle$ and $\langle 100 \rangle$ fiber texture, in rolling a more complex $\beta$-fiber texture is formed (Hirsch and Lücke, 1988) that recrystallizes into a dominant cube texture (Hirsch, 1986). In principle, these effects can be simulated with sophisticated simulation tools (Engler 2002, 2003), yet for most practical applications experimental data (from laboratory and plant rolling trials) have been applied and semi-empirical descriptions are mostly used.

Microstructure and texture evolution during hot rolling are controlled by the recrystallization processes involved, which critically depend on the four main process parameters temperature $\vartheta$, strain $\varepsilon$, strain rate $d\varepsilon/dt$, and inter-pass holding time $t_p$ (Figure 38.5a). Details
38.5 Microstructure Evolution During hot Rolling

Figure 38.4: Progress of recrystallization and resulting grain structure of AlMg5Mn after hot compression at various temperatures and quenching after holding times indicated (strain $\varepsilon = 1$, strain rate $d\varepsilon/dt \approx 50 \text{ s}^{-1}$).

Recrystallization kinetics can be obtained from the described laboratory hot forming tests, where the microstructure evolution is investigated as a function of these four main process parameters (Figure 38.5b).

Recrystallization kinetics are evaluated on a number of deformed samples held for various defined times at temperature and subsequently quenched (Figure 38.4). These experiments show the dependency of grain size on temperature, strain and strain rate. A quantitative description of the recrystallization kinetics and resulting microstructure evolution (grain size, texture) can be given by help of semi-empirical equations (Hirsch et al. 1994). Integrated into a computer model of the hot rolling process it can be applied to predict the corresponding microstructure evolution and the effects of variations in hot rolling parameters (Karhausen and Kopp 2002).

In most commercial aluminum alloys dynamic recrystallization does not occur. Only under very specific hot deformation conditions (high $\vartheta$, low $d\varepsilon/dt$) may grain boundary fluctua-
tions occur in specific alloys with high Mg content (Figure 38.5a). The fluctuations are caused by local grain boundary motions during hot deformation, resulting into geometrical effects of dynamically formed new grains. In most cases, however, genuine, or post-dynamic, recrystallization takes place as the main process occurring immediately after hot rolling or in-between hot rolling passes.

In modern tandem hot rolling lines the hot band is produced on connected rolling stands in steps of up to 50% reduction with increasing rolling speed so that towards the final passes high strains and strain rates can partially compensate the drastic decrease in temperature. Accordingly, mills can be run in a process window that ensures a desired amount of recrystallization on the coil (“self-annealing”) for most Al-alloys. The recrystallization characteristics (kinetics, texture changes) depend on the stored energy accumulated over the passes and on the temperature development during cooling on the coil. Hence, it is important to take the whole series of finishing passes into account which can be achieved best when integrated thermomechanical/microstructure simulation models are applied.

Figure 38.6 shows the simulated temperature evolution (and the experimental exit temperatures) of a four-stand tandem mill for two hot rolling schedules (“fast” and “slow”) in AA3104 together with the predicted amount of recrystallization after two hours of subsequent coil cooling. The results agree well with the observed differences in microstructure and texture obtained in corresponding plant trials, as shown in Figure 38.7. Due to the typical strain rate and temperature distribution in rolling, unrecrystallized structures are found predominantly in the center layer as indicated by elongated grains and the occurrence of a strong rolling texture. The somewhat slower recrystallization in the center is also represented by the simulation (Figure 38.6b).

Process control of hot tandem mills conventionally is performed by exit temperature measurements. Mill speed and pre-cooling regulations are adjusted accordingly. By this method,
38.5 Microstructure Evolution During hot Rolling

Figure 38.6: Simulation of two rolling trials for alloy AA3104. (a) Temperature development during four-stand tandem rolling, (b) subsequent recrystallization during 2 hours coil cooling.

Figure 38.7: Microstructure of industrially hot rolled AA3104 under different process conditions ($T_{hr}$: hot rolling temperature).
However, effects of the material history are neglected which may be considerable. As an example, Figure 38.8a shows a simulation of the effect of pre-cooling and rolling velocity on temperature profile in a four-stand tandem mill. Combinations of strong pre-cooling at high speeds and vice versa give rise to marked differences in rolling conditions, although the exit temperatures are similar. The simulation of the recrystallization behavior for these temperature histories shows that in one case a large amount of recrystallization occurs between the first two tandem passes (Figure 38.8b).

The importance of recrystallization during hot rolling is demonstrated in Figure 38.9, where the experimentally observed variations of texture and resulting anisotropy (in terms of the $0^\circ/90^\circ$ earing value) are plotted versus the recrystallized volume fraction at the entry of the second stand. In spite of the constant exit temperatures yielding fully recrystallized microstructures with very similar grain sizes, systematic texture variations occur. The corresponding anisotropy behavior of the cube recrystallization texture ($0^\circ/90^\circ$ earing) is important for the quality of the finish gauge sheet used e.g. for can production where a balanced earing behavior is required (see below). The parameter of recrystallized volume fraction contains all important information and was determined by coupling thermomechanical calculations of hot rolling with integrated recrystallization simulations (Karhausen 1995).

For all these reasons, high Mg containing alloys such as AA5182 that tend to rapidly recrystallize between hot rolling passes are usually characterized by a relatively low cube texture intensity. However, when process parameters are changed in a way to reduce inter-stand recrystallization (high speed and low entry temperatures) also here a strong cube texture can be generated. Vice versa, when slow recrystallizing materials (by additions of Mn or Cr) are produced in single-stand reversing hot rolling mills with long inter-pass times, the material may recrystallize between passes and, therewith, generate a weak cube texture.

**Figure 38.8:** Simulation of a four-stand tandem mill for two different operating conditions. Development of (a) temperature and (b) recrystallized volume fraction.
Figure 38.9: Earing and texture evolution of AlMg5Mn in a tandem hot rolling line as a function of inter-stand recrystallization after the 1st rolling pass.

have revealed that a strict temperature control must be applied to avoid inter-stand recrystallization which might require additional coil annealing in order to produce a sufficiently high cube texture required for good earing properties of the finish gauge sheet (Hirsch and Grenz 1998).

In most Al-alloys a cube texture is obtained after hot rolling and complete recrystallization (Figure 38.7) with specific variations as a function of temperature (e.g. specific rotations at high $\vartheta$). However, for each alloy and processing route there is an optimum process window where maximum cube intensity is obtained. As addressed earlier, at too high temperatures and/or low strain rates frequent inter-pass recrystallization reduces the amount of stored energy and the cube nucleation bands. Vice versa, at too low temperatures (and high strain rates) new recrystallization grains can form with more random orientations by the mechanism of particle stimulated nucleation, PSN (Engler et al. 1996; Marshall 1996). In cold rolled sheets PSN recrystallization is a common feature, because industrially processed alloys always contain sufficiently large particles. However, since PSN depends on the ability to pile up dislocations at particles, it is strongly reduced by recovery during thermomechanical processing, i.e. when the material is deformed at high $\vartheta$ and low $d\varepsilon/dt$. Nonetheless PSN can occur in hot rolling lines for thin gauge hot band. It has a strong randomization effect on recrystallization texture, i.e., it may drastically decrease cube texture sharpness. Furthermore, at very low temperatures ($<300^\circ$C) recrystallization cannot be completed and deformation texture remains, which also strongly affects properties like yield strength and earing.
38.6 Material Properties of Industrially Processed Aluminum Sheet

At sufficiently high coiling temperature and time after hot rolling a fully recrystallized grain structure is usually generated by self-annealing, i.e. a fully soft (O-temper) condition is achieved. During subsequent cold rolling strain hardening significantly increases flow stress and reduces formability, as can be seen in Figure 38.10a. Furthermore, the change from cube to a rolling texture implies a diametrical anisotropy behavior, which can be noted in deep drawn cups (Figure 38.10b, c).

Yield strength evolution of different materials is shown in Figure 38.11a as measured by tensile tests performed on cold rolled material for a number of different rolling strains and alloys (Wagner et al. 1998). The tensile data can be represented by the generalized Voce equation given in Equation (38.1). Also, a dependence of the Voce parameters on alloy composition can be formulated, which allows interpolating such data for different alloys.

After strain hardening and subsequent annealing the material undergoes recovery due to annihilation of dislocations and dislocation re-arrangement; thus, strength is lowered. Figure 38.11b shows the effect of recovery on yield strength during annealing. In the logarithmic plot a linear behavior can be observed and recovery can be modeled on an empirical level by the following normalized form:

\[
R = \frac{\sigma(t, \vartheta)}{\sigma_0} = 1 - a \cdot \ln(1 + b \cdot \tau) \quad \tau = \int_1^t dt \cdot \exp \left( -\frac{Q}{R \cdot \vartheta(t)} \right)
\]  

(38.2)

However, different alloys and pre-treatments require different factors \( a \) and \( b \).
Figure 38.11: Modeling of (a) strain hardening for various Al-Mg alloys and (b) recovery in AlMg5Mn for different annealing temperatures.
An important practical application is the correct prediction of strength evolution of Al alloys under plant rolling conditions. Here always elevated temperatures (>100°C) occur during and after cold rolling due to the high reduction and speed of industrial rolling mills. So in multi-pass cold rolling sequences the coil requires cooling periods to avoid too high temperature levels in subsequent passes. As a consequence, the coil is stored for several hours before the subsequent rolling pass while the temperature generated by the previous pass gradually decreases and recovery takes place to some extent, depending on alloy, constitution, exit temperature, cooling conditions and previous pass reductions.

For example on a two-stand tandem mill the succession of two rapid passes with short intermediate time, followed by a long coil storage period before the material is brought to the next passes, bestows complex conditions where the strength of the alloy and the processing conditions strongly interact. Figure 38.12 shows the strength prediction of three double-passes on such a mill. It is demonstrated that at 50% reduced rolling speed the final yield strength is increased by 20 MPa as a result of increased cooling periods in-between stands and lower coil storage temperatures.

### 38.7 Simulation of Anisotropic Sheet Properties

Besides mechanical strength, in many applications, formability constitutes an important aspect of sheet material. Here, simulation is a well-established tool to analyze forming behavior and predict critical parameters, like local strains and failure (e.g. Banabic and Wagner 2002). The
texture introduced into a sheet during the thermomechanical processing affects its anisotropy properties like strength anisotropy (variation in strength in different directions of the sheet), or forming behavior (earing during deep drawing).

### 38.7.1 Strength Anisotropy

Figure 38.13a shows the anisotropy in $R_{p0.2}$ proof strength of AA 5182 can end stock under various angles $\alpha$ to the sheet rolling direction, RD. To sustain the internal pressure against buckling, the minimum strength level – here under $\sim 45^\circ$ to the RD – must meet the strength requirements of the can closures. Can end stock is heavily rolled to strength H19 so as to fulfill the strength requirements. Accordingly, can end stock is characterized by a pronounced rolling texture (Figure 38.13b). The yield strength anisotropy can readily be assessed from the Taylor factor $M$, which is defined as the sum of slip contributions $\Sigma \Delta \gamma_i$ on the $i$ active slip systems necessary to fulfill deformation under a given external shape change, divided by the strain $\Delta \varepsilon$ (Bunge 1982). This value, scaled to the yield strength for $\alpha = 0^\circ$, gives an idea on the texture-based part of the yield strength anisotropy (Figure 38.13a).

### 38.7.2 Tensile Test and r-Value Simulation

Tensile tests for material characterization are usually carried out parallel to the rolling direction. However, the material response also depends on the straining direction. Texture and anisotropy must be considered in subsequent forming operations, so more detailed measurements in different directions are required. For that purpose, commonly the material response in $0^\circ$, $45^\circ$ and $90^\circ$ to RD is investigated which shows the principle differences such as the earing effect. Based on sophisticated deformation texture models like the GIA model introduced in Chapter 22 of this book and suited strain hardening models, tensile tests can be simulated and flow stress can be predicted quite accurately as function of strain and texture (Figure 38.14a, b).
Figure 38.14: Simulation of (a), (b) orientation-dependent yield stress and (c) Lankford parameter $r$.

Besides strength and elongation, tensile tests also provide information about the formability of a sheet in the form the Lankford parameter (normal anisotropy) $r$, which is a measure for the relation of strain in width and thickness. A high $r$-value is beneficial for the material response in deep drawing operations. With the help of tensile test simulations in different angles to the rolling direction and with different prescribed deformation variations the $r$-value can be predicted at the conditions with the energy minimum of deformation as the optimum material response. Results of these calculations agree well with experimental data (Figure 38.14c).

38.7.3 Earing During Cup Deep Drawing

In can production highly cold rolled states are used (‘H19’) to meet the strength requirements (>275 MPa). Thus a strong $\beta$-fiber rolling texture is always formed. Figure 38.15a shows the evolution of (symmetrized) earing profiles where always 45° earing tips are formed after high cold rolling reductions. 45° ears also prevail in hot band if recrystallization is incomplete. The four ears in 0°/90° direction of the fully recrystallized hot band undergo the most drastic transition during which the typical intermediate eight-ears (at 0°/90° and 45°) and six-ears (at 0°/180° and 45°) profiles may occur. The individual ear heights correspond to the intensities of the related texture components. The cube orientation forms ears at 0°/90°, the RD-rotated cube (down towards Goss) at 0°/180° and the $\beta$-fiber rolling texture orientations all near 45°.

Thus, the formation of the rolling texture and the resulting earing profiles strongly depend on the amount of cube texture in the initial hot band. Due to its slow rotation rate and its opposing 0°/90° earing characteristics the cube component balances the 45° ears up to sufficiently high rolling reductions. The range of minimal final earing percentage depends on the optimum combination of initial cube texture and rolling reduction, e.g. at about 85% for fully recrystallized AA3104 can stock (Hirsch et al. 1996). The evolution of earing profiles caused by the texture changes occurring during cold rolling consists of a general rotation of the initial orientations towards the final stable rolling texture components ($\beta$-fiber).
The influence of three initial (hot band) texture components can be distinguished:

- The already existing hot rolling $\beta$-fiber texture components. They stay in their stable position, causing only some sharpening of the $\beta$-fiber orientations and, therewith, an increase in $45^\circ$ earing.

- The random part of the initial texture. It rotates relatively quickly towards the stable rolling texture orientations ($\beta$-fiber), increasing their volume and contributing to the corresponding $45^\circ$ ears.

- The cube texture. Only it can form $0^\circ/90^\circ$ ears and with its relatively slow rotation rate efficiently balances the $45^\circ$ ears. Its rotation around the rolling direction (RD) causes a stronger decrease of $90^\circ$ ears and the resulting pronounced $0^\circ/180^\circ$ ears together with the evolving new $45^\circ$ ears form the characteristic “six-ear” profiles (above 75% reduction).

The amount of initial cube texture necessary to achieve minimum earing at highest strains is controlled by the hot rolling parameters. High hot rolling exit temperatures promote complete recrystallization on the coil (self-annealing), which is a prerequisite for the required strong cube hot band texture. With a balanced volume of the two main texture components (cube and $\beta$-fiber) a low earing can be achieved. This, however, is valid only in a limited range of rolling reductions. At small rolling reductions the $0^\circ/90^\circ$ cube texture ears still prevail, while after too high rolling reductions the rotating initial cube grains begin to meet the stable rolling texture components and, hence, contribute to a further increase in $45^\circ$ ears. Furthermore, the control of these effects needs correct adjustment with the requirements for a specific mechanical strength which also depends on the cold rolling strain (Figure 38.11a).
Due to its characteristic form and symmetry, earing profiles can be described quantitatively by (even) Fourier series expansion coefficients (Hirsch et al. 1996) in the form:

\[ h(\alpha) = \sum_{i} f_i \cdot \cos(i \cdot \alpha) \quad \text{expansion degree} \quad i = 0, 2, 4, \ldots, 14 \]  

(38.3)

Only very few coefficients are sufficient to quantitatively describe the entire profile. Furthermore, the complete earing profile, the corresponding quality data (Z%) and their transition during rolling can also be simulated quantitatively by the evolution of the expansion coefficients.

\[ f_i(\varphi) = k_{i,0} + k_{i,1} \cdot \varphi + k_{i,2} \cdot \varphi^2 + k_{i,3} \cdot \varphi^3 \]  

(38.4)

Earing profiles can also be simulated with the help of physically based models under certain simplifying assumptions with the help of polycrystal plasticity approaches (Engler and Kalz, 2004). Figure 38.15b shows the predicted earing profiles of a recrystallized hot band as function of cold rolling reduction up to 92%, simulated with a visco-plastic self consistent (VPSC) scheme (see Chapter 23). All details of the earing profiles are displayed with very good accuracy. Compared to phenomenological earing models, the polycrystal plasticity based approaches have the advantage to be applicable to all texture types without any additional regression efforts. Furthermore, physically-based earing models may operate with simulated rolling and/or recrystallization textures, which is important with a view to the demands for a full, coupled prediction of the evolution of microstructure, texture and the resulting formability during the thermomechanical processing of Al sheet products (Goerdeler et al. 2002; Engler 2002).

### 38.8 Formability of Aluminum Sheets

The conventional way to simulate sheet forming operations like deep drawing is to employ the finite element method (finite element method). In order to account for crystallographic texture, the material input to the FEM code – the so-called constitutive equations – must incorporate the anisotropic properties of the sheet; see, e.g., the recent review by Raabe et al. (2002). For this purpose, two methods are mainly used. In the first method, the plastic behavior is described by a phenomenological yield function. However, although such yield criteria can readily be implemented into FEM codes, the accuracy achieved with this method is generally not satisfactory. In the second approach, a polycrystal texture is associated with each element of the FE-mesh and texture, hardening and anisotropy are updated as deformation proceeds (Becker et al. 1993; Beaudoin et al. 1994; Zhao et al. 2001). Thus, the constitutive behavior is derived from polycrystal plasticity. However, these finite element method computations are extremely time consuming, which currently restrains their use from routine forming simulations.

In order to take advantage of both approaches it has been proposed to use an anisotropic yield locus, derived from the sheet texture, as an input to the finite element method code (Bacroix and Gilormini 1995; Li et al. 2001). Figure 38.16 shows the yield locus of AA 6181 car body sheet in the final solutionized T4 temper. The yield surface was derived from
Figure 38.16: Yield surface of solution annealed AA 6181 car body sheet, plotted in the $\sigma_1/\sigma_2$ plane.

the texture data with the VPSC formulation by probing the texture with equi-spaced strain rates contained in a given plane through the full five-dimensional stress space. Here, the conventional plane spanned by the two in-plane stresses $\sigma_1$ and $\sigma_2$ is used assuming that the through-thickness stress $\sigma_3$ is zero. Finally, the data are normalized with the critical resolved shear stress, $\tau_0$. The texture-derived yield surface fits very well with the experimental data points for uniaxial tensile tests in rolling and transverse direction as well as for different combinations of biaxial stretching (see Chapter 6).

38.9 Summary and Outlook

Microstructure and property control based on sophisticated models has developed into a new method for improvement in profitability, innovation and quality in modern production lines for Al-sheet. The quantitative description of deformation and recrystallization effects and their kinetics integrated into process-simulation models helps to assess critical variables and optimize process conditions and resulting final material properties. Many thermomechanical effects like inter-stand recrystallization in tandem mills can only be analyzed and controlled with the help of models since they are often complex functions of different process param-
eters (\(\vartheta, \varepsilon, d\varepsilon/dt, t\)) that determine microstructure and texture evolution. In order to obtain optimum and reproducible properties process models with integrated microstructure evolution must be applied. They can be used to aid investment planning and assess optimization strategies on existing plants to meet high quality standards and reduce the amount of costly empirical plant trials. Once established and validated, these models can be extrapolated for any prediction of imaginary, innovative rolling conditions and new mill equipment.

Thus modeling microstructure evolution during the whole production chain will revolutionize improvement strategies on quality and productivity of sheet material manufacturing. It is the most feasible way for an Al-sheet producer to meet today’s high quality demands of customers on tight property limits while keeping production and development costs on reasonable levels. Furthermore, interfacing the integrated process models to sheet forming models, typically applied by sheet customers, will open new perspectives for the utilization of aluminum sheet material and enable the user to test and qualify material prior to production trials.

References

39 Forging

Yvan Chastel and Roland Logé

39.1 Introduction

Forging spans a vast class of equipment, techniques and forming operations. From a metals processing and a modeling point of view, very similar features can be found in hammer or press forging, open die or closed die forging, roll forging, upsetting, and even in ring rolling, piercing or extrusion operations. A common goal of all these processes is to shape a work-piece by using tools. A succession of forming steps is often needed to obtain the desired final shape. The part can be hot, warm or cold worked, depending on the ductility of the material, on the effort required and on the desired final microstructure. Indeed, the thermal and mechanical conditions of a forging process can cause significant microstructure evolution, thus leading to specific grain flow, grain size distribution and metallurgy. In the end, when the metallurgy is properly controlled, forged products can meet high levels of strength, reliability and mechanical performance. The evolution of the microstructure during the process is also crucial since it is the prime reason for a good workability of the material in large strain.

Numerical simulation tools prove very useful when designing these forming operations. Finite element codes can solve complex thermal and mechanical boundary value problems such as forming operations. And they can now deal with very large strain when coupled with automatic 3D remeshing techniques. We focus in this section on methodologies which can be used to compute also microstructure evolutions in forming processes with large spatial heterogeneity of strain rate, strain and temperature fields. Forging is far from being a steady-state process and typically includes multi-stand operations. Therefore, the numerical models for forging must be able to deal with such variations in space and time (Karhausen and Roters 2002, Choi et al. 2002, Beynon and Sellars 1992). The microstructure of the metal changes significantly during thermal and mechanical paths in forging: dynamic or static recovery and recrystallization can take place (e.g. Roucoules et al. 2003, Shen et al. 1995). For metals such as steels, phase change takes place, sometimes during the forging stage itself. These evolutions need to be modeled when the final microstructure is to be optimized, or when the behavior of the metal must be described accurately during forging.

The finite element method (FEM) is also emphasized because it allows the modeling of microstructure evolution in two distinct but complementary ways. In a first classical way, each node is supposed to represent a material volume sufficiently large to be summarized by an Homogeneous Equivalent Medium (HEM). Laws for microstructure evolution of homogeneous media are then applied. This strategy is adequate for the modeling of a large piece of material, i.e. at the scale of the part being formed (e.g. Sommitsch et al. 2002, Kawano and Isogawa 2001, Miyake et al. 2001). A second approach may be taken at the smaller scale of the HEM.
It consists in meshing a (heterogeneous) microstructure, possibly of complex topology and geometry, for studying the detailed relationships between the microstructure and its evolution, or even the associated constitutive law (Thibaux 2001).

This chapter covers three situations where the microstructure must be considered when optimizing the process itself, or the final product properties. Case I refers to recrystallization phenomena in hot forging sequences, case II treats the problem of warm forming of two-phase steels, and case III illustrates the modelling of crystallographic texture evolution induced by the forming process.

39.2 Case I: Microstructure Evolution During Complex Hot Forging Sequences

The first example presents a study of multi-step hot forging of an austenitic stainless steel typically used for retaining rings (Tsukahara 1995, Sellars and Wadsworth 1990). Since these rings are to satisfy severe criteria, it is essential to control accurately the microstructure during the manufacturing sequences and in the final state. The rings are hot forged around 1000°C on a 3000 ton press and then cold worked under a 65000 ton press. The grain size is given by the hot forging conditions. The alloy is an austenitic stainless steel with a high content of manganese and chromium, called NMF18.

39.2.1 Equations for Microstructure Evolutions

Hot compression tests were performed to characterize the microstructure evolution, particularly the static recrystallization and grain growth kinetics. Cylindrical samples are heated up to the deformation temperature for one hour, then forged perpendicular to the longitudinal axis of the cylinders. Tests down to various different heights were performed. The range of test conditions was selected according to finite element simulations in plane strain with the code Forge2®, which allowed to determine the true strain and strain rates at each position within the part. Figure 39.1 shows a true strain contour plot of a hot forged sample. The samples were annealed at different temperatures and for different times. Four parameters were modified in order to study their influence on recrystallization kinetics and final grain size: annealing temperature (880–1180°C), true strain (0.02–0.7), deformation temperature (940–1180°C) and initial grain size (0.4–1.5 mm). For the process of interest the deformation rate was not considered as crucial, for small variations around 0.2 s⁻¹. The samples were then analyzed to determine grain size and recrystallized volume fractions. It was found that there was no dynamic recrystallization.

The usual Avrami-Sellars equations (Avrami 1939, Avrami 1940, Avrami 1941, Sellars and Whiteman 1979) allow to describe correctly the static recrystallization phenomena for this alloy. The best result was obtained with an Avrami exponent $k$ of 1.0:

$$X = 1 - \exp \left( - \ln 2 \left( \frac{t}{t_{0.5}} \right)^k \right),$$
where $X$ is the recrystallized volume fraction. The half recrystallization time is expressed as:

$$t_{0.5} = 7.0 \times 10^{-13} \varepsilon^{-2.1} D_0^{2.0} \exp\left(\frac{-85000}{R T_{\text{def}}}\right) \exp\left(\frac{434000}{R T_{\text{rex}}}\right)$$

where $T_{\text{def}}$ and $T_{\text{rex}}$ are the deformation temperature and annealing temperature respectively, in Kelvin, $D_0$ is the initial grain size expressed in mm, and $t_{0.5}$ is in seconds. The grain size obtained after 100% static recrystallization is well fit by:

$$D_{\text{rex}} = 6.4 \times 10^{-4} \varepsilon^{-0.5} D_0^{1.2} \exp\left(\frac{40000}{R T_{\text{def}}}\right)$$

where $D_{\text{rex}}$ and $D_0$ are in $\mu$m.

After recrystallization other slower but significant mechanism takes place: normal grain growth. Figure 39.2 shows grain growth in this alloy when heated at 1160°C. Ranges of annealing temperatures (980 – 1200°C) and times (15 minutes to 24 hours) were studied and an equation was successfully fitted for grain growth:

$$D^{3.3} - D_0^{3.3} = 2.4 \times 10^{20} \exp\left(\frac{-444000}{R T}\right) t$$

where $D_0$ is the initial grain size and $D$ is the mean grain size in microns, $t$ is the annealing time in seconds, $T$ the annealing temperature in Kelvin.
39.2.2 Integration into a Finite Element Code

Retaining rings cannot be hot forged in a single pass. They require complex forging sequences. As shown in Figure 39.1, the very first compression can quickly lead to heterogeneous strain, strain rate and temperature fields. In the large scale simulation of the process the workpiece is divided into material volumes which can match the finite element discretization. In subsequent heat treatments, part of a material volume could then recrystallize while the remaining part of this volume would not (Figure 39.3). In order to model the full sequence of forging, one can either average the grain size and the work hardening for the different material volumes after each pass, or keep track of the history of the heterogeneous nature of each material volume. This latter technique improves the accuracy in the prediction of both the final average grain size and the grain size standard deviation. Anan (Anan et al. 1992) introduced such a method to store an array of microstructural state for each material volume. Each cell of an array evolves with a different kinetics from that of the other cells of the array. After each forging step and interpass time, the kinetics laws are applied to each cell and the statistics of grains is then recalculated for the array according to the new populations of grains. No averaging is then needed at the level of the material volume and no homogeneous equivalent microstructure is calculated.

The microstructure evolution laws are typically obtained from isothermal experiments, such as the tests described in the previous section. However, the process leads to temperature evolution all along the forging passes and interpass times. Using the results of isothermal tests in order to calculate the microstructure evolution for any thermal path implies that an additivity rule can be applied. It is then assumed that a non isothermal path can be divided into a number of isothermal steps. Despite arguments against this assumption, additivity rules are used in recrystallization studies (Woldt 1992). The procedure uses a fictitious time method (Pumphrey and Jones 1948, Tsukahara 1995) to discretize the microstructure evolution equation. This means, using for instance the Avrami equation: at step i of the computation, the recrystallized volume fraction $X_i$ at time $t_i$ has been computed; the fictitious time $t^*_{i+1}$ is calculated as the total time needed to obtain a recrystallized volume fraction of $X_i$ at temperature $T_{i+1}$. In
order to calculate the next recrystallized volume fraction $X_{i+1}$, Avrami equation is used for temperature $T_{i+1}$ and time $(t^*_{i+1} + t_{i+1} - t_i)$. The reasoning is illustrated in Figure 39.4.

### 39.2.3 2D Simulation Results

The forging sequence used to test the finite element model with the microstructure laws built in is shown in Figure 39.5. The finite element code Forge2® was used to simulate the whole
sequence. The initial grain size is uniform ASTM −3. In order to model the microstructure evolution using Anan’s method, an array of 200 cells is defined for each node of the finite element. The mean grain size and the standard deviation are calculated over these 200 cells. The results of the simulation are plotted in Figure 39.7. Figure 39.8 shows 13 micrographs to compare with the calculation, in reference to the sketch of Figure 39.6. The computed mean grain size and standard deviation are close to the experimental values. In the center of the workpiece a significant grain refinement is observed, from −3 to 3 ASTM. The standard deviation is biggest in a transition zone. This shows how important it could be to rotate the part and optimize the sequences in order to obtain homogeneous grain size distribution in the whole ring.

Figure 39.5: Temperature versus time graph describing one of the tests performed to validate the recrystallization model.

Figure 39.6: One quarter of the transverse section of the cylinder and locations of microscopic analyses.
39.2.4 Extension to 3D Forging and Dynamic Recrystallization

Forging is inherently a 3D process, with the ability of producing very complex shapes under large deformations. The strategy defined above, for the modelling of recrystallization phenomena, has been used with the 3D Finite Element Code Forge3R as well, and with equal success thanks to its reliable automatic remeshing technique. Figure 39.9 illustrates the case of a 316L austenitic stainless steel part deformed at 1273 K and annealed at 1373 K for 50 seconds. The simulation allows to take into account both the strain and temperature variations throughout the part, leading to a complex map of (statically) recrystallized volume fractions, ranging from 10 to 100%.

While the presented 2D and 3D results concern forming operations without the occurrence of dynamic recrystallization, the adopted strategy can also be used to calculate the concurrent contributions of dynamic and static phenomena. When treating dynamic recrystallization, the same Avrami equation usually applies, with the concept of time replaced by the concept of strain. The half recrystallization time $t_{0.5}$ becomes the half recrystallization strain $\varepsilon_{0.5}$, which generally shows a dependency with grain size and the Zener-Hollomon parameter $Z$:

$$Z = \dot{\varepsilon} \exp\left(\frac{Q_{\text{def}}}{RT}\right),$$

where $Q_{\text{def}}$ is an activation energy for hot deformation. In a forging process, however, the local conditions of temperature and strain rate vary strongly with time. It is therefore useful to consider small deformation steps, for which these local conditions can be considered as constant at any point of the mesh. The fictitious time technique is then replaced by a fictitious strain technique, dealing not only with evolving temperatures but also with evolving strain rates (see Figure 39.4).

39.3 Case II: Warm Forming of Two-Phase Steels

The temperature range of warm forming of steels runs from above room temperature to below the recrystallization temperature. Compared to a cold operation, warm forming has potential advantages such as reduced tooling and press loads, increased material ductility, elimination
of annealing prior to forging and favorable as-forged properties. However, phase changes usually take place in this temperature range and may cause flow instabilities. Moreover these phenomena are very sensitive to temperature-time diagrams, so that the forming operations need to be well scheduled and temperature controlled.

This second example treats the problem of plastic deformation of a steel during phase transformation (Thibaux 2001). An austenite and ferrite composite is in fact a two-phase HEM which needs to be analyzed in order to determine the mechanical behavior of the steel and the strain distribution within each phase. Analytical bounds and estimates for the mechanical behavior of a two-phase aggregate are based on strong topological assumptions. Another large field of research is the extension of the linear homogenization models to non-linear materials such as warm or hot metals. Bounds have been developed for such materials but many predictions from extended models are beyond the bounds. Only the estimates of Ponte Castaneda extension (Ponte Castaneda 1991) remain between these bounds (Gilormini 1997). The calculation of the mechanical response of these composite materials can be calculated.
39.3 Case II: Warm Forming of Two-Phase Steels

Figure 39.9: Strain, temperature and recrystallized volume fraction maps of a forged 316L part.

using the finite element method (FEM) with a representative description of the microstructure, well posed boundary conditions and a powerful computer. The HEM must contain enough grains so that a larger number of grains inside the HEM would not change the global behavior. Predictions of such FEM simulations can be compared with analytical models.

An example of a two-phase topology is shown in Figure 39.10 with the corresponding 3D finite element discretization. The mesh is able to capture strong strain gradients at the interface between the two phases. A visco-plastic Norton-Hoff law is used to describe the non-linear behavior of each phase by relating the flow stress $\sigma$ to the strain rate $\dot{\varepsilon}$:

$$\sigma = \sqrt{3} K (\dot{\varepsilon})^{m-1} \dot{\varepsilon},$$

where $\dot{\varepsilon}$ is the Von Mises equivalent strain rate. The strain field is calculated for a ratio of 4 between the rigidity $K$ of the hard (austenite) and the soft (ferrite) phase, and a strain rate sensitivity of 0.2 for both phases (Figure 39.11).

The upper bound for the two-phase ensemble rigidity, $K_T$, is obtained by using the standard mixture rule (Taylor model). The Sachs model gives the lower bound. Variants of homogenization models can also be used and give rigidities which are of course lower than $K_T$. 
Figure 39.10: Austemite-ferrite topology and corresponding 3D finite element mesh.

The differences between the calculations are shown in Figure 39.12 for a number of homogenization models together with results of FEM simulations for material volumes such as the one presented in Figure 39.10.

Among several variants of homogenization models, the best model for this topology and this ratio of materials parameters is obtained using a self-consistent model with Ponte Castaneda extension to non-linear behavior (Figure 39.12). The classical secant approach is more rigid than the finite element results.

Once an appropriate analytical model is validated, it can be used as the constitutive law in large scale finite element simulations and coupled with the phase transformation diagram to calculate the local percentage of each phase. This numerical model is then able to calculate thermal, mechanical and metallurgical evolutions in forming processes inducing large strain such as forging or rolling. As an example comparisons of calculated and experimental warm rolling forces are shown in Figure 39.13 for a C-Mn steel.

39.4 Case III: Texture Evolution in an Hexagonal Alloy

The texture evolution of metals is of prime interest in forming operations because it largely dictates the mechanical anisotropy of the final products. Furthermore, establishing a relationship between texture and constitutive laws is crucial, because textured materials exhibit
anisotropic plastic flow in forming operations, with geometrical consequences when free surfaces exist.

Integrating the crystallographic textures into the modelling of forming operations requires to perform at least two length scale transitions. The first one relates the process boundary conditions to the local thermal and mechanical conditions of the material, supposed to be homogeneous. A second length scale transition then partitions the strain and stress fields among the various grains, which differ from each other by their crystallographic orientation, and therefore by a distinct “hardness” with respect to the local thermal and mechanical conditions. In fact, this length scale transition is similar to the one described for Case II, where the two phases of interest could simply represent two classes of crystallographic orientations. In general though, even in the presence of strong textures, an adequate description of the polycrystalline nature of the material requires a minimum of a few hundred classes of grains.

The simulation aspect of the texture evolution of low symmetry alloys has been largely discussed in the literature, because the usual Taylor-type assumptions of homogeneous deformations through the aggregate (see Chapter 22) is pertinent for cubic metals, but no longer describes accurately their behavior (Lebensohn and Tomé 1993, Chastel et al. 1993, Marin et al. 1995). In the example reported here, we deal with an alloy whose lattice symmetry is hexagonal close-packed (hcp).

Case III illustrates hot forging and extrusion operations on Zirconium alloys. In the temperature and strain rate range of interest, crystallographic slip is the principal mechanism of deformation (no twinning or grain boundary sliding). Like in the previous examples, we use
the FEM to perform a first length scale transition, with an element size compatible with uniform thermal and mechanical fields within one element. The second length scale transition is done assuming that each element is an aggregate made of $n$ grains, whose crystallographic orientations are sampled from the global measured initial texture (Toth and Van Houtte 1992). A self-consistent homogenization method (Lebensohn and Tomé 1993) is used for the calculation of the second length scale transition, providing the texture evolution and the associated constitutive law. This model has already been described in Chapter 23.

### 39.4.1 Calibrating the Polycrystalline Model with Simple Mechanical Tests

Uniaxial compression tests and torsion tests were performed in the 873–1023 K temperature interval to measure the texture evolution at various strains. Considering four possible slip systems (Logé et al. 2000), the values of slip system hardnnesses were chosen in such a way that the calculated textures gave the best match with experiment. Numerical tools based on an inverse method were developed for that purpose (Signorelli et al. 2000). They allow to identify parameters of the polycrystalline model using experimental data like the texture or the stress-strain curve. For high temperature experiments it is not recommended to use stress-strain curves, due to the presence of thermally activated phenomena which are not directly accounted for by the model.

Plane strain compression tests were also designed to check the influence of other microstructural aspects like the precipitate structure or recrystallization, on the slip system hardnnesses and the final textures (Logé et al. 2000).
39.4.2 Using the Texture-Induced Anisotropic Plastic Flow to Validate the FEM Results

Besides the experimental textures, other geometrical measurements can be made to validate the calculation, this time in the context of FEM simulations where evolving constitutive laws are calculated at each integration point with the polycrystalline model. In Figure 39.14, cylindrical specimens were compressed, and the cross section evolved towards an elliptic shape, as a consequence of the non-random initial texture and its evolution, leading to local and global anisotropic plastic flow.

The “ovaling ratio” \( r \), defined by \( \frac{D-d}{d} \) where \( D \) and \( d \) are the elliptic long and short axes respectively, can serve as a good indicator of the accuracy of the FEM calculation. However, considerable care should be taken regarding the reproducibility of the friction conditions in the experiments. Indeed, Figure 39.15 shows that \( r \) is very sensitive to the friction coefficient, with a final value increasing from about 12% to 28%, when the friction coefficient is reduced from a typical value at high temperature, to zero.

39.4.3 Application to Hot Forming

Once the set of simple tests have been validated, more complex forming operations can be simulated based on the identified values of slip system hardesses and friction coefficients. Other metallurgical parameters can have an impact on the simulations, like the hardening.
coefficients which may vary from one slip system to another, the inhomogeneity of precipitate structure, recrystallization/recovery phenomena, etc. While the first item is usually not crucial in hot forming conditions, the last two items may have a significant influence on the final textures. They have been taken into account to model the evolution of texture during the hot extrusion of tubes. The comparison between predictions and experiments is given in Figure 39.16, at three different locations in the tube (Logé et al. 2000), with explicit consideration of the precipitate structure. The agreement in terms of final crystallographic textures is very good, because it was shown that the (0002) poles near the axial direction (AD) tend to disappear with the post-deformation (or metadynamic) recrystallization occurring directly after extrusion.

39.5 Conclusions

Three case studies have been presented in order to show the influence of microstructural features on hot forming operations like forging. These examples are clearly not exhaustive: their purpose is to illustrate the typical modelling challenges of a large class of problems dealing with the microstructure, at high or medium temperature.
While the concepts behind Cases I, II and III are usually applicable to any kind of hot forming, the particular features of forging impose specific ways of integrating well-known modelling aspects into the numerical computations:

- Forging is often a multi-step operation, with therefore a need to model both dynamic and static microstructure evolutions, in particular when looking at recrystallization. Keeping track of the local history of microstructure, within a specific data structure, consequently improves the results accuracy.

- Homogenization techniques are a central theme, because they allow to summarize the complexity of microstructures into simple constitutive laws, at an intermediate length scale between the scales of metallurgy and of the process. This was illustrated in Cases II and III, but similar strategies are needed in Case I, where recrystallization introduces softening (e.g. El Wahabi et al. 2003, Roucoules et al. 2003, Karhausen and Roters 2002) and possible strain localization. When applying the homogenization techniques to forging, the challenge is to obtain formulations which can integrate the inherent large variations in time and space of the mechanical and thermal fields.

- Forging means complex shapes, usually in 3D, and large strain. When using the FEM method, this implies automatic remeshing within a Lagrangian formulation, since the process cannot be described as a steady state, i.e. a Eulerian formulation is not applicable. This has important implications, because remeshing is usually done by transferring all the relevant variables, i.e. mechanical, thermal, and microstructural variables, from the old node positions to the new ones. Simple interpolations are usually performed. However, when microstructures are described by large arrays of variables (Anan’s method in recrystallization, crystallographic texture, etc.), new and reliable interpolation techniques must be developed.
Figure 39.16: Comparison between predicted and experimental crystallographic textures following hot extrusion, at three locations in the tube: (A) inner surface, (B) mid-thickness, and (C) outer surface. Contour lines represent multiples of the uniform distribution (mud). AD and TD refer to the Axial and Transverse Directions of the tube, respectively; the third direction in the center of the pole figure being the Radial Direction. Calculated poles near the AD are the ones disappearing during recrystallization.
References


Pumphrey, W.L., Jones, F.W., 1948. Inter-relation of hardenability and isothermal transformation data. ISIJ 159, 137-144.


References

40 Numerical Simulation of Solidification Structures During Fusion Welding

Vitali Pavlyk and Ulrich Dilthey

A numerical model based on a combination of the cellular automaton and finite difference methods was developed to simulate the formation of solidification morphologies under constrained solidification conditions, typical for welding. The model is based on solution of the sharp interface solidification problem. Solute partition between solid and liquid phase, solute diffusion in both liquid and solid, solutal and curvature undercoolings and the anisotropy of surface tension are considered. The model reproduces qualitatively the parabolic shape of the dendrite tip, development of secondary and ternary arms, instability of the planar interface, coarsening driven by curvature reduction and dendrite selection under constrained solidification conditions. The primary dendrite arm spacing and microsegregation patterns can be numerically investigated with the model. The model was applied to welding conditions. The solidification parameters, growth velocity and temperature gradient, were calculated taking into account the macroscopic fluid flow in the weld pool. The macro- and micro-modelling agree with each other and with experimental findings.

40.1 Introduction

The achievable mechanical and technological properties of weld joints are essentially affected by the microstructure, which forms in material during heating, melting, solidification and cooling. The solidification behavior of welds plays an important role in the whole chain of the microstructure evolution, which starts from inclusion formation already in the molten pool, goes through dendritic growth, grain formation, micro- and macro-segregations in course of solidification and continue by the solid state transformations at cooling (David and Babu, 1997). Especially, in austenitic stainless steels and aluminum alloys, the final microstructure is formed to a great extent during solidification, since the solid state transformations in this materials are basically restricted to the grain growth and precipititation. Also, in ferritic steels and other materials, in which the final microstructure is formed at a phase transition in solid ($\gamma \rightarrow \alpha$ transformation in iron alloys, for example), the primary solidification structure remains relevant, because it determines the staring conditions for the subsequent transformation Bhadeshia and Svensson (1993).

The importance of microstructure for mechanical and technological properties of weld joints and a great number of parameters, which influence its formation, facilitated the microstructure modelling to grow to a significant branch of the welding science (David and Vitek (1989), DebRoy and David (1995), Zacharia et al. (1995)). A traditional way to model...
the microstructure is based on phenomenological, analytical and/or statistical approach. A comprehensive description of this important methodology, which provides a deeper understanding of the microstructure assisting in the welding process optimization and control, is presented in the book of Grong (1997). Another, rapidly developing approach to model the microstructure in casting and other material processes can be thought as direct microstructure simulation. Any microstructure is associated with a steep gradient of some physical property at the boundary between different phases or different microstructure constituents. The boundary evolution in course of the microstructure formation is affected by temperature, concentration of reagents and, in some cases, also stress fields. Thus, in order to simulate the microstructure development directly, the evolution of this boundary in both space and time, coupled to the above physical fields has to be calculated. Either direct or implicit information about the boundary location allows to visualize and numerically investigate the process of the microstructure development. To these methods belong very popular phase-field (PF) (Steinbach et al., 1996; Steinbach and Schmitz, 1998; Tiaden and Grafe, 1999) and cellular-automata (CA) (Gandin and Rappaz, 1994; Dilthey et al., 1996, 1997a,b; Pavlyk and Dilthey, 1999; Gandin and Rappaz, 1997; Desbioles et al., 1998) techniques. Both of the above techniques operate on continuum scale from about 0.1\(\mu\)m to 1 mm, and can represent the microstructure features with characteristic length ranging from 1\(\mu\)m (dendrites, inclusions) to several centimeters (dendritic grains in ingots).

Besides the process relevant material properties and the alloy composition, the solidification behavior, the morphology of solidification front and the grain structure in fusion zone of weldments are determined, to a great extent, by the weld pool shape and the time dependent temperature field (David and Vitek, 1989). These in turn are influenced by a strong melt circulation in the molten pool, which contributes to the overall heat and mass transfer (DebRoy and David, 1995). Therefore, in order to perform a consistent microstructure simulation during welding, the temperature field and the weld pool geometry have to be calculated, considering coupled heat and fluid flow phenomena, i.e. the thermo-fluid problem. In view of this, a space-resolved simulation of solidification microstructure under welding conditions represents a multi-scale problem, in which the coupling between the macroscopic scale of the temperature and the microscopic scale of the microstructure has to be considered.

In the last two decades considerable progress has been made in modelling and numerical simulation of fluid flow and heat transfer during arc welding. Numerous models and a great number of numerical results have been published on this subject (see Jönsson et al. (1994) and Weckman (1999) for references). Although, a number of questions are still to be clarified, and some phenomena, like heat input and turbulence, require a more detailed investigation, basically, the physical models of driving forces and numerical methods for solution of the thermo-fluid problem of the weld pool are well developed (Weckman, 1999).

In this contribution, we aim to present some recent developments in the field of computer modelling of microstructure during weld pool solidification. We will consider the results obtained using a modified cellular automata method, coupled to numerical finite difference (FDM) and finite element (FEM) methods. For the numerical modelling of thermo-fluid phenomena general-purpose computational fluid dynamics (CFD) code FIDAP, which is commercially available, was used. The above simulation techniques and models were also described in (Dilthey et al., 1996, 1997a; Nestler et al., 1998; Pavlyk and Dilthey, 1999, 2001).
40.2 Modell of Dendrite Growth under Constrained Solidification Conditions

The physics of the developed solidification model is described by the sharp interface solidification problem for binary system (Kurz and Fisher, 2001). For the numerical solution of the problem under specific conditions of directional solidification, we suggested an algorithm, which is based on a combination of traditional finite difference method and the cellular-automaton technique. While the first of them belongs to well established computing methods, application of the second one for modelling various metallurgical phenomena has received extensive attention only during the last decade. The basics of this method is presented in Chapter 3. In the following, the essentials of the model and numerics are described.

40.2.1 Solidification Problem with the Sharp Interface

We consider the solidification of a binary alloy in a two-dimensional spatial region \( \Omega \) with the interface \( \Gamma(t) \) (Fig.40.1). The phase transition is governed by the evolution of the temperature \( T(x, y, t) \) and concentration \( C(x, y, t) \) fields which have to satisfy several conditions on the moving interface. The interface represented by a curve \( \Gamma(t) \) is not known a priori but has to be found as a part of the solution. The following equations describe the process (Kurz and Fisher, 2001):

Heat diffusion equation

\[
\rho c_v \frac{\partial T}{\partial t} = \nabla \cdot (\lambda \nabla T) \quad \text{in } \Omega \setminus \Gamma(t),
\]

Solute diffusion equation

\[
\frac{\partial C}{\partial t} = \nabla \cdot (D \nabla C) \quad \text{in } \Omega \setminus \Gamma(t),
\]

Heat conservation (release of the latent heat)

\[
\rho \Delta H \psi = \lambda_l \nabla T|_{\text{liquid}} \cdot \bar{n} - \lambda_s \nabla T|_{\text{solid}} \cdot \bar{n} \quad \text{on } \Gamma(t),
\]

Solute conservation

\[
(C_l - C_s) \psi = D_l \nabla C|_{\text{liquid}} \cdot \bar{n} - D_s \nabla C|_{\text{solid}} \cdot \bar{n} \quad \text{on } \Gamma(t),
\]
solute partitioning between liquid and solid phases

\[ C_s = kC_l \quad \text{on } \Gamma(t). \] (40.5)

Here \( \rho \) is the density, \( c_m \) is the specific heat per unit of mass, \( \lambda \) is the thermal conductivity, \( C \) is the solute concentration, \( D \) is the diffusion coefficient of solute, \( \Delta H \) is the enthalpy of melting, \( \vec{n} \) is the normal unit vector to the interface directed from solid to liquid, \( v \) is the normal interface velocity and \( k \) is the partition coefficient. The subscripts “l” and “s” denote the values of respective variables and coefficients in the liquid and the solid phase, respectively. The interface is suggested to be in a local equilibrium with both phases, the equilibrium temperature being defined as

\[ T = T_M + (C_l - C_0)m - \frac{\sigma(\theta) + \sigma''(\theta)}{S_m} \kappa - \alpha \Gamma \kappa v, \] (40.6)

where \( T_M \) is the equilibrium liquidus temperature of alloy with the nominal composition \( C_0 \), \( m \) is the liquidus slope on a linearized phase diagram, \( \sigma \) is the surface tension, \( S_m \) is the melting entropy, \( \Gamma \kappa = \gamma S_m \) is the Gibbs-Thomson coefficient, \( \theta \) – is the mis-orientation angle between the interface normal vector and one of the principal crystallographic orientation of crystal, \( \kappa \) is the interface curvature, \( \alpha \) is the kinetic coefficient of atom attachment. The three last terms in equation 40.6 represent the solutal, curvature and kinetic undercoolings respectively. In the curvature undercooling term, the anisotropy of the surface tension is taken into account (Caginalp and Socolovsky, 1994) and is assumed to have the following form

\[ \sigma(\theta) = \sigma_0 \left(1 + \delta_\sigma \cos(4(\theta - \theta_0))\right), \] (40.7)

where \( \sigma_0 \) is the surface tension coefficient, \( \delta_\sigma \) is the amplitude of the anisotropy, \( \theta \) is the orientation angle of the interface normal, \( \theta_0 \) is the orientation angle of a principal crystallographic direction. This expression is commonly used in the literature on the dendrite modelling (Kobayashi, 1993; Ihle and Müller-Krumbhaar, 1994; Caginalp and Socolovsky, 1994). The third term in equation 40.6 describes the kinetic undercooling, which is essential only at rapid solidification conditions, and can be neglected in case of usual welding processes with moderate solidification velocities (Kurz and Fisher, 2001). Thus, the equilibrium interface temperature is expressed by:

\[ T = T_M + (C_l - C_0)m + \Gamma \kappa \left(1 - 15\delta_\alpha \cos(4(\theta - \theta_0))\right) \kappa. \] (40.8)

For constrained solidification conditions, which is the case in welding, the temperature profile can be assumed as given, because the heat diffuses much faster than the solute, only the latter being relevant for solidification morphology (Kurz and Fisher, 2001). Therefore, the solidification behavior at the fusion boundary of weld pool is assumed to be governed by equation set (40.2), (40.4), (40.5) and (40.8).

### 40.2.2 Numerical Solution

The simulation domain is represented by the net of the similar, regularly arranged rectangular cells. Each cell is characterized by a certain state and two additional fields, where the values of
the concentration, \( C \), and the fraction of solid, \( f \), are stored. The latter denotes the fraction of the cell volume filled with the solid phase. The evolution of the solid fraction variable together with the changes of the cell state reflect the evolution of the solidification structure. Five possible cell states are suggested: “liquid”, “solid”, “front”, “melting” and “freezing”. The latter three of them correspond to transition cells in the vicinity of the solid-liquid interface.

### 40.2.2.1 Determination of Curvature and Orientation

In order to reduce the influence of artificial anisotropy caused by the rectangular numerical mesh (Rappaz et al., 1995; Sasikumar and Sreenivasan, 1994), more than the nearest cell neighbors are involved in the definition of transition rules and in determination of the curvature and orientation. Namely, the neighborhood with radius of 3 cells is considered. Additionally, the influence of cells is weighted with the area, which the circle of radius 3.5\( a \) (where \( a \) is the the cell size), which is centered in the middle of the regarded cell cuts within each of the neighbor cells. This is illustrated in Figure 40.3. It is seen from the figure that 44 neighbors influence the behavior of each individual cell. The overlap areas are represented by the weight matrix, \( M^w \). The numbers shown within the cells in Figure 40.3 are the values of the weight matrix members. In order to formalize the calculation procedure, we define a so-called \( \vec{F}_s \)-vector. The value of this vector is equal to the total amount of the solid fraction within the area of radius 3.5\( a \) around the cell (Fig.40.2), weighted with the matrix, defined above:

\[
|\vec{F}_s| = \sum_{i,j=-3}^{3} f_{i,j} M_{i,j}^w, \tag{40.9}
\]

where \( f_{i,j} \) denotes the solid fraction of the cell with indexes \([i,j]\). The orientation of this vector is defined along the line connecting the center of mass of the “total solid” and the center of the regarded cell. The coordinates of the center of mass in a Cartesian coordinate system, parallel to the cell net with the origin in the cell for which the orientation is to be determined, are easily calculated from the solid fraction and the weight matrix elements:

\[
\Delta x = \sum_{i,j=-3}^{3} i \cdot f_{i,j} \cdot M_{i,j}^w, \quad \Delta y = \sum_{i,j=-3}^{3} j \cdot f_{i,j} \cdot M_{i,j}^w \tag{40.10}
\]

The interface normal vector is considered to be represented by

\[
\vec{n} = \frac{\vec{F}_s}{|\vec{F}_s|} \tag{40.11}
\]
Thus, the orientation angle of the crystallographic lattice in Equation (40.7), $\theta$, is calculated as

$$\theta = \arctan\left(\frac{\Delta x}{\Delta y}\right)$$

The interface curvature in a cell with the solid fraction $f$ is calculated from the following semi-empirical expression:

$$\kappa = A + Bf - |\vec{F}_i| \cdot \frac{2}{a}$$

where

$$S = 3 \sum_{i,j=-3}^{3} M_{i,j}^{w} = 38.484$$

$$A = \sum_{i=-1}^{-3} \sum_{j=-3}^{3} M_{i,j}^{w} = 15.754, \quad B = \sum_{j=-3}^{3} M_{0,j}^{w} = 6.976$$

are constants determined from the weight matrix. Geometrically, these constants correspond to the areas of different sectors of the circle with the radius of 3.5, as illustrated in Figure 40.2: $S = 2A + B = 3.5^2 \pi$ is the total area of the circle, $A$ is the area of sector “fde” and $B$ is the area of domain “bcdf”. In order to resolve the dendrite morphology, the radius of the weight matrix, 3.5a, has to be small in comparison with the characteristic length scales of the structure, such as dendrite tip radius and arm spacing.

### 40.2.2 Attachment Conditions: CA-rules

The attachment conditions, e.g. definitions of cell states, are summarized in the Table 40.1. A cell obtains the state “liquid” if the solid fraction within this cell and in its neighborhood is equal to zero. The state “solid” is assigned to a cell in the case, when the solid fraction within them and in its neighborhood is equal to one. Neither “liquid” nor “solid” cells are considered...
to change the value of the solid fraction during a simulation step, e.g. they do not belong to the interface. The state “freezing” allows the solid fraction only to increase. In contrast to this, a “melting” cell is capable only to decrease its solid fraction. The state “front” means that the solid fraction may either increase or decrease. But the decrease in $f$ in a cell is only allowed, if at least one of its nearest neighbors has already solid fraction equal to zero. The increase in solid fraction for a cell with the state “front” may only happen, if at least one of the nearest neighbors around it is already totally filled with the solid phase. These rules ensure the continuity of the interface and keep the calculation procedure stable.

From the other hand, the above rules make the “numerical” interface diffused over several cells. They well represent the dynamic of solid-liquid interface, which may reflect both, freezing or melting. Indeed, the coarsening of dendrite arms in the mushy zone driven by curvature deduction, for example, proceeds by melting convex dendrite arms and freezing concave region near the dendrite trunk. Therefore, a window is introduced between $F_{\text{min}}$ and $F_{\text{max}}$, within which the solid fraction may take both, positive or negative increment. If $|\vec{F}_s| > F_{\text{max}}$, the interface is very concave, that means melting is unlikely to occur, and vice versa, the condition $|\vec{F}_s| < F_{\text{max}}$ can take place on a sharp tip, which will melt because of the high positive curvature. The values of these parameters, $F_{\text{min}}$ and $F_{\text{max}}$, were determined in such way, that the minimum possible $F_{\text{min}}$ and the maximum possible $F_{\text{max}}$, under which the algorithm remains stable, were chosen.

### 40.2.2.3 Numerical Scheme for Solution of Diffusion Equation

The diffusion equation (40.2) was solved with a modified explicit finite difference scheme. The nodal points, situated at the middle of the cells were used as the finite difference mesh. In order to avoid complications due to the discontinuity of the concentration field at the interface,
the diffusion equation was solved in terms of a potential, \( \Pi \), being defined as
\[
\Pi = \begin{cases} 
C & \text{for liquid phase} \\
\frac{C}{k} & \text{for solid phase} 
\end{cases}
\] (40.16)

The potential has no discontinuity at the interface, since the concentrations in solid and liquid are in equilibrium (equation 40.5). Since it is the gradient of the potential to be the driving force of diffusion, the diffusion equation were solved throughout the entire simulation domain with a similar algorithm regardless the presence of different phases. The value of the diffusion coefficient must be, however, adapted with account for different phases which are present in the domain. The finite difference form of the diffusion equation is following:
\[
\Pi_{t,i,j} + \Delta t \left( D_{i,i} \left( f_{i+1,j} - f_{i,j} \right) + D_{j,j} \left( f_{i,j+1} - f_{i,j} \right) \right) = \Pi_{t-1,i,j} + \Delta t \left( D_{i,i} \left( f_{i-1,j} - f_{i,j} \right) + D_{j,j} \left( f_{i,j-1} - f_{i,j} \right) \right)
\] (40.17)

If one of two cells, for example with indices \( i, i-1 \), involved in the finite difference expression, has the solid fraction 1, then the diffusion coefficient was taken to be equal to the one for the solid \( D_s \), otherwise a weighted mean coefficient was used:
\[
D_{i,i-1} = \begin{cases} 
D_s & \text{if } f_i = 1 \text{ or } f_{i-1} = 1 \\
\frac{D_s f_i + D_l (1 - f_i) + D_s f_{i-1} + D_l (1 - f_{i-1})}{2} & \text{if } f_i < 1 \text{ and } f_{i-1} < 1 
\end{cases}
\] (40.18)

When the solid fraction in a cell changes during a simulation step \( \Delta t \), the solute is rejected from the solid to the liquid phase, or is absorbed by the liquid if freezing or melting takes place, respectively. In both cases, the solute is assumed to be redistributed within this cell only. Therefore the total amount of solute in the cell remains constant:
\[
C_{t+1}^l \left( 1 - f^t \right) + C_{t+1}^l k f^t = C_{t}^l \left( 1 - f^{t+1} \Delta t \right) + C_{t}^l k f^{t+1} \Delta t
\] (40.19)

The new value of the potential, which is equivalent in value to \( C_l \), is adjusted in accordance with the change of the solid fraction
\[
\Pi_{t+1}^t = \Pi_t \frac{1 - (1 - k) f^t}{1 - (1 - k) f^{t+1} \Delta t}
\] (40.20)

The explicit FDM scheme is stable under time step iteration when the well known condition \( \Delta t < \frac{a^2}{4 D_l} \) is satisfied. The approximation error is \( O(\Delta t + a^2) \). Because the moving boundary is involved in the calculations and for decreasing the approximation error the time increment in simulations was chosen not to exceed the half of the critical value.

### 40.2.2.4 Determination of Interface Velocity and Solid Fraction Increment

The interface velocity components in the \( x \)- and \( y \)-directions, \( V_x \) and \( V_y \), are determined from the mass conservation equation (eq.40.4). In order to include the undercooling effect,
the following method is used. Let us consider a cell \((i, j)\) in which the velocities are calculated. When evaluating the gradients in equation (40.4), the actual concentration in this cell is replaced by the equilibrium concentration in the liquid \(C_e\), which is determined from the equation (40.8):

\[
C_{e,i,j} = C_0 + \frac{T_{i,j} - T_M - \Gamma \kappa \left(1 - 15\delta_a \cos(4(\theta - \theta_0))\right) \kappa}{m}
\]  

(40.21)

where \(T_{i,j}\) is the imposed temperature in the regarded cell. In the left part of equation (40.4), the equilibrium concentration \(C_e\) is also used. If we also take into account the relation between the concentrations in solid and liquid via the solute partition equation (eq. 40.5) and the solid fractions of the neighbor cells, the following expressions for the velocity components result:

\[
V_x = \frac{D_l}{(1 - k)\alpha} \left(\left(1 - \frac{C_{l,i-1,j}}{C_{e,i,j}}\right)(1 - f_{i-1,j}) + \left(1 - \frac{C_{l,i+1,j}}{C_{e,i,j}}\right)(1 - f_{i+1,j})\right)
\]

(40.22)

\[
V_y = \frac{D_l}{(1 - k)\alpha} \left(\left(1 - \frac{C_{l,i,j-1}}{C_{e,i,j}}\right)(1 - f_{i,j-1}) + \left(1 - \frac{C_{l,i,j+1}}{C_{e,i,j}}\right)(1 - f_{i,j+1})\right)
\]

(40.23)

Here \(C_l\) denotes the actual concentration in the liquid phase.

Knowing the velocity components in both directions, the solid fraction increment is determined as follows:

\[
\Delta f = \frac{\Delta t}{a} \left(V_x + V_y - \frac{V_x V_y}{a} \Delta t\right)
\]  

(40.24)

The solidification behavior of a real system is influenced by the thermal fluctuations which promote the development of the instabilities. The noise is incorporated into the model by introducing the fluctuations in the solid fraction increment. A random number \(p\) between 0 and 1 is generated and the right side of the equation (40.24) is then multiplied by the factor \((1 + \delta_n(1 - 2p))\), with \(\delta_n\) being the noise amplitude. It was found, that the noise amplitude of 0.01 is already sufficient in order to introduce a reasonable fluctuation to the system, without any undesirable influence on the simulated structures (Dilthey et al., 1997a).

### 40.2.2.5 Simulation Sequence

To simulate constrained solidification conditions, a constant temperature gradient was imposed on the domain. The temperature at each location decreases during the simulation step by the value defined by a constant solidification velocity and the time interval. A planar solid seed with the thickness of several cells (5 to 10) at the bottom of the domain was used by setting the solid fraction value for these cells to 1. In order to simulate the initial transient, the
liquid cells were initialized with the equilibrium concentration $C_0$ and the solid cells with $C_0k$. The concentration field may also be initialized using an analytical solution for the concentration profile ahead of the moving planar interface during the transient or steady-state. In this case, the solid cells obtained initial concentration, which is in equilibrium with the concentration in the first liquid cell ahead of the front. The simulation sequence, executed repeatedly, consists of the following procedures: 1) calculation of the concentration field; 2) determination of the state of the cells; 3) calculation of the solid fraction increment; 4) solute rejection (potential update for the cells with $\Delta f \neq 0$); 5) update of the solid fraction variable; 6) update of the imposed temperature field.

### 40.3 Verification of the CA-FDM Solidification Model

A direct experimental observation of structure formation during the solidification of metals and alloys is complicated because of the opacity of these materials and the high temperatures. However, there is a lot of transparent organic substances which behave at crystallization like metals. A principal reason of this is a small anisotropy of the surface tension on the solid-liquid interface, which results in a non-faceted growth with formation of dendrites, cells and other structures, depending on process conditions. To such substances belong, for example, succinonitrile-acetone and halogenated hydrocarbons, which are widely used in experiments on solidification (Kurz and Fisher, 2001). Such experiments can be applied for verification of dendritic modelling. In our case, the work of Akamatsu et al. (1995), who investigated the solidification behavior of thin films of the $\text{CBr}_4 - 8\text{mol}\% \text{C}_2\text{Cl}_6$ alloy, was chosen for this purpose. The experimental conditions corresponded well to the two-dimensional solidification (thin films of 15 $\mu$m thickness) and all relevant data and material properties were listed in that work, which makes it especially suitable for the validation of the CA-FDM solidification model. Below, the experiment for solidification velocity of 35 $\mu$m/s and temperature gradient of 80 K/cm is considered. The numerical domain of 600 $\times$ 3000 cells with the cell size of 1 $\mu$m was used. The comparison of simulated dendrite morphology with the experiment published by Akamatsu et al. (1995) is shown in Figure 40.4. The simulated dendrites look very similar to the dendrites observed in the experiment.

The primary dendrite arm spacing resulted from the simulation (98 to 104 $\mu$m) reasonably agrees with the experimental values (108 to 126 $\mu$m). It should also be pointed out that in both the simulation and the experiment a variation of the primary arm spacing between different dendrites takes place. It indicates an existence of a stable range of the primary dendrite spacing, the phenomenon being considered also in details in the work of Hunt and Lu (1996). Once formed, the dendrites keep the distance between their primary trunks in a certain range. The predicted by numerical modelling spacing reasonably corresponds to the experiment.

### 40.4 Model Application under Welding Conditions

#### 40.4.1 Macroscopic Modelling of Solidification Conditions

In order to calculate the weld pool dimensions and the solidification temperature conditions, the conservation equations of energy, mass and momentum were solved numerically with
40.4 Model Application under Welding Conditions

Figure 40.4: Comparison between the simulated (left) and the experimental (right) dendrite morphology. Experiment of Akamatsu et al. (1995).

the FEM method using the commercial CFD-code FIDAP\(^1\). An autogenous gas-tungsten-arc (GTA) welding with the moving heat source was considered. Detailed description of the model and solution technique is presented by Pavlyk and Dilthey (2001). The principal phenomena driving heat and mass transfer in the weld pool are shown in Figure 40.5. All of them, except viscous drag of plasma, have been included in the model. The following basic assumptions were made: (1) the problem is symmetric in longitudinal plane along the welding direction; and (2) the flow is laminar, incompressible and Newtonian.

The solid-to-liquid phase change was modelled artificially increasing viscosity by factor of \(10^4\) in the temperature range between the liquidus \(T_l = 1793\) K and the solidus \(T_s = 1758\) K for the high strength low alloyed steel S460M. Temperature dependent thermal conductivity, specific heat, density and viscosity were used. The surface tension as function of temperature and the surface active element sulphur was described with the model of Sahoo et al. (1988). In addition to the Marangoni effect, the electromagnetic body forces and buoyancy were considered to be the driving forces of the weld pool convection. The heat exchange between the workpiece and the environment were described by sub-models of the arc heat source, evaporation, radiation and convection. User supplied subroutines were used to implement the electromagnetic body forces, surface tension model and models of the overall heat flux. An adaptive 3-dimensional mesh with high refinement in the region of the weld pool was generated for solution of the conservation equations.

The mesh configuration in the region of the weld pool and the calculated velocity vector plot are shown in Figure 40.6. The velocity vector plot clearly shows the maximum velocities to be concentrated near the weld pool surface. This is caused by a dominating role of the thermo-capillary effect in the weld pool development. An interaction of all driving forces of fluid flow and heat transfer results in a complex weld pool circulation pattern and in a non-trivial weld pool bottom shape. An excellent agreement between calculated and experimental weld pool top shape is demonstrated in Figure 40.7. The isotherms between liquidus and solidus depict the mushy zone, in which the solidification takes place. In order to model solidification structure, the date about the temperature gradients and the local solidification velocity are needed. They can be readily obtained by tracking the calculated temperature profiles in directions perpendicular to the solidification front (“A”, “B” and “C” in Figure 40.7)

\(^{1}\) FIDAP is a registered trade mark of Fluent Inc., Lebanon
as examples). The temperature courses along “A”, “B” and “C” directions are presented in Figure 40.8. The release of the latent heat between the liquidus and the solidus temperatures results in a typical bends on the temperature curves. The temperature gradients were obtained differentiating the temperature curves. The gradients in the mushy zone and in its vicinity as functions of temperature are shown in the inner graph in Figure 40.8. Characteristic values of the calculated temperature gradients were used for the direct simulations of dendritic structure, which are presented below.

40.4.2 Microscopic Simulation of Solidification Structures

To simulate the dendritic structure coupled to the macroscopic weld pool and temperature calculations, the solidification velocity correlated to the welding speed and the specific temperature gradients along directions “A” and “C” (Figures 40.7 and 40.8) were used in the CA-FDM micro-model described above. The necessary material properties were taken for Fe-C system from Kurz and Fisher (2001).

Figure 40.9 shows the simulated dendritic structures for two locations in the mushy zone. The solute concentration is depicted with different grey shades and with the iso-concentrates. The domain height up to 3000 cells were used in order to catch the whole mushy zone and to reach a steady-state as well. The steady-state means, that the dendrite tips and the position, where 100% solid is reached, are at constant temperatures and move with the constant velocity, which is equal to the solidification velocity. Although, the primary dendrite arm spacing given by the simulations are equal for two cases, the size of the mushy zone is quite different. The higher temperature gradient in case “b” (Figure 40.9) results in a smaller distance between the dendrite tip and the region of total solid, which means a more compact mushy zone. From the microscopic CA-FDM simulation in case “a”, the mushy zone has the length of 0.46 mm (Fig. 40.9). From the other side, the macroscopic numerical modelling results in the same distance of 0.46 mm between the solidus and the liquidus in the direction along the symmetry plane (“A” in Figure 40.7) shown in Figure 40.8. In the direction “C” this distance is equal to 0.24 mm, which is slightly smaller than the length of the mushy zone given by the
40.5 Conclusions

Several novel modelling techniques for direct microstructure simulations in material processing were developed recently. Among them the coupled CA-FDM techniques, presented here, was used to simulate dendritic structures in welds. The new theoretical and numerical developments and enormous increase in computer power allow direct simulations of the micro-

Figure 40.6: Adaptive FEM mesh in the region of the weld pool and the calculated velocity field. Equations of fluid dynamics are solved only within the volume bounded with the thick black line ("fluid" entity). Result for 6 mm thick S460 M steel plate. GTA welding, shielding gas helium, welding velocity $V=4.17 \, \text{mm/s}$, current $I=250 \, \text{A}$, arc voltage $U=17 \, \text{V}$, arc efficiency and length 0.85 and 1.6 mm respectively.

micro-simulation (0.28 mm in Figure 40.9b). It proves the consistency between the macro-modelling of the fluid flow and temperature field from one side and the micro-modelling of dendritic structure from the other.

Finally, a non-direct experimental validation of the dendritic modelling can be derived from the comparison between the simulated primary dendrite arm spacing and the micro-graph, shown in Figure 40.10. The micro-graph was taken from the top in the middle of the weld shown in Figure 40.7. In the low alloy steel used, the original $\delta -$ or $\gamma -$ dendritic structure is naturally destroyed during the phase change on cooling ($\gamma - \alpha$ transformation). Therefore, only a special etching allowed to reveal the interdendritic carbides, which are visible in Figure 40.10 as white tracks. The distance between this tracks corresponds to the primary arm spacing, which is of the order of $50 \, \mu\text{m}$ in the experiment as well as in the simulation.

40.5 Conclusions
Figure 40.7: Calculated temperature distribution and mushy zone (top). Top view of the experimental weld pool after an abrupt termination of the arc (bottom).

Figure 40.8: Calculated temperature profiles along directions “A”, “B” and “C” (see Figure 40.7).

structural phenomena taking place during solidification. The direct microstructure modelling using space-resolved cellular automata technique can be thought as a `computational metal-
40.5 Conclusions

Figure 40.9: Simulated dendritic structure in the Fe-0.11%C system, which corresponds to S460M steel. Solidification velocity \( V = 4.17 \text{ mm/s} \), temperature gradient \( G = 100 \text{ K/mm} \) - (a); \( V = 3.0 \text{ mm/s} \), \( G = 200 \text{ K/mm} \) - (b). Grid size 600 × 3000 cells.

Figure 40.10: Experimental micro-graph taken from the top in the middle of the weld shown in Figure 40.7. A special etching reveals the interdendritic segregations, which enable the estimation of the primary dendrite arm spacing.
lography’, which can be useful for both, estimation of the influence of process parameters on the solidification microstructure and for teaching purposes. In order to make the simulation results more quantitative, than qualitative, and to extend the using of this methods many efforts still needed. The computing codes have to be optimized, in order to make simulations faster and available on common computers. The application of the direct microstructure modelling to the real welding conditions require a coupling between the macroscopic heat and fluid flow modelling in the weld pool and the micro-modelling techniques. An example of such coupling with respect to the dendrite solidification phenomenon was presented in this paper. A challenge for future research is a self consistent modelling, in which the results of micro-simulations, in terms of more precise latent heat release as a non-equilibrium function of temperature and non-equilibrium liquidus and solidus, will be implemented in the macro-simulations. It is to expect that for a high speed welding application non-equilibrium effects can be essential in determining both the weld pool dimensions and solidification structures.

Acknowledgements

The authors gratefully acknowledge the financial support of the Deutsche Forschungsgemeinschaft (DFG) within the Collaborative Research Center (SFB) 370 “Integrated Modelling of Materials”.

References


U. Dilthey, V. Pavlik and T. Reichel, Struktursimulation von Schweißgut und Wärmeinflußzone(Structure simulation of weld metal and heat affected zone), Blech Rohre Profile 43 (1996) 637-641

References


41 Forming Analysis and Design for Hydroforming

Kwansoo Chung

41.1 Introduction

Hydroforming has been one of the fundamental sheet metal forming processes for quite a long time, having been developed at least since pre-World War II, when it was applied to the German aircraft industry (Zhang 1999). About 30 years ago, the first tube hydroformed parts were fabricated by expanding straight or pre-bent tubes to make manifold elements and components for sanitary use. Hydroforming offers a way to save material and manufacturing costs while improving product performance in a variety of applications. Advantages include material savings, making shapes that cannot be made by conventional manufacturing methods, and the ability to shape dissimilar tailor-welded materials. It reduces tooling, parts, and labor costs, while significantly improving product performance (Anon 1997). Today, many hydroformed parts are already in mass-production in the automotive industry including exhaust manifolds, exhaust pipes and sub-frames, such as radiator enclosures, space frames, dash assemblies, frame rail, engine cradles etc. (Nakamura et al. 1994, Dohmann and Hartl 1997).

According to process features, hydroforming can be classified as one of the following processes (Zhang 1999, Siegert et al. 2000): (a) flat sheet hydroforming or hydroforming deep drawing (b) tube hydroforming or hydro bulging of tubular components (c) hydro-mechanical deep drawing process (d) new variation of hydroforming including the integral hydro bulge forming of sheet shell products and viscous pressure forming processes.

The hydroforming deep drawing process is one of special drawing processes, in which fluid (generally oil) is pressurized inside a pressure container (Thiruvarudchelvan and Lewis 1999, Kandil 2003). The hydroforming deep drawing is considered as the simplest technique to form complex components economically. It is possible to draw complex components in a single step by the hydroforming deep drawing, while several steps are required by conventional deep drawing processes. The hydroforming deep drawing process generally increases the drawability and minimizes the thickness variation of the cup formed, in addition to other advantages (Tirosh et al. 1977, Yossifon and Tirosh 1984, Yossifon and Tirosh 1985, Yossifon and Tirosh 1988).

Tube hydroforming offers several advantages compared to conventional manufacturing via stamping and welding (Ahmetoglu and Altan 2000): (a) part consolidation manufactured in one operation without welding from a hollow component (b) weight reduction through more efficient section design and tailoring of the wall thickness (c) improved structural strength and stiffness (d) lower tooling cost due to fewer parts (e) fewer secondary operation (f) tight dimensional tolerances and low spring-back (g) reduced scrap.
41.1 Introduction

While numerous variations of the tube hydroforming process exist, the basic procedure remains the same. First, tube preparation before hydroforming process is commonly fabricated through two methods, by rolling forming and welding and by extrusion. Especially, low carbon steel tubing is manufactured from flat sheets by roll forming and electric resistance welding and aluminum tubing is manufactured via extrusion (Ahmetoglu et al. 2000, Luche et al. 2001). Tube hydroforming process starts with a straight precut tube. The tube is often pre-bent in a rotary bending machine or a bent die to fit the hydroforming tool. A straight or pre-bent tube is placed in an open die, the die is then closed before hydroforming during which the tube is preformed with crushing. Then the tube ends are sealed and the tube is filled with hydraulic fluid. Internal hydraulic pressure forces the tube to conform to the shape of die cavities, producing a part having different cross-section shapes (Morphy 1997, Longhouse 1998).

The hydroforming operation comprises two stages: free forming and calibration. The portion of the deformation in which the tube expands without tool contact, is called free forming. As soon as tool contact is established, the calibration starts. During calibration, the tube is forced to conform to the inner contour of the tool by the internal pressure only (Asnafi 1999).

Tube hydroforming is a complex process combined with other processes such as the pre-bending, hydraulic bulging and stamping of tubes. Therefore, the deformation behavior of the tube is significantly complicated. In tube hydroforming, it is essential to understand the basic technology of each element in order to take advantage of the latent formability of tubes. It is also important to clarify the effect of material parameters, tool parameters and the process parameters on tube hydroforming. The material parameters represent the mechanical properties of tubes such as the strain-hardening behavior, anisotropy, elongation, tensile strength and the tool parameters include the shapes and dimensions of dies and the lubrication condition, while the process parameters represent the loading method of internal pressure and axial force, the amount of axial feeding and lateral compression, which cause the failure such as bursting, wrinkling and buckling.

Fuchizawa (1984) analyzed axisymmetric bulge forming of finite-length, thin-walled cylinders under internal pressure using the incremental plasticity theory. Later, he extended the study to explore the influence of plastic anisotropy (Fuchizawa, 1987). Sokolowski et al. (2000) proposed a practical test method to determine the flow stress of tube materials considering the change of material properties during the fabrication of tubes and evaluated the formability during the tube bulge process. Koc et al. (2001) applied the same method to evaluate tube properties in analyzing tube hydroforming.

In hydroforming, axial forces are applied to the ends of the blank tube at the same time when pressure is applied in order to overcome the frictional effect, which causes premature necking by strain localization. However, if the axial forces are too large, buckling may occur. Therefore, a proper combination of the internal pressure and the axial compressive force is important to prevent necking (therefore, tearing) as well as buckling. In order to investigate the buckling failure of the tube, elasto-plastic cylindrical shells were analyzed under the compressive axial force alone by Tvergaard (1983), while the diffuse necking criterion for hydroforming was obtained by Xing and Makinouchi (2001). Tirosh et al. (1996) and Xia (2001) as well as Nefussi and Combescure (2002) considered those two instabilities together. Dohmann and Hartl (1996) and Kim and Kim (2002) proposed a method to control the process by developing the diagram of the proper axial forces under different pressure, which was also used by Admed and Hashmi (1997) for tool design. Asnafi (1999), Asnafi and Skogsgardh
Forming Analysis and Design for Hydroforming

and Chow and Yang (2002) further considered the effect of tool parameters as well as the material and process parameters to study the failure criterion. Lei et al. (2001) applied a ductile fracture criterion to optimize two typical hydroforming processes, a bumper rail and a subframe, under different forming conditions using the finite element method. Lubrication and friction conditions in hydroforming are critical during calibration stage especially in parts where substantial axial feeding is required (Ahmetoglu and Altan 2000, Koc and Altan 2001). Vollertsen and Plancak (2002) proposed a method to determine the friction coefficient at the feeding zone in tube hydroforming.

Recently, the finite element analysis of tube hydroforming process is commonly used to design tools and forming processes. Guan (2000), Kridili (2000) and Koc and Altan (2002) simulated axi-symmetric tube bulging, while Hwang and Altan (2002) simulated the crushing performance of circular tubes into triangular tubes. Koc et al. (2000) also predicted the protrusion height of Tee-shaped parts using three-dimensional finite element analysis. Quite often, a pre-bending process precedes the hydroforming. Therefore, it is proper to take it into account the deformation of the pre-bending for hydroforming simulations. Using finite element analysis, some researchers investigated the cross-section shape change and thickness thinning during pre-bending and integrated the results with subsequent hydroforming simulation (Wu and Yu 1996, Dwyer et al. 2002, Trana 2002). Yang et al. (2001a) further included the spring-back effect after the pre-bending operation. Berg et al. (1996) developed a numerical scheme to resolve concave cross-section problems introduced by the local buckling during prebending and preforming processes.

Hydroforming processes in general involve stretching, bending and various combinations of basic deformation modes. In view of mechanics, the analysis of hydroforming operations involves the non-linearity of kinematics as well as the behaviors of materials and the contact algorithm. Designing die and preform shapes especially in the early design stage is one of the most difficult tasks. The optimal design is essential to produce final products without defects and also to effectively reduce production time. Design procedures however would require a tremendous amount of trials for completion. For the optimization of tube hydroforming processes in the design stage, several design schemes were introduced. Ghouati et al. (1999) and Yang et al. (2001b) used the sensitivity analysis based on the finite element scheme as a design tool. Kim et al. (2002) designed the preform to get the uniform wall thickness after hydroforming using the backward tracing scheme. Yoon and Chung (2003) applied a direct design method based on the ideal forming theory to design the preform for tube hydroforming, which is summarized here.

In order to improve the conventional trial-and-error based practices for optimizing forming processes, a direct design method called the ideal forming theory, has been previously developed (Chung and Richmond 1992a, b, 1994, Chung et al. 2000). In this theory, materials are prescribed to deform following the minimum plastic work path (or the proportional true strain path for materials having smooth yield surfaces) and the initial blank shape is obtained from the global plastic work extremum condition as a one-step backward solution, in which the final product shape is specified. In order to consider the local thinning effect due to friction, the method based on the modified extremum work criterion was also developed (Chung et al. 2000). The ideal forming theory has been successfully applied for sheet forming processes to optimize the flat sheet blank (Barlat et al. 1994, Chung et al. 1997, Richmond and Chung 2000) and also for bulk forming in steady (Richmond and Devenpeck 1962, Richmond and
41.2 Ideal Forming Design Theory for Tube Hydroforming


Anisotropy in plasticity is represented by the yield stress potential and the strain-rate potential: the yield stress potential is convenient to use for the elasto-plastic formulation, while the strain-rate potential is for the rigid-plastic formulation. The stress potential Yld91 (Barlat et al. 1991) and the strain-rate potential Srp93 (Barlat et al. 1993) were developed as (quasi-) conjugate quantities to phenomenologically account for the structural difference in polycrystals. However, those do not have enough anisotropic coefficients to express the anisotropy both in the yield stress and the plastic strain ratio (r-value) in the simple tension. Later, the yield stress potential, Yld96 (Barlat et al. 1997), and its quasi-conjugate strain-rate potential Srp98 (Barlat et al. 1999) were proposed to capture the anisotropy both in the simple tension yield stress and the r-value. The strain-rate potential Srp98 was used in this work. For demonstration purposes, the preform design of aluminum alloy tubes was considered for the hydroforming of industrial parts to show the design capability including the optimization of the cross-sectional shape and the tube length.

41.2 Ideal Forming Design Theory for Tube Hydroforming

When a (sheet) part is discretized with meshes using shape functions and the surface traction is approximated by point forces, the total plastic work, \( W \), becomes a function of the initial configuration and the final configuration upon the assumed minimum plastic work path imposed on each material element; i.e.,

\[
W = W[\varepsilon(x^i, X^i)] = \int \int \sigma(\varepsilon) d\varepsilon dV_0
\]  

(41.1)

where \( \sigma, \varepsilon \) \((= \int d\varepsilon)\), \( V_0 \) are the effective stress, the effective strain and the initial volume, respectively, while \( x^i \) and \( X^i \) are the final and initial positions of the i-th node in the global Cartesian coordinate system. In the design stage of forming processes, part shapes are prescribed and initial blank shapes are pursued. Therefore, the optimization is performed with respect to the initial configuration.

In order to derive the effective strain \( \varepsilon \), the materially embedded local Cartesian coordinate system \( \hat{X} \) is considered for each linear triangular membrane element at the initial configuration as shown in Figure 41.1(a). Then, the deformation gradient \( F \) and the Cauchy strain tensor \( C \) become

\[
F(\hat{X}) = g_k \otimes G^k = (g_k \cdot e_i^c)(G_k \cdot \hat{E}_j)^{-1} e_i^c \otimes \hat{E}_j, \quad C(\hat{X}) = F^T F = C_{ij} \hat{E}_i \otimes \hat{E}_j, \quad (41.2)
\]

respectively, where \( g_i \) and \( G^i \) are the covariant and contravariant base vectors for the convected coordinate system, respectively. Also, \( e_i^c \) and \( \hat{E}_i \) are the unit base vectors of the global and local Cartesian coordinates, respectively. Now, the logarithmic (or true) strain tensor \( \varepsilon \) becomes

\[
\varepsilon(\hat{X}) = \varepsilon_{ij} \hat{E}_i \otimes \hat{E}_j = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} \frac{1}{2} \ln \lambda_1 & 0 \\ 0 & \frac{1}{2} \ln \lambda_2 \end{bmatrix} \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} \quad (41.3)
\]
where $\lambda_{i=1,2}$ and $\theta$ are the principal values and principal direction of the Cauchy strain tensor, respectively. When the material element deforms following the proportional true strain path (the minimum plastic work path for materials having smooth yield surfaces), the effective strain is obtained from the effective strain-rate by substituting the true strain rate $\dot{\varepsilon}$ with the true strain tensor $\varepsilon$ (Chung and Richmond 1993), which is obtained from Equation (41.3). As for the effective strain-rate, the anisotropic non-quadratic strain rate potential $\text{Srp98}$ (Barlat et al. 1999) was employed for this work.

In optimizing the total plastic work with respect the initial configuration, the following geometric constraint on the initial configuration is considered: all materials are supposed to stay on the surface of the initial cylindrical tube. This constraint can be effectively implemented into the optimization procedure using the penalty method based on the local Cartesian coordinate systems $\bar{X}$, which are defined on an assumed cross-section of the initial tube as shown in Figure 41.1(b). There are commonly the ‘$n_{\text{sect}}$’ number of nodes for each cross-section, while there are commonly the ‘$n_{\text{length}}$’ number of nodes along the tube length. The local coordinate systems are defined for each cross-sectional node at the one end of the tube and they are shared by all nodes on the same row (along the longitudinal direction). Note that $\bar{X}_1$ and $\bar{X}_2$ are aligned with the longitudinal and tangential directions of the tube, respectively.
Now, the penalty function to impose all nodes to stay on the initial cylindrical tube becomes

\[ \Pi = \sum_{i=1}^{n_{\text{node}}} \frac{1}{2} C \left( \bar{X}_i^3 - \bar{X}_i^3 \right)^2 + \sum_{i=1}^{n_{\text{sect}}-1} \frac{1}{2} C \left( \bar{X}_i^1 - \bar{X}_{i+1}^1 \right)^2 \]

\[ + \sum_{i=n_{\text{node}}-n_{\text{sect}}+1}^{n_{\text{node}}-1} \frac{1}{2} C \left( \bar{X}_i^1 - \bar{X}_{i+1}^1 \right)^2 \]

(41.4)

where “\(n_{\text{node}}\)” is the total number of nodes and “\(C\)” is a penalty parameter which is usually chosen to be \(10^5 \sim 10^8\). Note that the first term is to (approximately) satisfy the geometric constraint by imposing \(\bar{X}_3\) to be identical for all nodes along the same row, while the second and third terms are to impose the nodes on both ends of the tube stay on the cross-sectional planes.

An additional constraint is imposed in the optimization procedure to maintain the assumption that the designed initial cross section is less than or equal to the smallest cross section of the final part geometry. The cross-section of the initial tube is supposed to expand by internal pressure so that the initial cross section solution, which is larger than the smallest cross section of the final product is meaningless. Therefore, the following penalty function is further considered:

\[ \Psi = \sum_{k=1}^{n_{\text{length}}} \sum_{k=1}^{n_{\text{length}}} \frac{1}{2} C \left( P_k + \gamma \right)^2 \]

(41.5)

where

\[ P_k = L^2_{\text{Target min}} = \sum_{i=1}^{n_{\text{sect}}-1} \left( \bar{X}_i^{(k-1)*n_{\text{act}}+i} - \bar{X}_i^{(k-1)*n_{\text{act}}+(i+1)} \right)^2 \\
- \left( \bar{X}_i^{(k-1)*n_{\text{act}}+n_{\text{act}}} - \bar{X}_i^{(k-1)*n_{\text{act}}+1} \right)^2 \]

(41.6)

Here, if \(P_k + \gamma \geq 0\), then \(P_k + \gamma = 0\) and if \(P_k + \gamma < 0\), then \(P_k + \gamma = P_k + \gamma\) (where \(\gamma \ll 1\)). Also, \(L^2_{\text{Target min}}\) is the square of the smallest contour length of the cross section in the final product. Note that \(\bar{X}_i\) is the initial position vector of the \(i\)-th node in the global Cartesian coordinate system.

The initial nodal positions, which describe the optimum initial blank tube is obtained from the following extremum condition:

\[ \frac{d\Pi}{d\bar{X}_i} = 0 \quad \text{for } i = 1 \sim n_{\text{node}} \]

(41.7)

where

\[ \Pi = W(\hat{\bar{X}}(\bar{X})) + \Pi(\bar{X}) + \Psi(\bar{X}(\bar{X})) \]

(41.8)

Equation (41.7) is a nonlinear equation, which is iteratively solved using the Newton-Raphson scheme. Note that \(W\) and \(\Psi\) in Equation (41.7) are conveniently represented with respect to the
materially embedded coordinate system $\hat{\mathbf{X}}$ and the global coordinate system $\mathbf{X}$, respectively, therefore, their gradient with respect to the local coordinate system $\bar{\mathbf{X}}$ defined on the initial blank tube in Equation (41.7) should be performed considering the proper transformations between the unit base vectors of the three coordinate systems.

For the modified extremum condition to incorporate the Coulomb friction effect, the optimization is performed in Equation (41.7) in which the gradient of $W$ is replace with the following quantity:

$$\left[ \frac{dW(\hat{\mathbf{X}})}{d\bar{\mathbf{X}}} - \mathbf{b} \right]$$

(41.9)

where

$$\mathbf{b} = \mu |\mathbf{f}_n| \mathbf{s} \cdot \bar{\mathbf{F}}.$$  

(41.10)

Here, $\mu$ is the friction coefficient, $\mathbf{f}_n$ is the nodal force normal to the surface in the final product, $\mathbf{s}$ is the unit tangential vector to define the node sliding direction and $\bar{\mathbf{F}}$ is the average deformation gradient of elements surrounding the node. Note that the third component of $\mathbf{b}$ in Equation (41.10) is vanished (Chung et al. 2000). The optimization procedure iteratively continues with the new $\mathbf{b}$, until the Coulomb friction condition is satisfied for the nodal force $f(\mathbf{x} = \partial W / \partial x_i)$ in the final product.

### 41.3 Strain-Rate Potential: Srp98

The strain-rate potential Srp98 (Barlat et al. 1999) for the plane stress state, denoted by $\Phi$, is

$$\Phi = \alpha_1 |\dot{e}_1|^a + \alpha_2 |\dot{e}_2|^a + \alpha_3 |\dot{e}_1 + \dot{e}_2|^a = 2k\dot{\varepsilon}^a$$

(41.11)

where $\dot{e}_1$ and $\dot{e}_2$ are the principal values of the isotropic plasticity equivalent (IPE) strain rate $\dot{\varepsilon}$, which is obtained from the plastic strain rate $\dot{\varepsilon}$ as

$$\begin{bmatrix} \dot{\varepsilon}_x \\ \dot{\varepsilon}_y \\ \dot{\varepsilon}_{xy} \end{bmatrix} = \begin{bmatrix} (2c_2 + c_3)/3 & (c_2 - c_3)/3 & 0 \\ (c_1 - c_3)/3 & (2c_1 + c_3)/3 & 0 \\ 0 & 0 & c_6 \end{bmatrix} \begin{bmatrix} \dot{\varepsilon}_x \\ \dot{\varepsilon}_y \\ \dot{\varepsilon}_{xy} \end{bmatrix}. $$

(41.12)

Also,

$$\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} = \begin{bmatrix} \alpha_x \cos^2 \beta + \alpha_y \sin^2 \beta \\ \alpha_x \sin^2 \beta + \alpha_y \cos^2 \beta \\ \alpha_{x0} + (\alpha_{z1} - \alpha_{z0}) |\sin 2\beta| \sin 2\beta | \end{bmatrix} $$

(41.13)

where the two angles $\beta$ and $\beta'$ are defined as

$$\tan \beta = \frac{\dot{\varepsilon}_1 - \dot{\varepsilon}_x}{\dot{\varepsilon}_{xy}}, \quad \tan 2\beta' = \frac{2\dot{\varepsilon}_{xy}}{\dot{\varepsilon}_x + \dot{\varepsilon}_y}. $$

(41.14)

Here, $c_i=1,2,3,6$ and $\alpha_x$, $\alpha_y$, $\alpha_{z0}$ and $\alpha_{z1}$ are the anisotropic coefficients, which become all identical, typically with the value of 1.0, for isotropic materials. The exponent $a$ in Equation (41.11) is used to match the shapes of strain-rate potentials with those calculated from
polycrystal models. For FCC and BCC polycrystals, the exponents are $4/3$ and $3/2$, respectively. Therefore, in the absence of more specific information, $4/3$ is a recommended value for aluminum alloys, especially for isotropic materials. The value $k$ in Equation (41.11) is a constant to accommodate the difference between the reference strain-rate state being used to define the effective strain rate with that being used for the stress-strain hardening curve.

### 41.4 Preform Design for Hydroforming Parts

For demonstration purposes, two industrial parts were considered to design optimum preform tubes. The following hardening data was used along with the isotropic strain rate potential $\text{Srp98}$ when $a = 4/3$: $\sigma = 359.84(0.001 + \varepsilon)^{0.223}$(MPa). The uniform thickness of the preform tube was 2 mm. It took only about two minutes to finish computations. The final part shape after hydroforming of the first example is shown in Figure 41.4. Because of symmetry, only half section was considered.

![Figure 41.2: The final part shapes and the predicted optimum preform shapes.](image)

The design code showed that the amount of the side feeding was 3.3 mm with the average pressure 60.78 MPa, while experimental optimum feeding to make the final part was 3.7 mm with pressure 75 MPa. Figure 41.3 compares the thickness strain predicted with the friction coefficient 0.15 and that predicted without friction. With the friction effect incorporated, more thinning was observed.
Figure 41.3: Thickness strain contours: (a) no friction (b) friction coefficient = 0.15.

In the second example, an automotive part shape shown in Figure 41.4 was considered. Figure 41.5 shows the predicted optimum preform tube obtained without considering friction. The result shows that the optimum cross section is not circular but close to an ellipse. The
predicted thickness strain contour is plotted on the final part geometry in Figure 41.4. The two example solutions confirm that the ideal forming design code can provide useful design information in the early design stage, which ultimately reduces the number of trials and errors in the finite element analysis to complete the design.

41.5 Summary

The ideal forming theory, developed as a direct design method to guide iterative design practices based on analysis, was applied for the tube hydroforming (having non-flat preform shapes). The penalty method was incorporated to impose geometric constraints in the preform shape. The frictional effect was implemented by employing the modified extremum work criterion. For material anisotropy, the strain-rate potential Srp98 was utilized. The proposed method was successfully applied for hydroforming of two industrial parts, confirming that the direct design approach based on the ideal forming theory is effective in providing design information on the optimum preform shape and the feasibility of the final part shape.

References

References


Koc., M, Allen, T., Jiratheranat, S., Altan, T., 2000. The Use of FEA and design of experiments to establish design
References


42 Sheet Springback

Robert H. Wagoner

42.1 Introduction

As used in this chapter, springback refers to the usually undesirable and often deleterious shape change that occurs following a press-forming operation when the forming loads are removed from the work piece. It is driven by internal stress patterns that are equilibrated with external loads at the end of the forming operation. Springback is particularly important for long, slender structures such as sheet-formed parts because small changes in stress pattern across the thickness translate into significant changes of the radius of curvature, which in turn translate into large displacements normal to the thickness.

The large displacements inherent in sheet springback lead to many problems downstream of the forming operation, including difficulties in subsequent forming operations, final assembly (particularly welding), appearance, and quality. These problems have a significant economic impact. A recent estimate (Wenner 2001) for the U.S. automotive industry alone put the cost of springback at more than $50 million per year, and a similar estimate has been made for U.S. Air Force applications (Semiatin 2001). In spite of this, there has been relatively little research done on springback related to typical sheet forming practice.

There are two basic approaches to avoiding the drawbacks inherent in springback: 1) minimizing the springback displacements, usually by increasing sheet tension, and 2) compensating for predicted or measured springback by tooling or process changes. (A third possible approach, designing subsequent operations and final products for springback tolerance, doesn’t address the root problem and will not be considered here.) For operations where the sheet tension cannot be controlled (e.g. in pure bending), compensation is the only alternative. Fortunately, pure bending is the simplest case to model and can usually be predicted and compensated for, although differences in material properties and process variables lead to unavoidable scatter.

For typical press-forming operations, such as those encountered in the automotive industry, control of sheet tension via drawbeads and other tooling features has been the historical mainstay of successful forming practice. Tool design optimization has usually been accomplished using rules-of-thumb for the initial design and then trial-and-error tooling changes until the final desired part shape is obtained. However, as product cycle times continue to decrease, the need to understand, predict, and compensate for springback at the die design stage has become increasingly important. While simulation techniques for sheet forming operations have been nearly universally adopted, the current challenge is applying these techniques in a reliable, predictive way to incorporate springback prediction.
At the time of this writing, the challenges facing control of springback in automotive-type forming operations concentrate into several areas: 1) reliable and robust numerical techniques for predicting springback following a sheet forming operation of a complex part, 2) material properties and mechanics formulations capturing the essential material behavior of the work piece during and after forming, and 3) design methods that make use of springback simulations of presumed accuracy to guide changes in tooling without the need for experimental intervention (i.e. trial and error). Because of space limitations, design methods making use of springback simulations will not be reviewed in this chapter, but the reader is referred to recent papers that have been submitted for publication in the open literature (Gan and Wagoner 2003a, Gan et al. 2003b).

This chapter is aimed at providing a brief introduction of the principal literature of springback, both analytical and experimental. The extensive references should be sufficient to enable the reader to trace the known knowledge in the area. An even briefer review of some recent developments is presented based on the draw-bend test results obtained by the author and co-workers. This chapter is thus organized into three principal parts, as follows:

1. Review of Simulation Literature
2. Review of Experimental Literature
3. Draw-Bend Springback

42.2 Review of Simulation Literature

A variety of analytical and simulation techniques have been applied to springback. These generally may be classified as analytical (usually closed form) or numerical (typically finite element) techniques.

**Analytical Solutions**  Analytical solutions for springback of plane stress, pure bending were presented for elastic-perfectly plastic material response (Gardiner 1957), and were extended to plane-strain pure bending (Queener and De Angelis 1968), and initially curved cases (Shaffer and House 1955). Later solutions were for plane-strain bending with superimposed tension and with a wider range arbitrary R/t (radius of curvature-to-thickness ratio) (Baba and Tozawa 1964, Ingrarson 1975, Ueda et al. 1981, Yu and Johnson 1982, Yuen 1990, Tozawa 1990, Wang et al. 1993, El-Domiaty et al. 1996, Zhang and Hu 1997). Narrow plane-stress strips have been analyzed for bending as well (Chan and Wang 1999). Reviews of pre-1950’s developments have been presented (Nadai 1950, Phillips 1956), as have more recent analytical springback methods and approaches (Huang and Gerdeen 1994, Yu and Zhang 1996).

Analytical methods other than finite element modeling have been applied to springback in die forming, often with limitations to pure bending, as appropriate for U-bending (Sudo et al. 1974), flanging (Wang 1984), sidewall curl (Thompson and Ellen 1995), repeated bending unbending (Chu 1986, Kuwabara et al. 1996a) or stamping with deformable tools (Zhang and Lin 1997). These analyses rely on radii of curvature during the forming operation established by geometrical considerations of the dies, but do not usually solve explicitly for such shapes consistent with the material mechanics and contact friction with the tools. With superimposed
tension, these methods may be applied to plane-strain draw die forming (Yoshida 1965, Duncan and Bird 1978, Wenner 1983) with gentle tooling curvatures. Such methods have also been applied to stretch bending of channels (Ueda et al. 1981, Mickalich and Wenner 1988) or draw bending of top-hat sections (Zhang and Lee 1995). Empirical rules have also been used to treat springback (Levy 1984).

**Finite Element Modeling**  With the rapid increase in computation power, finite element methods (FEM) for analyzing and predicting springback have become more attractive. Recent benchmark tests and accompanying papers (Makinouchi et al. 1993, Shen et al. 1995, Lee et al. 1996, Gelin and Picart 1999, Yang et al. 2002) illustrate the state of the art in predicting springback with FEM. In particular, the 1993 benchmark (Makinouchi et al. 1993), Figure 42.1a, represents a flanged channel forming operation that was simulated by many commercial and special-purpose programs with widely varying results, Figure 42.1b. This result was surprising in view of the reliability developed in forming simulations by that period, however, it should be noted that the experimental scatter, as approximated by the error bars on Figure 42.1b, generally exceeded the scatter of the simulations.

![U-Channel Benchmark](image)

**Figure 42.1**: U-channel springback benchmark of NUMISHEET ’93 (Makinouchi et al. 1993). Each point on the plot represents a simulation result while the error bars represent the approximate scatter of the experimental results.

FEM simulations of springback are much more sensitive to numerical tolerances than are forming simulations (Mattisson et al. 1995, Wagoner et al. 1997, Lee and Yang 1998, Li et al. 1999b), including effects of element type (Li et al. 1999a), integration scheme (He and Wagoner 1996, Focellese et al. 1998, Lee and Yang 1998, Li et al. 1999a,b, Narasimhan and Lovell 1999), and unloading scheme (Tang 1987, Yuen 1990, Li et al. 1999b). Because of
the inherent numerical sensitivity, implicit schemes for both loading and unloading i.e. implicit/implicit) have been popular (Wagoner et al. 1997, Hu and Du 1999, Li et al. 1999a,b, Geng and Wagoner 2000), as well as attempts to link dynamic explicit simulations of forming operations (upon which many commercial programs are based) to static implicit simulations of springback (Mattisson et al. 1995, He and Wagoner 1996, Park et al. 1999, Valente and Traversa 1999). Nearly every possible approach to simulating springback has been recently reported as successful, including explicit/explicit (Montmayeur and Staub 1999), static explicit/static explicit (Kazama et al. 1999), and even one-step methods (Abdelsalam et al. 1999). Hybrid approaches have been developed for post-processing of forming FEM results for springback (Pourboghrat and Chu 1995) and for optimizing die design by iterating with analysis (Karafillis and Boyce 1992a,b, Karafillis and Boyce 1996, Ghouati et al. 1998).


### 42.3 Review of the Experimental Literature

Corrections for springback are essential during die design in order to obtain specified final shapes. Correction curves based on empirical information or simple theories for small-springback in pure bending cases have been available for many decades (Schroeder 1943, Sachs 1951, Levy 1984) based on the basic assumptions of engineering beam bending. Many analytical solutions have been derived, and reviews have appeared both for the early work (Nadai 1950, Phillips 1956), and more recent contributions (Huang and Gerdeen 1994, Yu and Zhang 1996).

While the literature dealing with analysis of springback of metal sheets is extensive, carefully-controlled experiments of springback under realistic forming conditions (i.e. involving bending and unbending simultaneous with imposed tension and sliding over the tooling) are less common. Many experiments have been carried out under pure-bending (i.e. with minimal tension) conditions of various kinds: cylindrical tooling (Yu and Johnson 1983, Yuen 1990, Sanchez et al. 1996), U-bending/channel bending (Sudo et al. 1974, Chakhari and Jalinier 1984, Hino et al. 1999), V bending (Chakhari and Jalinier 1984, Zhang and Hu 1997, Hino et al. 1999), and flanging (Wang 1984). Such experiments show springback increasing with $R/t$ (tool radius/sheet thickness), but such results have little application to situations where significant sheet tension is present, because sheet tension dominates other process variables in determining springback.

Stretch-bend tests (Ueda et al. 1981, Kuwabara et al. 1996b, Hino et al. 1999) allow careful control of sheet tension during bending, but do not typically exhibit bending and unbending, nor large sliding over the tooling common in press forming operations. These tests do show, however, the pervasive effect of sheet tension on springback, especially for sheet tensile stresses in the range of the material yield stress.
Actual forming operations, such as a two-dimensional idealization of draw/stretch bending of a flanged channel (also called a top-hat section) are used to assess practical springback. This geometry is the most-studied of springback cases (Ayres 1984, Davies 1984, Hayashi 1984, Umehara 1990, Schmoeckel and Beth 1993, Pourboghrat and Chu 1995, Bayraktar and Altintas 1996) (including benchmark testing and simulations (Makinouchi et al. 1993)) because of its practical importance, and because of the obvious, measurable presence of sidewall curl. However, in such operations, the sheet tension is determined only indirectly, either from the blank holder force, or from a drawbead simulation (Nine 1978, Nine 1982, Wang 1982, Nine 1983), both of which depend on the coefficient of friction, which is usually known only approximately. Because of the dominant role of sheet tension in springback, lack of control or direct measurement of this quantity is a serious drawback for scientific use of such experiments in verifying simulation techniques. Such operations have common features, including reduced springback and curl for increased blank holder forces (especially when sheet tension approaches the yield stress). Curl has been shown to disappear as R/t approaches 2 for a variety of steels (Davies 1984, Hayashi 1984, Umehara 1990). For smaller R/t, the curl may reverse direction.

The exceptional work is by Liu (1988) for flanged channels, and Kuwabara et al. (1996b) for draw bending. Liu devised a special channel flanging experiment in which the restraining force was applied directly by a controlled hydraulic cylinder. In this way, sidewall curl and springback could be assessed in terms of known restraining forces for several materials. Kuwabara et al. used a draw-bend apparatus that appears similar to the one in the current work, although dimensions and specifications were not provided (Kuwabara et al. 1996b). Reported springback angles decreased with increasing R/t and the usual dependence on sheet tension was observed. The current draw/bend test (Wenzloff et al. 1992, Haruff et al. 1993, Vallance and Matlock 1992), conceived for friction testing, is similar in concept to Liu’s experiment, with four significant differences of the current device:

1. much longer draw distances are attainable, with correspondingly increased precision of shape measurement,
2. there is no bend at the bottom of the channel to complicate interpretation,
3. there is only a single tool, thus avoiding changing clearances and geometries during the punch stroke, and
4. the material hardening curves were measured and fit in some detail.

The sensitivity of springback to a Bauschinger effect, particularly for bend/unbend operations (as is the case for draw/bend tests and flanged channel forming) has been noted in several analyses (Baba and Tozawa 1964, Yoshiida 1965, Tozawa 1990, Pourboghrat and Chu 1995, Zhang and Lee 1995, Tang SC, 1996, Focellese et al. 1998, Kuwabara et al. 1999). However, until recently, the magnitude of the Bauschinger effect in sheet metal tension-compression (Kuwabara et al. 1995c, Balakrishnan 1999) or bend/reverse bend (Jiang 1997, Shen 1999) had seldom been measured because of the instability of metal sheets subjected to in-plane compression.

A variety of manufacturing techniques has been developed for dealing with springback. The simplest involves designing tools to over-bend the sheet to compensate for springback.
However, this is not applicable to complex curved parts and, without additional control, this approach can lead to large variations in practice as material properties and thickness vary. Other approaches include the use of deformable tools (Zhang and Lin 1997, Zhang et al. 1997), adjustments to small die radii and clearances (Hayashi 1984, Umehara 1990), throughput-thickness deformation (Chou and Hung 1999), variable or stepped blank holder force control (Tozawa 1990, Hishida and Wagoner, 1993, Schmoeckel and Beth 1993, Sunseri et al. 1994, Sunseri et al. 1996, Han and Park 1999), multiple forming steps (Ayres 1984, Nagai 1987, Tozawa 1990), or reconfigurable tooling (Kutt et al. 1999). Each of these methods can benefit from accurate knowledge of springback from measurement and analysis.

### 42.4 Draw-Bend Springback

As outlined above, the existing springback literature has either a) limited relevance to standard sheet forming because of different forming conditions, or b) high scatter arising from imprecise control of sheet tension in industrial configurations. The draw-bend test (Vallance and Matlock 1992, Wenzloff et al. 1992, Haruff et al. 1993) avoids the shortcomings of many of the previous test configurations. As shown schematically in Figure 42.2a, the test reproduces closely the geometry of drawing a sheet over a die radius (such as into a die cavity). The strip is bent around a fixed radius and then straightened as it is drawn for large distances under significant sheet tension. The test not only reproduces industrial sheet forming conditions, it may be conducted with precise closed-loop control of the sheet tension. The large specimen geometry allows measurement of springback angles to within $0.3^\circ$, corresponding to an outer fiber strain of 0.00001 (Wagoner et al. 1997).

**Draw-Bend Results** The author and coworkers have used the draw-bend test as a basis for springback experiments and corresponding simulations (Li and Wagoner 1998, Li et al. 1999a,b, Geng and Wagoner 2000, Wagoner et al. 2000a,b, Carden et al. 2002, Li et al. 2002, Geng and Wagoner 2002). The simulations make use of either research finite element programs SHEET-S (Wagoner et al. 1989, Saran et al. 1991c, Wang and Wagoner 1991a,b, Keum et al. 1992) and SHEET-3 (Germain et al. 1989, Keum et al. 1989, Wagoner et al. 1990, Kim and Wagoner 1991, Saran and Wagoner 1991a,b, Wagoner and Zhou 1992, Wagoner and Zhou 1993, Zhou and Wagoner 1994, Zhou and Wagoner 1995), or commercial program ABAQUS (ABAQUS User Manual, version 6.2, version 6.2.1.). Comparison of measurements with simulations, Figure 42.2b, reveals the complexity of the springback behavior and its simulation under these conditions. (Note that the normalized back force in Figure 42.2b represents the controlled sheet tension expressed as a fraction of the sheet tension required to yield the strip specimen.)

**Numerical Sensitivity** Initial 2-D simulations, using standard numerical parameters and material laws for sheet forming simulations (Figure 42.2b, Curve 1), showed dramatic discrepancies with the experiments. A 2D sensitivity analysis (Li and Wagoner 1998, Li et al. 1999a, b) of the simulation results in terms of numerical parameters was undertaken to verify the consistency of the simulations themselves. On this basis and considering later 3D simulations, the following guidelines were reached for simulating sheet springback:
42.4 Draw-Bend Springback

Figure 42.2: Improvement of draw-bend springback prediction by choice of numerical procedure and better material understanding (Li and Wagoner 1998, Li et al. 1999a,b, Geng and Wagoner 2000, Wagoner et al. 2000a,b, Carden et al. 2002, Li et al. 2002, Geng and Wagoner 2002). \(<\sigma>\) represents the standard error of fit for each simulation relative to measurements. Unless otherwise noted, von Mises yield and isotropic hardening were employed. Notation:

1. Plane-stress simulation, 5 integration points (IP), 600 elements along length (El), \(<\sigma>=26^\circ\)
2. Plane-strain simulation, 51 IP, 600 El (remaining results use 51 IP, 600 El), \(<\sigma>=19^\circ\)
3. Plane-stress simulation, \(<\sigma>=11^\circ\)
4. 3-D shell simulation (w/persistent anticlastic curvature), \(<\sigma>=5.7^\circ\)
5. 3-D shell/Geng-Wagoner anisotropic hardening law, Barlat '96 yield function, \(<\sigma>=2.3^\circ\) and 1.2\(^{\circ}\), respectively.

- Up to 51 integration points through the sheet thickness are required (compared with 3 to 5 normally used for sheet forming simulation)
- Fine meshes, typically 4 times finer (or less than 5 degrees of turning angle per element) as compared with sheet-forming meshes
- Convergence must be insured to very tight tolerances, typically to one part in 10,000
- Because of the numerical sensitivity, it is essential to perform sensitivity analyses
- Implicit time integration appears to be essential in view of the numerical sensitivity to convergence tolerances (He and Wagoner 1996)
- 3-D elements are required because of persistent anticlastic curvature
- In some cases, it is necessary to consider elastic-plastic springback
- For R/t ratios less than 5, shell elements are not sufficient and solid elements are required
It should be noted that these requirements increase the computational requirements for forming/springback simulation considerably in comparison to forming alone, not only in raw computation time, but in the reliability of the algorithms employed. The time for simulation of forming and springback can increase by a factor of 100 times that required for the forming step alone.

**Anticlastic Curvature** Once the numerical parameters and procedures had been satisfactorily established, the 2D simulation results were still hundreds of percent in error for those cases where the sheet tensile stresses approached the material yield stress, although for other back forces far removed from this value the agreement was much better. (Compare Curve 3 with the experimental results for a normalized back force near 0.9 and 0.5, for example.)

3-D shell simulations and re-measurement of the specimens revealed the presence of persistent anticlastic (secondary) curvature, the role of which in nominally 2D springback testing was not previously appreciated (Carden et al. 2002, Li et al. 2002). The sudden depression of springback angle for sheet tension approaching yield was shown to be a result of this effect.

**Role of Yield Function** The magnitude of anticlastic curvature, and therefore the springback angle, were shown to depend intimately on the shape and anisotropy of the yield function. Maintaining an isotropic hardening rule but changing the yield function form yielded the following standard errors of fit (not shown on Figure 42.2b for clarity reasons):

- Hill quadratic yield (Hill 1948, Hill 1950): 11.3°
- Barlat YLD89 (Barlat and Lian 1989): 10.2°
- Von Mises: 5.7°
- Barlat YLD96 (Barlat 1997): 2.0°

As discussed more fully elsewhere (Geng and Wagoner 2002), the springback simulation is intimately tied to the form of the yield function in two respects: 1) the lateral (TD) yield stress controls the persistence of anticlastic curvature, with a lower yield enhancing the permanent secondary curvature that remains after springback (thus increasing the moment of inertia and decreasing springback drastically), and 2) the strain ratio arising from the normal to the yield function control the under-load anticlastic curvature in the same way that Poisson’s ratio controls anticlastic curvature in an elastic beam.

It is because of these competing effects that the Von Mises yield function works fortuitously while the Hill quadratic (Hill 1948, Hill 1950) and Barlat YLD 89 (Barlat and Lian 1989) do not. These latter yield functions, which are nearly equivalent, capture either the r-value variation with direction well but not the yield stress variation, or vice-versa. Figure 42.3 shows that only the Barlat YLD 96 (Barlat 1997) yield function matches both yield stress and plastic strain ratio variations accurately, and thus reproduces all of the aspects of material behavior necessary to understand and predict draw-bend springback.

3-D simulations with appropriate numerical tolerances and standard material laws continued to show significant variance with the experiments: 5.7° standard deviation vs. 1° experimental scatter. Two aspects of the material constitutive law were found responsible for the errors: the shape of the initial yield function and the assumption of isotropic hardening (Geng and Wagoner 2000, Wagoner et al. 2000b). When these aspects were investigated with
42.4 Draw-Bend Springback

Figure 42.3: Plastic anisotropy of aluminum alloy (Barlat et al. 1997, Shen 1999, Geng and Wagoner 2002).

independent experiments and taken into account by algorithmic development and implementation, the simulations and measurements agreed within the experimental scatter (Figure 42.2b, Curves 5 and 6).

Role of Bauschinger Effect A second aspect of the material law was found to influence draw-bend springback results significantly. Because nearly every element of material through the sheet thickness in Region 3 of the specimen, the dominant area, undergoes stress and strain reversal, the presence of a Bauschinger effect must be considered. In order to investigate the Bauschinger effect for the 6022-T4 alloy tested, a stabilized tensile test, Figure 42.4a, was constructed (Balakrishnan 1999, Geng and Wagoner 2002) using principles presented in the literature (Kuwabara et al. 1995a, b). Typical results of these tests are shown in Figure 42.4b.

Bauschinger effects of the type shown in Figure 42.4b are not reproduced well by continuum isotropic hardening models, nor by most forms of kinematic hardening (Armstrong and Frederick 1966, Krieg 1975, Dafalias and Popov 1976, Chaboche et al. 1979, Ristinmaa 1995, Jiang and Kurath 1996). Based on Armstrong-Frederick-type hardening rules (Chaboche et al., 1979, Chaboche and Rousselier 1983, Chaboche 1987, Chaboche and Nouailhas 1989, Chaboche 1991, Ohno and Wang 1991, Ohno and Wang 1993a, b, Jiang and Sehitoglu 1995a, b, Khan and Huang 1995), Geng and Wagoner (2000) introduced a variation based on a two-surface plasticity model which reproduces the strain-hardening behavior adequately. To model the permanent softening shown on reverse loading at larger prestrains, the bounding surface translates and expands according to a mixed hardening rule (Hodge 1957, Crisfield 1991). This hardening law has been implemented for plane-stress thin-shell elements in conjunction with three anisotropic yield criteria: Hill’48 (Hill 1948, Hill 1950, Mellor and Parmar 1978), Barlat’s three-parameter yield function (Barlat Yld89) (Barlat and Lian 1989) and Barlat’s Yld96 (Barlat 1997).
When this hardening law was introduced along with the Barlat YLD96 yield function, the springback (and anticlastic curvature) predictions matched the experiments within the scatter of the experiments, 1.2° (Curve 6, Figure 42.2b).

The results presented above illustrate the complex and sensitive dependence of springback phenomena on the detailed plasticity law, including anisotropy, complex hardening, and Bauschinger effect. It is therefore clear that these aspects of the material behavior must be accounted for accurately in order to obtain accurate springback predictions. However, one last and surprising aspect of material behavior that was observed with some draw-bend tests must be discussed: time-dependent springback.

**Time-Dependent Springback** The experimental results shown in Figure 42.2b were obtained for a tool radius-to-sheet thickness ratio (R/t) of 10. They showed little variation with forming rate or with time after the unloading step. However, process conditions which produce large springback angles (particularly for small R/t ratios) also produce significant time-dependent deformation for aluminum alloys after removal from the testing fixtures. Figure 42.5 shows the shapes for a single draw-bend specimen for times up to 7 months following testing.

Since first reporting these results (Wagoner et al. 1997), this phenomenon, which appears to have been previously neglected in the literature, has attracted interest in industry (Kuwabara et al.1996a, Tang 1996) and academia (Focellese et al. 1998). The impact of a part changing shape after forming has implications for replacement parts, which may meet specifications immediately after production, but then change during shipping and storage. Initial thoughts on the origin of time-dependent springback revolve around anelasticity (Zener 1948, Burton 1948).

Preliminary simulations based on a model of room-temperature creep to relieve internal stress at the end of the forming operation have shown reasonable qualitative agreement with the form of the time dependence of springback angle, Figure 42.6a. The corresponding rearrangement and relaxation of internal stress is shown in Figure 42.6b. It should be emphasized that these results are very preliminary; they make use of very simple material laws with isotropic yield, and the creep law is for pure aluminum. Nonetheless, they demonstrate that
such simulations can be carried out and that the form of the governing time dependence is unlikely to be simple.

42.5 Conclusions

Springback is an apparently simple phenomenon involving the elastic shape change of a body when deforming loads are removed. Nonetheless, the current inability to predict springback with accuracy costs manufacturers and consumers hundreds of millions of dollars per year in terms of lost production and remedial efforts. The complexities behind the simple façade of springback have been discussed.

The literature dealing with springback has been reviewed in terms of analytical methods, experiments, and numerical simulations. Much of the literature base is of limited use in predicting springback pertaining to industrial sheet forming operations. In many experiments the sheet tension has not been controlled reliably and many simulations apply to simpler situations than those encountered at the radius of a die cavity in common forming tooling.

Recent results by the author and coworkers for draw-bend springback has been summarized. Combined forming and springback simulations are shown to be much more sensitive and time-consuming than forming simulations alone, on the order of 100 times more CPU intensive. For even nominally 2D situations, 3D elements are required to take into account the anticlastic curvature that develops, and elastic-plastic unloading has been noted.

When numerical issues are addressed adequately, significant complexity of material behavior must also be considered to obtain accurate springback predictions. The material issues include the need for anisotropic yield functions that consistently reproduce both yield stress and strain ratio variation with direction, and novel hardening evolution formulations to reproduce the Bauschinger effect accurately. When these concerns are addressed, however, excellent springback predictions can be made.

The first reported observations of time-dependent springback of aluminum alloys were summarized and preliminary ideas of how this phenomenon can be modeled presented.

Acknowledgements

The author would like to thank the many individuals and sponsor agencies who carried out, assisted with the manuscript and supported the original work summarized in this chapter, as follows: Lumin Geng, Kaiping Li, David K. Matlock, Wendy P. Carden, William D. Carden, Vijay Balakrishnan, Yao Shen, Jianfeng Wang, Frederic Barlat, Michael P. Wenner, Christine D. Putnam, the National Institute of Science and Technology (ATP), Partnership for a New Generation of Vehicle (PNGV), National Science Foundation (DMR 0139045), Ohio Supercomputer Center (PAS080), Edison Materials Technology Center (CT-72), and the Ctr. Adv. Mats. Mfg. of Auto. Components (CAMMAC).

References


References


References


Quecner CA, and De Angelis RJ, 1968. Elastic springback and residual stresses in sheet metal formed by bending. Trans ASM 61, 757-768.
References

References


43 The ESI-Wilkins-Kamoulakos (EWK) Rupture Model

Argiris Kamoulakos

Certain facts behind metal rupture prediction are exposed together with the classical ways they have been approached in the past. The limitations of these approaches are highlighted and a new approach is proposed. The way this approach is manifested through the new EWK model in PAM CRASH is shown. Academic examples are presented with this new modeling capability and they are compared favorably with the existing academic literature evidence. Finally pre-industrial application examples are shown to support the potential of the new EWK model, especially in the field of automotive crash and Foreign Object Damage (FOD) in aeronautics.

43.1 Background

Predictive fracture mechanics within numerical codes has become a necessity for the virtual design of an automotive but at the same time is still very difficult to achieve. There are two main difficulties to predict material rupture:

- the inability of today’s most commonly used theoretical models in capturing the necessary physics that lie behind “crack initiation and propagation”
- the mesh dependency that all commonly used models exhibit

In order to progress in achieving predictive fracture capabilities within a numerical code to be used in industrial production, three basic issues have to be investigated:

- the material description and calibration (“the physics”)
- the numerical description (“algorithmic numerics”)
- the solution cost (for modeling and computing)

The material description issue comprises of identifying material models and stable (“intrinsic”) parameters, easy to calibrate and reasonably robust, that can represent the basic physics behind the initiation and propagation of fracture. But what are these “basic physics” and associated processes?

Engineering structures usually contain flaws or cracks consequently engineering design often requires evaluation of the maximum flaw size and operating stress level for safe operation. Large flaws and/or high stresses can lead to crack growth and ultimately to unstable propagation and structural failure. Knowledge of the “fracture toughness” (in general terms) of a material, ie. the measure of its resistance to crack growth, is required to design against unstable crack propagation.
Small-scale specimens are normally used to determine the resistance of a material to crack propagation, but measurements taken at small scale do not necessarily coincide with large-scale results. Structures that are large enough fail by “brittle” fracture (where with the term brittle we do not necessarily mean intragranular fracture). In this brittle-fracture regime, the failure stress varies inversely proportional to the square root of the (crack) size, so that larger geometrically similar structures will fail by brittle fracture at a lower average stress. Smaller similar structures will fail at a higher average stress, until a certain size is reached. Any further reduction in size will provoke the change of the failure mode from “brittle” to ductile. For small-scale testing to be successful, the size-effect scaling laws must be understood.

Linear elastic-fracture mechanics (LEFM) successfully describes the scaling law for brittle fracture. The material is treated as a linear elastic solid, and the stress field at the crack tip due to a remote stress is calculated.

The stress field is then characterized by the stress-intensity factor $K_I$. For certain geometries, the value of $K_I$ at the crack tip has been calculated in terms of the geometry and far-field stress. Thus, the applied load and external geometry of the component are connected by linear elastic theory to a value of $K_I$ at the crack tip.

A reasonable postulate is that crack growth occurs when a critical stress-intensity factor $K_I$ has been reached and it is called the plane strain fracture toughness $K_{IC}$. Since a single parameter $K_{IC}$ describes fracture initiation we would call that a “one parameter” fracture model. The advantages of the $K_{IC}$ concept for evaluating fracture toughness are:

- an analysis can be made that relates external physical measurements on a test specimen to a reasonable postulate for fracture initiation
- the method is conservative, in that it evaluates the materials resistance to fracture at worst-case conditions, i.e., those under which brittle fracture occurs.

Alternatively, the $J$ integral (originally defined by Rice) is a path-independent energy line integral for two-dimensional problems. Using the “HRR” model developed by Hutchinson and by Rice and Rosengren to describe the stress-strain field in the vicinity of a crack, Mc Clintock concluded that $J$ is a measure of the plastic stress and strain singularity near a crack tip. With such an interpretation, the field-characterizing parameter $J$ could be regarded, for the plastic case, as analogous to the stress-intensity factor $K$ in LEFM. As in the $K$ method, the applied load and the external geometry of the specimen are connected by theory to a value of $J$ at the crack tip. A critical value for crack extension, $J_{IC}$, can then be identified. Since the $J$ integral accounts for plasticity, the method permits compact tension tests on specimens many times smaller than permitted by the $K_{IC}$ method.

The $J$ method is based on a “one-parameter model” of fracture, just like the $K$ method. It overcomes a major limitation of the $K$ method by including stable crack growth with increased load. A limitation of the $J$ method is that deformation plasticity is used to describe plastic flow. This non-physical model of plastic behavior can only provide an analysis for modest crack growth and is certainly erroneous in describing the strain localization that happens after the onset of the plastic instability in a structural component.

If $J_{IC}$ is accepted as a fracture criterion, and if conditions for ductile behavior are met, larger flaws than would be permitted by the $K_{IC}$ criterion can be tolerated for the safe operation of a structure. Nevertheless, the $J$ method must be applied conservatively, because of the limitations of the analysis and because of incomplete experimental data.
43.2 The EWK Fracture Model

The EWK model of ductile fracture correctly predicts fracture initiation and propagation. Fracture occurs in the model when the time-integrated product between the equivalent plastic-strain rate and two functions (one of the “mean tensile” stress and the other of the asymmetric stress distribution) exceeds a critical value over a critical dimension.

This cumulative-strain-damage model assumes that fracture is the result of the history of strain damage to the material. It is well known that ductility increases when hydrostatic tension decreases. This behavior can be interpreted in terms of hole growth by using the theoretical models of Mc Clintock and Rice and Tracey. Their models suggest a rapidly increasing ductility with decreasing hydrostatic tension. This result is simply built into the EWK fracture model by the hydrostatic tension term (“mean tensile” stress) which accounts for the growth of holes during loadings that consist of large triaxial stress and small strain. The asymmetric strain term accounts for the observation that, after initiation, the holes can link up as a band if subsequent loading is shear.

Because the spacing and size of the inclusions control the cracking, it is not enough for this critical value to be reached at a single point; it must occur over some length $r_c$ that is characteristic of the microstructure of the material.

A fundamental assumption behind the EWK model is that the initiation or the extension of a crack can be treated as a constitutive property of materials. Fracture will then depend explicitly on the current mechanical state and past history of a local region, and will be independent of geometry and boundary conditions except as they affect the local state. Thus crack extension and crack initiation are treated alike.

The EWK model is based on the fracture model suggested by Wilkins (1999) which in turn is based on the ductile-fracture theory of McClintock. The Wilkins model contains material-dependent parameters that are linked semi-empirically to other physical properties and they are adjusted by trial and error until a single set of values is obtained that fit experimental fracture initiations in several different small size specimen geometries.

The basic definition of the Wilkins (and consequently the EWK) model is as below:

$$\frac{\Delta p}{\Delta C} = \int w_1 w_2 d\bar{\varepsilon}_p \,$$

$$w_1 = \text{hydrostatic pressure weighting term} = \left( \frac{1}{1+aP} \right)^\alpha$$

$$w_2 = \text{asymmetric-strain weighting term} = (2-A)^\beta$$

where

$$a = \frac{1}{P_{lim}}$$

$$A = \text{Max} \left( \frac{S_2}{S_3}, \frac{S_2}{S_1} \right) \quad S_1 > S_2 > S_3.$$

Fracture starts when

$$\frac{\Delta p}{\Delta C} > r_c$$

over a distance $r > r_c$. 
It is interesting to note that the above Damage function is a stress measure multiplied by strain and, in that sense it suggests energy density. Energy density, in various forms, has been used by other researchers, like Sih and Nemat-Nasser, as a fracture criterion.

### 43.3 Academic Validation

The following three academic cases (first described by Wilkins) have been tested using the solid element version of the EWK model (MAT 71 of PAM CRASH).

The first one consists of a small circular cylinder loaded under tension beyond the plastic instability point. This case was originally conceived to test the validity of the elastoplastic constitutive law used rather than to test for fracture. However, it is an excellent case also for fracture since after the onset of the plastic instability, the provoked necking induces a profound state of hydrostatic tension in the specimen. The maximum value of this hydrostatic tension appears at the centerline and consequently this is the area where voids do nucleate and grow thus provoking fracture.

Figure 43.1 below shows the state of hydrostatic stress at the neck just prior to fracture, while Figure 43.2 shows the fracture sequence as element an elimination process.

![Figure 43.1: Cylindrical specimen during plasticity law calibration (showing specimen necking).](image1)

![Figure 43.2: Cylindrical specimen during failure calibration.](image2)

Figure 43.3 shows a fatigue pre-cracked compact tension specimen, a quarter of it been modeled (ie. half symmetry and half through the thickness). This is a case of brittle fracture because a very powerful hydrostatic tension field develops at the centerline ahead of the crack and thus precipitates the rapid crack growth at the interior of the specimen (tunneling) while most of the specimen is still in the elastic regime.

In Figure 43.4 a thin notched plate is shown, again a quarter of it been modeled (ie. half symmetry and half through the thickness). This is a case of ductile fracture because the crack occurs after the plate has heavily plastified and the crack growth is not only conditioned from the hydrostatic tension but the asymmetric strain, too.
43.4 Semi-Industrial Validation

A simple validation was performed against Self Piercing Rivet (SPR) testing. The basic idea was to demonstrate that a single set of rupture parameters and without any re-calibration for mesh dependency can be used to investigate the rupture of the three basic test assemblies (Coach Peel, Lap Shear and U-Tension). More than that, one set of meshes was created for each case based on “engineering judgment” and no mesh modification and recalculation was allowed after the initial calculations were performed, in order to see the level of reliability and robustness one could have in such an approach.

In Figures 43.5a and b the comparison between a typical test and a simulation for Lap Shear are shown. The SPR is modeled as a rigid body (since it was made of steel) while the two plates with aluminum. Damaged parameters are automatically treated internally by PAM CRASH to remove the mesh dependency and provide a coherent answer within the limits of the mesh idealization chosen by the user, which in this case was 3 to 4 elements through the thickness of each laminate.

Similarly in Figures 43.6a and b we see the Coach Peel results. In all cases the assemblies were loaded to failure.
Finally and most importantly in Figures 43.7a and b we see the force displacement curves for these one-off trials. The comparison between experimental results and simulations is evident.

Figures 43.7a and b depict a typical rivet lap shear test. Note that in this simulation the formation of the rivet was also simulated, hence the lap shear test was performed with the corresponding constituents under the correct pre-stress state.

Another case of interest is a beam from Nissan subjected to a substantial impact load. This case tests the Belytchko-type shell implementation of the EWK model (MAT 171 in PAM CRASH). It is quite interesting because a shell implementation of such a criterion is not straightforward. Figure 43.9 below shows a comparison between the experiment and the
**Figure 43.8:** General rivet Lap Shear test simulation (half model shown).

Simulation. The damage and failure of the simulated model was very comparable qualitatively with the experiment and this was achieved without any specific calibration for failure regarding the mesh used, since all that is automatic and internal to PAM CRASH.

**Figure 43.9:** Nissan Beam impact testing. Experiment versus simulation. Example published at PAM Conference in 2002.

Within the European project CRAHVI where ESI Group is currently participating, various birdstrike experiments were performed by Airbus and CEAT and were simulated with PAM CRASH. Figure 43.10 shows a comparison between experiment simulations for birdstrike upon an aluminum plate for a case where the plate was not perforated by the bird but instead
it failed along the riveted sides. The simulation was reasonably close to the final result. The simulation also used the Belytschko version of the EWK model.

However, when the plate was machined at its center until its thickness was halved then a clear perforation occurred at the test. For the simulation the same input file was used but with the thickness reduced at the same area where the plate was machined. The same EWK model parameters were used. The results are shown in Figure 43.11 and confirm that the EWK constitutes a promising predictive model that can identify the different loading regimes that provoked failure.

![Plate failure under birdstrike (experiment versus simulation)](image1)

Figure 43.10: Plate failure under birdstrike (experiment versus simulation) Courtesy of Airbus and CEAT, CRAHVI EC project.

Finally full birdstrike tests were performed upon real aerofoils under speeds high enough to provoke perforation of the leading edge. A typical comparison between some experiments and simulation regarding the perforation of the aerofoils is as shown in Figure 43.12 (from two complementary views).

### 43.5 Conclusions

A “two parameter” fracture model, the EWK model, has been introduced into PAM CRASH in order to aid the reliable prediction of metal fracture in engineering simulations. It is a cumulative plastic strain damage model which includes the effects of hydrostatic tension and shear in the phenomenological treatment of fracture. It has been introduced in both solid element and shell element idealization forms. It is based on an intrinsic representation of the
physical data for fracture ie. once they are defined for a specific material, they do not need to be re-adjusted according to the specific analysis model of the user, instead they are valid for all scenarios. Academic and pre-industrial applications are very encouraging. Current work at ESI Group focuses on improving the model for a wider class of problems.

References

Damage percolation modeling provides a means to probe the role that microstructural heterogeneity plays in determining the progression or “percolation” of damage across a microstructure. Tessellation-based methods are used to characterize length scales present in second phase particle fields, acquired from 5000-series aluminum alloy sheet. Inter-particle dilational spacing (IPDS) spectra are used to identify the level and scale of particle clustering present in the particle field. A linked damage percolation–continuum damage numerical simulation is performed to simulate the damage percolation process within forming operations. The percolation code accepts detailed second phase particle fields, allowing damage nucleation and coalescence events to be modeled at the individual particle level. Continuum calculations of stress state, constitutive softening and void expansion are performed using a Gurson-based constitutive model. The linked model utilizes a Gurson-based yield surface to account for the softening effects of void damage, while the local damage development and void linkage events are modeled using the damage percolation code. This coupled approach offers a comprehensive treatment of ductile damage evolution. The model is applied to a stretch flange sheet forming operation incorporating drawbeads.

44.1 Introduction

The drive for lower emissions and higher fuel economy within the automotive industry has instigated a great deal of research into the use of aluminum alloy sheet as a lightweight alternative in the fabrication of automotive structural and closure panels. With anticipated increases in aluminum usage, there will be a corresponding increase in the level of recycled material that will lead to higher impurity content, such as Fe and Mn. The presence of Fe- and Mn-based intermetallic particles in aluminum alloys leads to ductile damage during forming processes (Gelin 1986, Tang et al. 1999, Hu et al. 2000). Experimentation has demonstrated that ductile damage can have a strong impact on formability and bendability of automotive aluminum alloy sheet (Pilkey et al. 1998, Lievers et al. 2000, Spencer et al. 2002). The nucleation, growth and coalescence of microvoids inherently promote the formation of through-thickness shear bands ultimately leading to failure (Lievers et al. 2002). Therefore, understanding damage evolution during forming of aluminum alloy sheet is necessary.

The distribution of second phase particles and associated ductile damage in aluminum alloy sheet is highly localized (Pilkey et al. 1998); thus it is questionable whether continuum-based damage models, such as Gurson’s porous plasticity model (Gurson 1977) can properly capture the details of the damage evolution process. Furthermore, ductile damage accumulates...
in a stable fashion until failure occurs abruptly through catastrophic local coalescence of voids. In continuum models, this instant is generally modeled in terms of a threshold value of bulk damage; however, this average value does not reflect local damage levels within second phase particle clusters, for example.

In the past decade, considerable research efforts have contributed to the understanding of the effect of heterogeneity of damage distribution on ductile fracture. To investigate the effect of particle/void clusters, Benson (1995) employed a 2D void cluster unit cell model. The results showed that the fracture strain decreases monotonically as the cluster radius increases and that the fracture pattern is dependent on the cluster size. Thomson et al. (1999, 2003) employed a 3D unit cell approach to investigate the effect of particle clustering. By varying the particle distribution in the cluster, three types of particle cluster are invoked. Their results show that the orientation of clusters with respect to the major principal strain dominates the void damage rates. Huang and Hutchinson (1989) studied the effect of non-uniform void distribution on shear localization and argued that shear localization is controlled largely by the most critical void cluster. Based on this conclusion, Huang et al. (1998) utilized a unit cell model with a penny-shaped particle cluster to investigate the effect of the cluster on the plastic localization strain. Their results show that larger clusters result in smaller localization strain and vice versa. Thomason (1993) modeled a 2D unit cell with a non-periodic array of voids and discovered that the resulting void growth and localization behavior is highly sensitive to the size and spatial distribution of voids. Interestingly, the results of Thomson et al. (2003) demonstrate that coalescence of voids nucleated within so-called first-order particle clusters is not a sufficient condition for catastrophic failure. Coalescence or link-up over higher-order length scales (e.g. cluster-to-cluster link-up) is required to trigger catastrophic ductile fracture.

It is worth noting that the majority of previous damage models have adopted some form of unit cell concept, which pre-supposes a periodic microstructure of materials. However, in real damaged materials, micro-defects always present non-uniform distributions. Recently, a damage percolation model has been developed (Worswick et al. 1998, 2001, Chen et al. 2003) utilizing measured second phase particle distributions as a starting point to model the damage evolution in aluminum alloy sheet under various forming conditions. Tessellated particle fields are used to capture the real spatial distributions of particles as well as nearest neighbor and particle clustering data. Unlike continuum-based models, in this damage percolation model, final failure occurs as a chain reaction with widespread void coalescence without any further straining once a number of primary particle clusters link-up to form a dominant crack.

In the current study, the damage percolation code is linked with LS-DYNA, an explicit dynamic FEM code used to account for local compliance effects due to damage development. The linked model utilizes a Gurson-based yield surface within the FEM model to account for the softening effects of void damage, while the local damage development and void linkage events are modeled using the percolation code. The model is applied to stretch-flange forming which is known to be a damage-sensitive operation. The nucleation, growth and coalescence of damage are predicted for the most critical region of the stretch-flange, using the coupled percolation damage approach.
44.2 Experiment

Figure 44.1 is a schematic of the stretch flange tooling and blank, while Figure 44.2 shows an example of a blank tested to failure at the punch profile radius. The blanks were AA5754 and AA5182 aluminum alloy sheet, with a thickness of 1 mm. The cut-out is circular of radius ranging from 88 mm to 98 mm. The punch diameter is 290 mm and the male drawbead profile has a 12 mm height and 5 mm profile radius.

In the stretch flange forming process, the outer region of the blank is clamped against the die while the main punch is pressed against the backup punch to close the drawbead. Once the tooling is closed, the punch moves downwards to form a so-called z-flange through expansion of the central cut-out. This type of operation is common in the corner of window cutouts in automotive inner panels and the drawbead is introduced to limit the rate of cut-out expansion. In aluminum alloy sheet, the repeated bending and unbending of the sheet as it flows through the drawbead can lead to damage nucleation followed by necking or tearing in the drawbead or punch profile radius region of the part. Further details regarding these experiments are given by Cinotti (2001).

![Figure 44.1: A schematic of the stretch flange tooling and blank.](image)

44.3 Material – Characterization of Second Phase Particle Fields

Sections of the two as-received microstructures were prepared for metallographic examination and image analysis. AA5754 contains an array of Fe intermetallic particles, whereas the AA5182 microstructure has a dual population of Fe and Mn intermetallic particles. For
simplicity, no attempt was made to distinguish between these different particle types during the acquisition and processing of the particle fields.

For each of the two alloys under study, a large-scale high-resolution digital image of a second-phase particle field was acquired. These rather large images are roughly $5500 \times 4250$ pixels for AA5182 and $5140 \times 4160$ pixels for AA5754 and have a resolution of 0.372 microns/pixel, corresponding to a physical size of about $2.0 \times 1.6$ mm and $1.9 \times 1.6$ mm, respectively. The constituent particles have been separated from the matrix background using standard thresholding techniques. Relevant particle and clustering characteristics have been extracted from the large-scale high-resolution particle fields through the application of a matrix erosion tessellation algorithm (Pilkey et al. 1998). The matrix erosion tessellation produced for the AA5182 microstructure is shown in Figure 44.3.

The matrix erosion tessellation algorithm also provides inter-particle dilational spacing (IPDS) data that is used to identify characteristic length scales of clustering within the particle field. Further information regarding the extraction and interpretation of IPDS data can be found in Pilkey et al. (1998). The IPDS frequency spectra for the planar view of the AA5754 and AA5182 alloy sheets are plotted in Figure 44.4. Overall, the differences in particle clustering characteristics are subtle. Both alloys exhibit a prominent first-order cluster peak at an IPDS of 2 microns, suggesting that a relatively large number of closely-spaced particles are available to promote void nucleation and interaction. Furthermore, this first-order peak is found to be marginally higher in the AA5182 alloy. A secondary peak can also be discerned at 10 micron spacing that is attributed to cluster-to-cluster spacing.

### 44.4 GTN-based FE Model

The well-known Gurson-Tvergaard-Needleman (GTN) constitutive equation (Gurson 1977, Tvergaard 1981, Tvergaard and Needleman 1984) is employed in the finite element analysis.
The GTN yield function is given as

$$\Phi = \left( \frac{\Sigma_{eq}}{\sigma} \right)^2 + 2f q_1 \cosh \left( \frac{3\Sigma_m}{2\sigma} \right) - 1 - q_1^2 f^2 = 0$$  \hspace{1cm} (44.1)$$

where \( f \) is the void volume fraction, \( \Sigma_{eq} \) and \( \Sigma_m \) are the von Mises effective stress and hydrostatic stress, respectively, and \( \sigma \) is the matrix flow stress. The “\( q \)” coefficients are “calibration
coefficients” introduced by Tvergaard (1981) to better represent void interaction effects in plastically deforming materials.

The rate of increase of void volume fraction is due to the growth of existing voids and the nucleation of new voids:

\[ \dot{f} = \dot{f}_{\text{growth}} + \dot{f}_{\text{nucleation}} \]  

Assuming the material is plastically incompressible, the growth of existing voids is related to the hydrostatic component of plastic strain, \( \varepsilon_{kk} \), by:

\[ \dot{f}_{\text{growth}} = (1 - f) \dot{\varepsilon}_{kk} \]  

The contribution of void nucleation will be material dependent. For the AA5182 and AA5754 alloys studied in the current paper, a plastic strain-controlled nucleation rule is adopted, which assumes that voids nucleate at second phase particles and there is a normal distribution of nucleation strain for the total population of particles (Gurson 1975),

\[ \dot{f}_{\text{nucleation}} = A \dot{\varepsilon}^P \]  

in which, \( \varepsilon^P \) is the effective plastic strain, and

\[ A = \frac{f_N}{s_N \sqrt{2\pi}} \exp \left[ -\frac{1}{2} \left( \frac{\varepsilon^P - \varepsilon_N}{s_N} \right)^2 \right] \]

The term \( f_N \) represents the volume fraction of void-nucleating particles while \( \varepsilon_N \) and \( s_N \) are the average and standard deviation of the strains at which particles nucleate voids. The mechanical and void nucleation parameters adopted for these two aluminum alloys are listed in Table 44.1. It is worth noting that the void nucleation strain is approximated as 90% of the elongation of uniaxial tensile specimens, based on the assumption that void nucleation dominates ductile fracture in these aluminum alloys.

<table>
<thead>
<tr>
<th>Material</th>
<th>AA5182 (1.0 mm)</th>
<th>AA5182 (1.6 mm)</th>
<th>AA5754 (1.0 mm)</th>
<th>AA5754 (1.6 mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( E ) (GPa)</td>
<td>71.71</td>
<td>71.71</td>
<td>71.71</td>
<td>71.71</td>
</tr>
<tr>
<td>( \nu )</td>
<td>0.33</td>
<td>0.33</td>
<td>0.33</td>
<td>0.33</td>
</tr>
<tr>
<td>( \sigma_0 ) (MPa)</td>
<td>120.17</td>
<td>117.34</td>
<td>87.91</td>
<td>102.78</td>
</tr>
<tr>
<td>( \varepsilon_N )</td>
<td>0.20</td>
<td>0.20</td>
<td>0.20</td>
<td>0.20</td>
</tr>
<tr>
<td>( s_N )</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
</tr>
<tr>
<td>( f_N )</td>
<td>0.016</td>
<td>0.016</td>
<td>0.014</td>
<td>0.014</td>
</tr>
<tr>
<td>( f_0 )</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

A common approach to predicting void coalescence is to specify critical or threshold levels of porosity at which coalescence is assumed to initiate (Tvergaard and Needleman 1984). This approach was not adopted in the current work; instead, the damage percolation model is
used to predict the onset of coalescence based upon local interaction of second phase particle and voids as described in the following section. The prediction of the formability of stretch flanges, using a more conventional form of the GTN model, is given by Cinotti et al. (2001). That paper also provides greater detail regarding the meshing and boundary conditions used in the stretch flange models.

44.5 Coupled damage percolation model

The damage percolation model is that proposed by Worswick et al. (1998, 2001) to handle damage evolution processes in aluminum alloy sheet. This model considers the nucleation and growth of voids at each particle within the microstructure. Coalescence of individual voids to form cracks and the subsequent growth of cracks and coalescence of cracks with other cracks or isolated voids is also modelled (Worswick et al. 2001). Particle feature data is read directly from the tessellated second phase particle fields (Figure 44.3) and comprises particle centroid coordinates, particle principal axes (and orientation), nearest neighbor list, and cluster list (particles comprising each cluster). Based upon experimental observation, individual particles are represented as ellipses with principal axes aligned with the sheet rolling and transverse directions.

To accommodate the deformation experienced by the sheet metal, the tessellated second phase particle field is mapped directed onto the finite element mesh. Due to the limited size of the tessellated particle field, only a small area of interest within the finite element mesh is selected for the mapping process, generally corresponding to the area of highest predicted damage from the GTN calculations. The mapping includes assignment of a unique “parent element” and location for each particle. In the current implementation of the model, the coupling is limited to a one-way “loose coupling” in which the GTN-based FE calculation is performed first and the damage percolation code is then run using the FE results as input data. The model is also two-dimension in nature, axi-symmetric in the case of the stretch flange models. A bi-directional coupling of the two codes is currently under development and three-dimensional calculations are possible, pending availability of measured three-dimensional particle fields.

In the percolation code, void nucleation is treated as strain controlled and as being sensitive to particle size, as outlined by Chen et al. (2003). In general, nucleation is assumed to occur preferentially at larger particles and the fraction of nucleated particles increases with strain (Embry, 1985). This treatment is consistent with the model suggested by Embry (1985), in which void nucleation “sweeps” through a particle population initiating preferentially at larger particles. It should be noted that the nucleation rule within the percolation code differs from that in the continuum-level FE simulations, although attempts have been made to harmonize the two approaches to minimize differences in predicted nucleation response. This difference between nucleation model treatments will be eliminated once a two-way coupling of the continuum damage and percolation codes is completed.

Once nucleated, voids will expand at the rate dictated by the void growth within the parent element predicted by the GTN model. Prediction of void coalescence is performed exclusively within the percolation damage code. Void coalescence is predicted using a version of the Brown and Embry (1973) ligament-to-void-size ratio criterion, as modified by Pilkey et al.
The search algorithm required to predict void coalescence utilizes the nearest neighbor list generated by the tessellation software.

Once two voids coalesce, they are treated as a single larger void or “crack”. Note that the term “crack” is used here for convenience and does not imply a sharp-tipped crack in a fracture mechanics sense. To account for the amplified interaction between voids and cracks due to the increase of void size and interaction distance when two or more voids coalesce to form a large void, groups of coalesced voids are treated as a single-larger elliptical void. The size of this elliptical void is such that its bounding rectangle encloses all voids comprising the crack. Once established, this elliptical crack can then grow and coalesce with other voids or cracks. The significant increase in size of this crack serves to introduce the amplification of the interaction distance one would expect to occur with the onset of void coalescence.

For each time step, particle, void and crack positions are updated through linear interpolation of the nodal displacements of their parent elements. Particles, voids and cracks are also rotated about their parent element centroid to account for rigid body rotation. Cracks formed by coalesced voids or cracks are assigned parent elements based upon the location of its centroid. Cracks may shift from one element to another due to coalescence of additional voids or cracks amongst different elements.

44.6 Results

Figure 44.5 is a series of plots showing the deformed cross-section of the stretch flange for a 1.0 mm AA5182 blank with a 88 mm radius cutout.

Bending/unbending effects lead to the worst damage situation for the elements initially located immediately above the top of the drawbead. Accordingly, this region was selected for subsequent percolation damage simulation. Figure 44.6 is a plot of this “region of interest”, comprising the blank, male drawbead, and particle field overlaid on the mesh.

Figure 44.7 shows the predicted porosity history from the GTN continuum calculation for the element at the bottom left of the area of interest. Figure 44.8 shows the corresponding percolation calculation at several stages during the deformation process. The particles are shaded black. The approximate size of a crack is represented as a grey ellipse sized to fit within the bounding box encompassing all voids within a crack. This ellipse corresponds to the “crack size” used to determine crack-crack coalescence.

The porosity remains low as the drawbead closes (Figure 44.7, 44.8a). Voids nucleate at larger second phase particles on the outside of the bend radii and primary coalescence occurs amongst clustered voids. The first sharp increase in porosity occurs as the material is pulled off the male drawbead and approaches the profile radius at the drawbead exit (Figure 44.8b, c).

As the region of interest exits the drawbead, the upper elements experience compressive bending while lower elements are placed in tension. Figure 44.8(d) shows the material just before the state of coalescence becomes critical, after which coalescence occurs throughout the lower elements. Coalescence of voids in two adjacent elements results in a large crack over 10 neighboring void clusters (not shown). After this event, macro-cracking through void link-up across the particle field leads to the catastrophic failure without additional straining.

In order to examine the effect of cutout radius on the predicted formability of stretch flange, six different cutout radii, varying from 88 mm to 98 mm, were modeled using coupled
Figure 44.5: A series of plots of deformed geometry for the stretch flange: (a) initial position of blank; (b) after drawbead closure; and (c)–(g) at punch depths of 13, 25, 37, 46 and 50 mm.

Figure 44.6: Close-up view of the undeformed finite element mesh in the region of highest void growth, located above the male drawbead, and the particle field as mapped onto the mesh.
calculations for the two different aluminum sheets. Predictions of limit punch depth for each cutout are plotted as a function of cutout radius in Figure 44.9. Interestingly, the formability of AA5754 is well above that of AA5754 sheet for the same cut-out size. This is attributed to the lower initial particle area fraction and more uniform particle distribution in AA5754. Experimental validation of the predictions is currently underway.

44.7 Discussion

Coupling of the damage percolation model with the GTN-based FE model has allowed consideration of local micro-heterogeneity in damage calculations for the stretch flange forming process. Cracks initiate through coalescence between two or three nucleated neighboring voids in the same cluster. Further coalescence of voids or cracks is limited to about 10 neighboring void clusters after which catastrophic failure initiates.

The fracture behavior of the stretch flanges depicted in Figure 44.8 is phenomenological in nature, since the results are extremely sensitive to the choice of void interaction and nucleation criteria (Worswick et al. 2001) and the choice of area of interest may greatly affect the predictions. In addition, particle fields will vary in different samples. However, the present model is beneficial in interpreting the nature of crack progression through the thickness of the blank during the forming process and may serve as useful tool in evaluating or ranking different microstructures in alloy development.
Figure 44.8: Sequence of deformed particle field and corresponding blank-tooling positions.

Figure 44.9: Predicted limit punch depth for 1 mm AA5754 and AA5182.
Acknowledgement

Funding from the Natural Science and Engineering Research Council of Canada and ALCAN International Limited is gratefully acknowledged.

References

Benson D.J., 1995. The effects of void cluster size on ductile fracture, Int. J. Plast., 11, 571-582


45 Structure Damage Simulation

Dirk Steglich

45.1 Introduction

Nowadays one of the main ambitions in designing structures and components is to avoid their failure. After stress and strain concepts based on continuum theory were developed in the 19th century and used from then on, “fracture mechanics” gained importance in the second half of the 20th century. It was – and still is – treated separately from continuum mechanics, because the underlying assumption of a macroscopic discontinuity (a crack) in a material body at least complicates the solution of field equations. However, sophisticated assessment methods have been developed to accomplish the demands for reliability and safety, e.g. the so-called “global approach to fracture”, which is either based on global parameters such as the K-factor (Irwin, 1962) or the J-integral (Rice, 1968), or the more recently introduced two parameter approach J-Q (O’Dowd and Shih, 1992). Although they are successfully used to compute safety margins for cracked structures, they are unable to capture experimentally observed geometry effects, see also Chapter 32 of this book.

On the other hand, the micromechanical character of damage – deterioration in the material which occurs prior to failure – was first accounted for by Kachanov (Kachanov, 1958) in 1958 by introducing an internal field variable. Later, “Local approaches” and “micromechanical modeling” of damage and fracture (Beremin, 1983, 1981; Lemaitre, 1986) have found increasing interest and application. The general advantage, compared with classical fracture mechanics, is that, in principle, the parameters of the respective models are only material and not geometry dependent. Thus, these concepts guarantee transferability from specimens to structures over a wide range of sizes and geometries and can still be applied when only small pieces of material are available which do not allow for standard fracture specimens. It is not even necessary only to consider specimens with an initial crack as, of course, initially uncracked structures will also break if the local degradation of material has exceeded a critical state. The identification and determination of the “micromechanical” parameters usually require a hybrid methodology of combined testing and numerical simulation. Different from classical fracture mechanics, this procedure is not subject to any size requirements for the specimens as long as the same fracture phenomena occur.

Generally, “local approach” is based on finite element simulations of stress and strain fields. So-called locally coupled models (Beremin, 1983) are based on a rather simple elastic–plastic calculation with a post-processing treatment and can be used to derive crack initiation criteria. Fully coupled models are based on continuum damage mechanics and account for the softening effect induced by cavity growth (ductile fracture), micro cracking or decohesion between aggregates. These models use internal variables (scalar, vectors or tensors) to
quantify the amount of damage at a material point. Beside this, an evolution equation for the damage variable has to be formulated and a correlation between damage variable and stresses and strains has to be established. This can be done via a pure phenomenological procedure by postulating thermodynamic potentials (Lemaitre, 1987; Rousselier, 1987) or via a homogenization method (Eshelby, 1957; Hill, 1965; Hutchinson, 1976), which may account for the micromechanical mechanisms of damage (Rice and Tracey, 1969; Gurson, 1975). In the first case the damage variable is solely used to describe the influence of deterioration on the stresses and a physical interpretation of the variable can only be given as a vague approximation. In the second case a representative volume element (RVE) has to be defined and constitutive relations have to be derived on a “meso-scale”. An RVE is a material volume which is statistically representative of a material neighborhood of that material point. This micromechanical modeling encounters the problem that the material is not uniform on the microscale but consists of various constituents with differing properties and shapes. Micromechanics is a general methodology of expressing continuum quantities in terms of the parameters which characterize the microstructure and properties of the microconstituents of the material neighborhood (Nemat-Nasser and Hori, 1993). A material element therefore has its own complex and evolving microstructure. This concept has the potential to cover the physical effects of damage, such as particle cracking, matrix–particle debonding or intergranular decohesion. Unfortunately, with increasing complexity of the RVE under consideration, the likelihood of finding an analytical form for the constitutive behavior decreases. In practice, relatively simple unit cells are used, e.g. a plastically deforming matrix containing a spherical void (Koplik and Needleman, 1988) for the description of ductile damage or a bundle of fibers with statistically distributed strength to investigate fiber or brittle fracture (Rosen, 1964).

In the following, emphasis is put on the ductile damage and fracture of metals under quasistatic loading conditions. The underlying micromechanical mechanisms are the nucleation, growth, and coalescence of micro voids with significant plastic deformation. Figure 45.1 shows as an example the fracture surface of a Al−Al₃Ti material with clearly visible dimples, which stem from voids nucleated at inclusions during the loading process.

### 45.2 Plastic Potentials and Porosity

The constitutive equations which are used to describe ductile fracture processes are based on relatively simple models of the growth and coalescence of voids (Rice and Tracey, 1969; Koplik and Needleman, 1988). On the meso-scale, a homogenization process is applied which relates the microstructural properties and processes in a systematic way to field quantities of continuum mechanics. In the case of an initially spherical void in a plastically deforming matrix, the void volume fraction, \( f \), which is defined as the ratio of the total volume of all cavities to the volume of the body is introduced as an internal scalar variable to characterize the damage. In fact, this definition allows the damage variable to be denoted as “porosity”, in accordance with the usual meaning of the word in common speech. The presence of a void results in a reduction of the stresses to be transferred in a material volume, and therefore the plastic flow potential has to be a function of the porosity,

\[
\phi = \phi \left( \sigma_{\text{eff}} - R(\varepsilon_{\text{eq}}^p), f \right),
\]  

(45.1)
45.2 Plastic Potentials and Porosity

where $\sigma_{\text{eff}}$ is an effective scalar stress which is a function of both the macroscopic stress tensor $\sigma_{ij}$ and the porosity, $f$. $R$ is the yield stress of the undamaged (matrix) material and $\varepsilon^\text{pl}_{eq}$ an effective plastic strain representative of the matrix hardening.

As a consequence of mechanical loading, the void may grow (in the case of tensile loading) or shrink (in the case of compression loading). The evolution equation for the porosity is derived from the assumption, that volume changes in the plastic regime are solely due to changes of the porosity (conservation of mass). From this, void growth has to follow the differential equation

$$\dot{f}_{\text{growth}} = (1 - f)\dot{\varepsilon}^\text{pl}_{kk},$$

(45.2)

and is therefore proportional to the plastic volume dilatation rate, $\dot{\varepsilon}^\text{pl}_{kk}$.

Models for porous materials proposed by Gurson (Koplik and Needleman, 1988; Tvergaard and Needleman, 1984) and (Rousselier, 1987) both use this methodology. Whereas the first one was derived from a micromechanical description of the porous material, the second one was developed based on thermodynamic considerations. In both cases, damage is represented by the porosity, $f$. Although the origin of both models is different, they can be described using a unified framework. $\sigma_{\text{eff}}$ is defined by the following equations in the case of
the Gurson–Tvergaard–Needleman model (GTN) and the Rousselier model, respectively:

$$\Phi = \frac{\sigma_{eq}^2}{\sigma_{eff}} + 2q_1 f^* \cosh \left( \frac{q_3 \sigma_{kk}}{2 \sigma_{eff}} \right) - 1 - q_3 f^{*2} = 0$$ (45.3)

$$\Phi = \frac{\sigma_{eq}}{(1 - f)\sigma_{eff}} + \frac{\sigma_1}{\sigma_{eff}} fD \exp \left( \frac{\sigma_{kk}}{3(1 - f)\sigma_1} \right) - 1 = 0. \quad (45.4)$$

where $\sigma_{eq}$ is the von Mises equivalent stress and $\sigma_{kk}$ the trace of the stress tensor. $q_1$, $q_2$, $D$ and $\sigma_1$ are model parameters which are assumed to be constant. $f^*$ is a function of the porosity $f$ which was introduced on a purely phenomenological basis to represent void coalescence for accelerated damage occurring with the onset of coalescence of voids when a critical void volume fraction, $f_c$, is reached (Tvergaard and Needleman, 1984):

$$f^* = \begin{cases} f & \text{for } f \leq f_c \\ f_c + K(f - f_c) & \text{for } f > f_c \end{cases} \quad K = \frac{f^*_u - f_c}{f_f - f_c}. \quad (45.5)$$

According to this model a crack appears if the final void volume fraction, $f_f$, is reached. In this case the material loses its stress carrying capacity and the damage variable achieves its ultimate value, $f^*_u = 1/q_1$. In the case of the GTN model (Equation (45.3)) the definition of $\sigma_{eff}$ is implicit whereas it is explicit in the case of the Rousselier model (Equation (45.4)). An extensive description of both models is given in Besson et al. (2001).

One of the challenging modeling problems in the theory of ductile fracture is concerned with the nucleation of microvoids at the sites of inclusions and second-phase particles in a plastically deforming matrix. An empirical approach for the nucleating part of void evolution was proposed by Needleman and Rice (1978). Chu and Needleman (1980) later suggested a normal distribution for the void nucleation rate:

$$\dot{f}_{nuc} = \frac{f_n}{s_n \sqrt{2\pi}} \exp \left[ -\frac{1}{2} \left( \frac{\varepsilon_{pl} - \varepsilon_n}{s_n} \right)^2 \right] \dot{\varepsilon}^p. \quad (45.6)$$

Usually it is assumed that the (total) increment of the volume fraction, $\dot{f}$, consists of two parts, one due to void growth (Equation (45.2)) and one due to the nucleation of voids, Equation (45.6).

In rate independent plasticity, the current flow stress of the matrix, $R(\varepsilon_{eq}^p)$, depends on the accumulated plastic strain alone. This describes strain hardening for slow deformation processes. If the damage model is to be extended to dynamic processes, the effect of strain rate on plastic hardening and the coupled effect of thermal softening have to be considered and $R(\varepsilon_{eq}^p, \dot{\varepsilon}^p, \Theta)$ is introduced, instead (Sun et al., 1994). Extensions for kinematic and combined isotropic–kinematic hardening exist (Besson and Guillemer-Neel, 2003), and special formulations for the yield function to account for a ratcheting effect of the porosity in the case of cyclic loading were proposed (Leblond et al., 1995). For the simulation of the deep drawing process of sheet metal, the anisotropy of deformation in combination with isotropic damage is accounted for by Seibert (1994) and the anisotropy of damage is considered via non-spherical voids in the model proposed by Gologanu (Gologanu and Leblond, 1993; Gologanu et al.,
45.3 Model Parameter Identification

For practical application the common way is to implement the respective material model into a Finite–Element code. Some commercial FE-packages (e.g. ANSYS and ABAQUS) provide special interfaces to allow the user to link a routine containing any material model. Although some limitations apply, this modus operandi combines the flexibility of USER-materials with high efficient solvers and convenient postprocessing features of the established programs.

45.3 Model Parameter Identification

The approach of continuum damage mechanics is a way to overcome the numerous problems of size and geometry dependence of the characteristic parameters used in conventional fracture mechanics. There are also problems and open questions with respect to this approach. There is still no general proof that, and under which conditions, a unique solution for the set of model data exists. This may look like a mathematically interesting problem rather than an engineering one. But it obviously affects the question whether or not the set of data may be transferred to any structural component in order to assess its structural integrity under service load conditions. For the use of damage models in practice, it is important to have a guideline for the determination of model parameters using input data from standard mechanical tests, metallographic investigations and fracture tests, but so far no generally accepted recommendations exist for how to identify and determine these parameters. They are obviously not independent of each other, see Needleman and Tvergaard (1987), Brocks et al. (1994), Xia and Shih (1995), Ruggieri et al. (1996), Bernauer and Brock (2002). In general, as many parameters as possible should be taken from literature on similar materials, and kept constant while the remaining are varied in the numerical simulations to fit experimental data which significantly reveal failure events. In the following some experiences concerning the determination of parameters of the GTN-model are given and discussed. Some of them have fundamental character and can therefore applied to every model in the “ductile-damage-family”.

The GTN model described above includes a total of nine parameters. Three parameters \( f_0, f_c, f_f \) describe the evolution of damage up to coalescence and final failure, three \( \varepsilon_n, f_n, s_n \) are used to model void nucleation, and the three remaining \( q_1, q_2, q_3 \) characterize the yield behavior of the material. Furthermore, the strain hardening behavior of the (undamaged) matrix material \( R(\varepsilon_{pl}^{eq}) \) is an unknown function.

Commonly it is assumed that the stress–strain curve of the matrix material can be directly calculated from the response of the “compound”–material (matrix and voids) in a tensile test. In fact this can be done without any error as long as the void volume fraction is small and hence the difference between “pure” matrix behavior and porous material is negligible. If this condition is not met, the porosity causes a significant reduction of the yield stress and an iterative fitting procedure with the aid of a damage model has to be used. In this case at least the model parameters affecting the early yielding behavior are needed, which complicates the fitting procedure to a certain extent.

The initial void volume fraction, \( f_0 \), is often regarded as a microstructural feature and can, hence, in principle be determined from metallurgical observations. This relation could be verified most easily for nodular cast iron, for instance, where void nucleation occurs by decohesion of the graphite particles from the ferritic matrix (Steglich and Brocks, 1998). For
many steels it can be assumed that decohesion is the dominant damage mechanism, and so the fraction of particles determined by metallurgical investigations can be set equal to the initial void volume fraction, \( f_0 \). The relation between \( f_0 \) and the volume fraction of second phase particles is much more complex for materials having more than one population of inclusions and for void nucleation by breaking of particles. Ruggieri et al. (1996) state that for all ferritic steels studied thus far, \( f_0 \), ranges from \( 10^{-4} \) to \( 4 \times 10^{-3} \). Void coalescence is described on a purely phenomenological level by accelerated damage evolution, see Equation (45.5). For a given \( f_0 \) without considering void nucleation, the parameter \( f_c \) can be determined by cell model calculations, see Koplik and Needleman (1988), Brocks et al. (1995). In addition, the onset of macroscopic fracture in a tensile test coincides with the beginning of void coalescence in the first finite element. On the macroscopic level this is associated with a sudden drop of the load. Fitting the numerical results to the experimental data at this point has therefore become a common technique to determine \( f_c \). \( f_c \) interacts with the initial void volume fraction, \( f_0 \), and the volume fraction of void nucleating particles, \( f_n \). Hence, the effects are difficult to separate in the numerical simulations, which speaks in favor of neglecting one of the two contributions if this is consistent with microstructural observations. The “damage acceleration factor” after coalescence, \( K \), in Equation (45.5) – which can be computed from the void volume fraction at final fracture, \( f_f \), and \( f_c \) – is of more numerical significance for crack extension simulations and should be taken as high as numerical convergence can be assured. Realistic values range from 2 up to 9.

A question which may arise is whether the analysis should consider void nucleation or start with an initial void volume fraction without accounting for nucleation, or both. It was shown (Steglich, 1999) that for an Al – Al\textsubscript{3}Ti-compound material, the fracture resistance of a SE(B)-specimen can be modeled using two different sets of parameters – one considering an initial volume fraction, the other assuming that voids nucleate at particles by particle cracking. The latter was accounted for via the nucleation function, Equation (45.6), and the respective parameters were taken from micromechanical analyses. In this case, the use of the GTN-model is “micromechanically” motivated, and \( f_n \) as well as \( \varepsilon_n \) bear its mechanical meaning: \( f_n \) as the volume fraction of voids to be nucleated and \( \varepsilon_n \) as the mean nucleating strain. The standard deviation of nucleating strain, \( s_n \), may then be chosen to reflect realistically the nucleation rate. The values of 0.3 and 0.1 for \( \varepsilon_n \) and \( s_n \), respectively, as often used in literature (Tvergaard, 1982a,b; Needleman and Tvergaard, 1994), seem to be an upper limit for engineering materials.

If the strain controlled nucleation rate (\( \dot{f}_{nuc} \)) considered is high, its contribution to the total increase of the volume fraction may become dominant and regions of large void growth coincide with regions of high plastic strains. As a consequence, strain controlled nucleation induces damage in the plastic “wings” formed ahead of a crack tip and thus favors a slanted fracture mode (Besson et al., 2001).

The parameter \( q_1 \) may depend on the hardening of the matrix material and is commonly set to 1.5. Perrin and Leblond (1991) calculated by a self consistent approach that in the case of small porosity \( q_1 \) equals \( 4/e \approx 1.47 \). Furthermore, it is often assumed that \( q_2 = 1 \), which generally appears to be a good choice for structures having a comparatively high thickness. In the case of thin structural components like sheet metal, where the stress triaxiality is close to plane stress, it may become necessary to increase the value of \( q_2 \) in order to boost the
influence of triaxiality and damage on the yield surface (Chabanet et al., 2003). For $q_3$ most of the authors follow the recommendation to set this equal to $q_1^2$.

For the Rousselier model in its basic form (Equation (45.4)) two parameters have to be determined: $\sigma_1$ and $D$. Following Rousselier (1987) one should set $D = 2$ and $\sigma_1 = 1/3(R_p + R_m)$ with $R_p$ and $R_m$ being the yield stress and ultimate strength of the matrix material, respectively. If the model accounts for additional void nucleation, the above mentioned holds analogously.

45.4 Strain Softening, Damage and Lengthscale

In the theory of finite elements it is considered that a numerical solution converges to the solution of the set of partial differential equations with increasing number of finite elements used for the discretization of a body. This means – practically speaking – that the accuracy of a numerically obtained solution increases with mesh refinement and the user has to judge whether the fidelity of the solution obtained (e.g. a stress field) is sufficient or not. This holds in the case of an elastic or strain hardening material. The results obtained with models showing strain softening do not converge in the above mentioned sense. Localization effects arise for any softening materials which lead to mesh dependence of the results in finite element calculations. This becomes especially significant in crack problems and can be illustrated by the following consideration:

The mechanical work dissipated per increment in one finite element having the dimensions $w$ (width), $b$ (thickness) and $h$ (height) by inelastic deformation results from

$$\Delta U_{diss} = \int_{\Delta V}^{t+\Delta t} \int_{t}^{t+\Delta t} \mathbf{T} : \dot{\mathbf{e}}^{pl} \, dt \, dV = \bar{u}_{diss} \Delta V$$

with the element volume $\Delta V = wbh$. The energy release rate per crack extension increment, i.e. one element width, $w = \Delta a$, is

$$\Gamma_0 = \frac{\bar{u}_{diss} \Delta V}{\Delta A} = \frac{\bar{u}_{diss} \Delta V}{b \Delta a} \sim h.$$  \hspace{1cm} (45.8)

which is the specific “work of separation”, see Siegmund and Brocks (1999). By increasing the element height, more mechanical work is dissipated – or energy “released” – per increment of crack extension. This affects, of course, the global behavior of a structure. Figure 45.2 illustrates this by comparing calculations of the force–CMOD- and the crack propagation behavior of a C(T)-specimen with varying element height (namely 150 $\mu$m, 70 $\mu$m and 35 $\mu$m). The larger the elements are, the higher the maximum load and the steeper the fracture resistance curve. As the results are mesh size dependent, the element size becomes an additional material parameter. Despite the general agreement that the element size has to be related to the microstructure of the material, namely the average spacing of void nucleating particles, the actual physics of the void nucleation process do not allow for a simple and direct correlation of material and model length scales. Several authors have introduced the term “computational cells” for the elements in the ligament, see e.g. Xia and Shih (1995), Xia et al. (1995),
Ruggieri et al. (1996), Faleskog et al. (1998), to point out the fundamental difference from the common concept of finite elements which are supposed to be numerical entities solving a boundary value problem but having no physical significance determining their size. The dominant effect of the “cell size” on the $J_R$-curve can be seen in Figure 45.2. From this, the necessity of introducing a characteristic length or volume is evident. The element width, $w$, affects the averaging of stress and strain gradients at the crack tip but plays a minor role for the $J_R$-curve. The correlation between length scales of the material and the respective model is established via $\Gamma_0$ and, hence, depends on the hardening behavior, but also on the type of finite elements used, linear or quadratic and their integration order.

![Figure 45.2: Results of simulations of C(T)-specimens behavior showing the influence of finite element height on the global response.](image)

The simplest way to account for an internal length is to control the mesh size and to correlate it with a (measurable) micromechanical feature, e.g. mean inclusion spacing, particle diameter or cluster spacing. As already mentioned, the element size is not “universal” but depends on the element formulation. For steel and aluminum, reasonable results were achieved with a ratio of nearest neighbor distance of particles and element height equal to 4 up to 6 (Steglich and Brocks, 1998).

In order to minimize mesh-dependence and to account for the influence of a material lengthscale on the material behavior, non-local extensions of damage models for rate-independent material behavior have been proposed, e.g. (Tvergaard and Needleman, 1995; Ramaswamy and Aravas, 1998a,b). The idea of non-local formulations is that the evolution of the void volume fraction at a material point depends on its variation in a finite neighborhood of this point. By defining the size of the “neighborhood” a lengthscale is introduced. Results of
simulations show localization patterns which may spread over more than one row of elements. As a consequence, the dissipated energy is not a function of the element dimension and the global response does not depend much on the element size (Reusch et al., 2003), as long as the preferred size is smaller than the material length.

45.5 Hints for Application

In this section some tips for users of damage models are given. They claim not to be comprehensive – they should help the user to overcome some of the problems encountering with the analysis.

Meshing

Generally, two different types of an analysis can be performed. The simpler one shows qualitatively at which points in a structure highly damaged regions occur, the more complex one should provide quantitative information about the crack propagation behavior. For the first type no specific meshing technique is required, as long the user keeps in mind that the stress triaxiality at sharp notches triggers void growth. Therefore, a careful meshing of notches is essential. For the modeling of crack propagation, a mesh as depicted in Figure 45.3 is required. It shows a model of a C(T)-specimen and uses twofold symmetry. The specimens ligament is meshed using brick elements of constant size in order to ensure the same circumstances for damage evolution at any point in the ligament. Crack propagation will most likely take place in the symmetry surface and is limited to one layer of elements (or integration points). To save computation time the element size should be increased outside the ligament. Singular elements, often used to capture crack tip blunting before initiation in the analysis of stationary cracks, may not be used for the simulation of crack propagation.

![Figure 45.3: Typical 3D-mesh used to simulate crack propagation.](image-url)
Element formulation

Depending on the problem considered, different “solid” elements may be used in the model: axisymmetric, plane or 3-D-solids. Whereas the result of the simulation of a tensile test is not affected by the choice of either a 4-node or an 8-node axisymmetric element, the situation for predicting crack extension is different. Generally, field quantities in the process zone ahead of the crack tip are neither plane strain nor plane stress, which makes it necessary to use a 3-D model. Nevertheless, in many applications good results based on plane strain simulations were achieved (e.g. for specimens having side grooves). Plane stress (or shell) elements can not be recommended to use together with damage models: the triaxiality is limited to 2/3 and therefore void growth will be underestimated. Nevertheless, some authors use shell elements for the simulation of deep drawing in sheet metal and take damage evolution to identify critical regions in the sheet.

Numerical strategy

In almost every realization of a user-defined constitutive model within a commercial FE-code, parameters affecting the integration procedure, equilibrium iteration and time stepping are used to control the simulation. The manuals of the respective program provide information – sometimes more than the user is able to handle. In case of convergence problems the user should first check whether the model parameters used have a physical significance or are pathological. A simulation based on a Dirac-function in the void nucleating part will definitely not work! In some cases it appears to be appropriate to decrease the time increment length compared to the default values in order to overcome convergence problems. An increase in the number of possible equilibrium iterations within one time increment and/or to loosen the convergence criteria (e.g. the force tolerance) for equilibrium may also be indicated. The latter should be done with great care, because the tolerances might affect the (global) solution.

“snap back”

a numerical problem caused by a “physical” effect which appears at crack initiation is the so-called “snap back”. When the failure criterion \( f^* = f_u^* \) is reached in the element ahead of the crack tip, its stiffness approaches to a value close to zero. In this moment a part of the elastic energy stored in the whole structure is released and causes a displacement jump in the nodes at the crack tip, which takes place instantaneously. As a consequence, equilibrium in the respective increment is difficult to achieve. Analogous to an experiment with a testing facility of high compliance (and so high elastic energy) it is difficult to get a crack grow in a stable manner. In the case a simulation stops at crack initiation, the user may check if Young’s modulus has in impact on the stability of the respective simulation – the less elastic energy is stored in the structure, the more easily stable crack growth can be obtained. Therefore, a high elastic constant will favor the stability of the solution. For a test, a drastic increase of Young’s modulus will do. In practice, of course, Young’s modulus has to be chosen according to the material. To overcome the convergence problems, either the displacement of the crack tip nodes has to be controlled, or a viscous (velocity dependent) material law can be introduced.
References


N. O’Dowd and C. Shih. Family of crack-tip fields characterized by a triaxiality parameter - i. structure of fields. *J.
References

46 Microstructure Modeling using Artificial Neural Networks

Hans-Ulrich Löffler

Within process automation systems, artificial neural networks are well-established as supplements to physical models. Wherever possible, physical models are used to describe the process. However, to attain the best possible model accuracy, the models have to be adapted to measured data. In many applications artificial neural networks turned out to be rather effective for such adaptations, with respect to both accuracy and computing time. There are different ways to use neural networks for process modeling: model parameter networks, model correction networks, or even black box networks that represent models of their own.

After an introductory survey of the various applications of neural networks for process automation tasks, microstructure modeling in hot strip rolling using both physical models and artificial neural networks will be described in detail. In the context of microstructure modeling the physical phenomena like precipitation, recrystallization, grain growth and phase transformation have all been studied in depth. Yet establishing the link of these phenomena to the mechanical properties (like tensile strength or yield strength) of the material is still a challenging task. A hybrid model to predict the mechanical properties of hot rolled steel will be presented here. The model comprises a comprehensive physical microstructure model as well as a neural network for linking the chemical composition and process parameters with the mechanical properties of the hot rolled steel strip.

Results obtained at the rolling mills of two German steel companies as well as from several other hot strip rolling mills show good agreement between measured and calculated values for a wide range of steel grades. Integrated into the process automation system of hot strip mills, the model can be used to reduce the number of tensile tests required and to save time by being able to release the coils immediately after coiling. The model may also be employed to improve product quality by optimizing process parameters like the coiling temperature so that target mechanical properties of the material being rolled are obtained.

46.1 Introduction

Since their emergence in the mid 1980’s, artificial neural networks have been applied to almost any branch of science and technology. After a big enthusiasm in the beginning, where under the heading of ‘artificial intelligence’ neural networks were claimed to model the human brain, applications nowadays are somewhat more down to earth: pattern recognition, empirical modeling, and adaptation/control. We will focus on their use within process automation systems,
i.e., for empirical modeling, adaptation and control purposes. More general descriptions can be found elsewhere (e.g., Hagan et al. 1996; Zell 1997; Zell 1999).

In today’s hot rolling mills the material enters the mill as a slab of about 250 mm thickness, which is schematically depicted in the lower part of Figure 46.1. In a so-called ‘conventional’ hot rolling mill the slab is usually reheated in a furnace to about 1473–1573 K (1200–1300°C), which makes plastic deformation of the material possible (respectively easier). Complemented by edgers that enable to set the width of the final strip product, a (usually reversing) roughing mill reduces the thickness of the strip to about 45 mm in typically 5 to 9 passes, with the length of the strip being increased accordingly. In a continuous finishing mill the strip thickness is further reduced to about 1 to 12 mm. In the finishing mill the quality of the final product with respect to geometric properties (thickness, width, profile, flatness, ...) can be ensured by shaping the roll contours load, e.g. by roll shifting, roll bending, and others. In addition, the microstructure of the material, e.g. ‘fine-grained’, can be achieved by selecting appropriate temperature and deformation sequences. Finally, a cooling section cools the strip from its finishing temperature (at the finishing mill exit) of about 1173 K (900°C) down to the coiling temperature (in front of the coiler) of 773 K (500°C) to 973 K (700°C). Then the coils are transported to the coilyard, where they cool down to room temperature within two to three days.

The essential processes occurring in steel mills are characterized by the often non-linear interaction of many influences. The processes must be made reproducible and sufficiently ac-
curate to meet the demanding requirements of the customer with regard to product quality and price. As shown in Figure 46.1, the basic automation system takes cyclic measurements of a number of process parameters (like material thickness and temperature). On this basis most of the open-loop and closed-loop control actions are performed. However, these measured data are usually available only at certain locations, e.g. in front of and behind the finishing mill, but not within. When a new strip enters the mill, a model-based setup calculation has to determine the correct setup values for certain process parameters, e.g. the screw-down position or the angular roll velocity. Thus, targeted physical parameters like thickness, profile, flatness and strip temperature are met. This setup calculation (also called ‘pre-calculation’) is performed repeatedly with increasing accuracy as additional measurement information becomes available. Precise models are especially needed to meet the required properties at the strip head, where no relevant measured data are known in advance. A so-called ‘post-calculation’ or ‘feedback calculation’ recalculates the model with all the input data now available as measured values (as opposed to the setup values in the ‘pre-calculation’). By comparing the model output with the measured data, an adaptation of the models, i.e. a correction of model parameters like heat transfer between strip and work roll, can be performed. Deficiencies in either the models or the measured data can be amended by so-called ‘correction networks’ that ‘learn’ the corrections and compensate for the errors. But even within the models themselves, neural networks can play an important role in representing complex model parameter dependencies. An example is the frictional coefficient in each roll bite as a function of the circumferential work roll speed, the amount of lubrication, the roll surface roughness and the reduction.

Nowadays not only the physical dimensions of the produced strip, but also its mechanical properties like tensile strength, yield strength or toughness must be met with increasing accuracy in order to stay competitive. Typically, the product quality is ascertained by taking samples from a certain proportion of all strips. These samples undergo expensive and time-consuming tensile tests after which the coils are released if they meet the customer’s requirements. In addition, quality departments are responsible for seeking ways of further improving mechanical properties, which is done by carrying out lots of experimental work. To overcome these problems, models to predict the mechanical properties of hot-rolled steel strip have been developed. It is obvious that models for calculating such properties must somehow deal with the microstructure of the steel strip. In other words, microstructure is key to the mechanical properties. Probably for this reason, numerous microstructure models can be found in the literature (see e.g., Sellars and Whiteman 1979; Hodgson and Gibbs 1992; Gladman 1997, Ginzburg 2003). They characteristically work for certain steel grades or groups of steel grades and describe phenomena such as precipitation, recrystallization, grain growth and phase transformation. Superposition of different strengthening mechanisms (all dependent on the microstructure) then leads to the mechanical properties as shown, for example, by Düfert et al. 1992, Herman 1999, or Gladman 1997. Especially when, like in the case of the steel industry, the mechanical properties of a huge variety of steel grades must be inferred from the chemical composition and the manufacturing process, the accuracy of the superposition principle is often poor. As an alternative, one can employ a neural network to link the microstructure to the mechanical properties, a procedure which has been applied quite successfully by several groups (e.g., Löffler et al. 2000; Linkens et al. 2000; Yang et al. 2000). We combined both approaches to form a joint microstructure model and neural network system called ‘MicroStructureMonitor’.
The specific benefits of such a system are the following: a reduced need of hot strip sampling; a faster release of the coils after rolling if the steel strip meets customer specifications; the possibility of optimizing the production process (e.g. the cooling schedule) with respect to final product quality and overall production costs. For instance, it is imaginable to alter the production process such that steel strip is produced that contains reduced amounts of costly alloying elements and yet exhibits the original mechanical properties.

The first part of this paper describes the application of neural networks in process simulation. Their applicability in the context of microstructure modeling is shown in the second part, where a microstructure model consisting of a physical model combined with neural networks is demonstrated. Subsequently, results obtained from different hot strip mills are presented: a comparison of measured and calculated values for the mechanical properties of low-carbon manganese steel grades and high strength low alloy (HSLA) steel grades shows a suitable accuracy of the microstructure monitor. In the depicted example for tensile strength, a range from below 300 MPa up to about 850 MPa is covered. Finally, a conclusion summarizes the paper.

46.2 Artificial Neural Networks in Process Simulation

The key features of a rolling mill process automation system are reliability and accuracy. Neural networks are known for the capability of accurate modeling of large data sets. But, as for all empirical models, the validity of a neural network is restricted to the data range, the network has been trained with. In process automation safe production is crucial, even if the model is supplied with new or incorrect input data. Thus, major efforts have been made to use the strength of neural networks while avoiding their weakness. By combining neural networks with physical/analytical models the accuracy of the latter could be increased without loosing physical information. But also the existing neural methods have been developed further to ensure the required reliability for applications in steel industry. First application attempts of standard neural networks failed primarily due to process drift, the high dimensionality and the strongly clustered nature of process data (Schlang et al. 2000). To overcome these problems, special neural networks architectures and training algorithms have been developed since.

Typically, neural networks must be trained with large amounts of data, where the required number of ‘training patterns’ scales with the dimension of the input space, the network size and the noise level of the measurements. At least 3 to 10 training patterns are necessary for each free network parameter (Broese and Löffler 2004). A high-dimensional input space additionally leads to the fact that the data cluster and that most of the space is empty, i.e. not filled with data. Often, the characteristics of the individual clusters are quite different such that local approaches like radial basis functions (RBF) networks must be favored over global ones like the classical multilayer perceptron (MLP).

There are different ways to use neural networks within the process automation system (see Figure 46.2):

(a) correction networks, correcting the output of (physics-based) process models

(b) parameter networks, serving as empirical sub-models

(c) ‘black box’ networks (not depicted), replacing the physical model entirely
46.2 Artificial Neural Networks in Process Simulation

Figure 46.2: Two ways to use neural networks. (a) The output of a given algorithm can be corrected by an additive (‘+’) or a multiplicative (‘∗’) correction network. (b) A parameter network supplies the physical model with parameters \( p \). The algorithm calculates its output \( y \) (usually used as setup values for the actuators like screwdown) from the data \( x \) (measured data and target data) using the parameters \( p \). The input data \( x \) can consist of measured data like temperature or strip thickness before rolling as well as target values like the desired strip thickness after rolling, which is different from the screwdown due to roll flattening, stand stretch etc.

Here is a simplified example to explain the different possibilities: Suppose you want to predict the exact elongation of a spring when loaded with different weights. According to Hooke’s law your physical model would be \( l = l_0 \left( 1 + m \cdot g/(A \cdot E) \right) \) with elongation \( l \), elongation without weight \( l_0 \), mass \( m \), Young’s modulus \( E \), unit area \( A \) and gravity constant \( g \). Especially \( l_0 \) will vary with temperature \( T \) as the spring steel expands or contracts but also \( E \) might change with temperature due to microscopic effects. For different weights your model will predict the expected elongation quite accurately. The model can be improved in the following ways:

(a) You could compute a temperature dependent correction factor by a correction neural network.

(b) You also could add additional physical models or parameter neural networks to calculate \( E(T) \) and \( l_0(T) \).

(c) A black box neural network in this context would learn the elongation as a function of mass and temperature.

A completely different approach is the so-called ‘short-term inheritance’. You simply measure the exact elongation and use this measurement as prediction for the next mass, if a series of same masses is to be predicted. When a different mass is to be calculated, the physical model (maybe supported by correction or parameter networks) must be used.

In process automation correction networks will be used as follows: if calculated using the physical models on their own, the setup of the mill actuators will usually be good enough such that the mill can operate, but the final product quality has not yet reached its optimum.
By taking into account additional parameters, which have not been modeled due to the not justified effort in modeling and/or data acquisition, an adaptive correction network usually can improve the final results. In case of a multiplicative correction we start with a factor of one. Online adaptation allows to train the correction network with a certain learning rate $\eta$ to fit the measured results (training patterns) better. To reduce the risk of loosing previously learned information for the sake of better accuracy, we distinguish between ‘short-term’ and ‘long-term’ inheritance. To obtain a robust result usually a smaller learning rate is chosen for long-term inheritance. Short-term inheritance often applies the following principle: if suitable, just use the measured parameters of the previous strip for subsequent strips of the same kind.

An example for a parameter network is the determination of the heat transfer coefficient in the context of heat conduction. To calculate the temperature of a hot rolled steel strip, the principal physical equations are known: Fourier’s equation of heat conduction, Stefan-Boltzmann’s law of heat radiation and Newton’s law of heat convection. To solve these equations, parameters like the heat transfer coefficient $\alpha$, the heat capacity $c$ or the thermal conductivity $\lambda$ need to be known. These parameters can either be taken from the literature (as fixed values) or can be provided by separate models (as functions of other parameters). In the case of the heat transfer coefficient, this parameter actually depends on the strip surface properties, which in turn might depend on the chemical composition of the steel strip and the amount of scale, dirt, steam or grease on the surface. Since these parameters are not known exactly, it is difficult to develop a physical model for the heat transfer coefficient $\alpha$. Instead an empirical model in the form of a parameter network seems to be more suited for this task. The chemical composition of the steel is correlated with the typical amount of scale on the surface and with the typical surface quality and therefore also with the heat transfer coefficient $\alpha$. Thus, a parameter network can model $\alpha$ as a function of the chemical composition of the steel strip only.

Another example for a parameter network is a neural network supplying the yield strength $\sigma_f$ and the frictional coefficient $\mu$ for a rolling model calculating rolling forces, rolling torques, rolling powers etc. (Kurz et al. 2001; Schlang et al. 2000; Lindhoff et al. 1994).

In some applications of parameter networks one has to deal with a specific problem: the target values for the network training can neither be estimated by direct measurements nor by inverting the physical model. The latter case may occur, if the parameter network has several outputs or if more than one parameter network is used for the same physical model (e.g. friction network and yield stress network for the rolling force model). In this case several combinations of network outputs can produce the same model output, i.e. there is no unique solution for the network targets. If this happens, the networks can only be trained by a special gradient descent algorithm. The quality of the resulting networks, i.e. of the resulting empirical models, depends on the variance of input vectors: the more different working points will be used to train the parameter networks, the better the networks will represent the reality and the lower the danger of physically implausible results will be. Nevertheless it is recommended to integrate boundary conditions (if available) into the network training, like boundaries of the input-output-gradients or of the outputs themselves. Despite of the difficulties described above a parameter network is very useful, especially if there is no other way to determine this particular model parameter. As the network output represents a physical parameter, an intuitive interpretation of the result becomes possible.
Common features of all the presented combinations of analytical models with neural networks are the following:

- the neural network is used for (online) adaptation

- experience with physical modeling can fully made use of. Also, the network training effort and the amount of data required for the actual training process is reduced

- results of the neural network, especially outside of the trained input domain, can be double checked by the physical model (for example the correction factors should be reasonably small)

- the accuracy of the physical model is often improved by taking into account second-order effects that were not included in the physical model

- a global neural network is often regarded as a black box. In combination with physical models physical coherences are still visible and can be analyzed.

- fast commissioning, because models start with physical information and adaptation takes place later. The output of a correction network can initially be set equal to a constant value and can be ‘fine-tuned’ later on.

- trained networks may hint at deficiencies in the physical models. For instance, if the correction factor shows larger deviations from its standard value of one for a certain group of strips, the physics relevant for this group may not be represented adequately by the physical model. Such a case occurs, for example, when a standard rolling model (such as the Bland-Ellis-Ford model) is applied to predict rolling forces for thin strips. As the simplifications made in the physical model are no longer adequate for the group of thin strips (indicated by unusual small correction factors), the physical model had to be corrected substantially (Kurz et al. 2001).

To sum up, neural networks are capable of learning complex correlations even when the underlying physical processes are not yet completely understood. The training of the neural networks can be performed both off-line and on-line. Due to this learning capability, neural networks have proved to be powerful tools in control applications for complex industrial processes, as shown for example by Schlang et al. 1997 and Jansen et al. 1998. If compared with other empirical methods, neural networks usually yield a better accuracy, because nonlinear phenomena can be modeled better. Using a diagnostic program which can display the input-output behavior, the transparency of a neural network is as good as the one of a sufficiently complex linear model, where a direct interpretation of coefficient values is almost as difficult as the evaluation of the weights of a multilayer perceptron network. Only by a deeper analysis of the underlying physics, which should be contained in a physical model, further insights might be gained.

Neural networks have been shown particularly advantageous in cases where either no or only fundamental physical models were available, or where parts of the physical models are still so complex that computation time and/or parameter estimation would be too computationally expensive in an industrial application. In such cases, empirical or semi-empirical models are often used in an industrial process automation system. A good example is the modeling of microstructure, which is described in the next section.
46.3 Joint Microstructure Model and Neural Network System

As indicated above, several approaches to microstructure modeling can be found in the literature. The basic empirical models for recrystallization and grain growth proposed by Sellars in the early 1980's (Sellars and Whiteman 1979; Sellars 1980) are still used today and form the starting point of many modern investigations (Lotter et al. 2002; Löffler et al. 2001; Verlinden et al. 2001; Radović 2001; Ginzburg 2003). But with the ever increasing computing capabilities also more sophisticated models have been developed: The Johnson, Mehl, Avrami, Kolmogorov (JMAK) approach to describe the recrystallization kinetics by a ‘sigmoidal’ curve \( X = 1 - \exp (-B \cdot t^n) \) was no longer considered sufficient by some researchers in the field. Models for interface migration, relaxed JMAK-models, nucleation rate and growth rate models, Monte-Carlo methods, cellular automata models and node displacement network models have been developed in the past 20 years (e.g. Vandermeer 2001). Also, grain growth has been studied in more detail over recent years. Whilst in the simple empirical models the growth exponent of the theoretically predicted parabolic growth rate (e.g. Gladman 1997) was adjusted to measured data, recent theoretical approaches tried to take into account particle pinning, solute drag and anisotropic boundary conditions. A good overview can be found for example in Maurice 2001. But in spite of the amazing computational power available these days, most analytical models still lack a generally precise answer to the question: how many austenite grains have become recrystallized after the third stand of a finishing mill and what is the average grain size? The problem of finding appropriate parameters for the empirical models has been superseded by the challenge to determine the correct initial grain size or the concentration of vacancies (Estrin 2001).

Because a direct application of the new models to a real rolling mill appears not yet feasible from a modeling point of view, and since the computational efforts are still too demanding for an online application, models used for process automation purposes still rely on semi-empirical models. Parameter values are taken from the literature and are subsequently adjusted using laboratory experiments (Lotter et al. 2002; Löffler et al. 2001; Andorfer et al. 1997).

All models consider at least parts of the following phenomena: dissolution of precipitates and grain growth in the reheating furnace; recrystallization, grain growth and strain-induced precipitation in the roughing and finishing trains; and \( \gamma - \alpha \)-transformation in the cooling section (see Figure 46.3). The results are characteristic parameters describing the microstructure of the material, such as the ferrite grain size and the fractions of ferrite, pearlite and bainite. A recent review article on microstructure models has been presented by Ginzburg 2003.

46.3.1 Physical Model

The microstructure model used in the MicroStructureMonitor system covers the influence of process parameters such as the temperature evolution in the reheating furnace, the temperatures in the stands, the strains and strain rates in each pass, the break time between two subsequent passes, as well as the strip temperature evolution in the cooling line. The various influences are schematically depicted in Figure 46.4.
In a first step, only low-carbon manganese steels were investigated. For these steel grades, algorithms based on the works of Sellars and Whiteman 1997 as well as Choquet et al. 1985 were shown to work reliably in several rolling mills. In a second step, the algorithms were extended to niobium microalloyed steels. As compared to low-carbon manganese steels, recrystallization and grain growth are hindered or at least delayed. This is due to solute Nb and precipitation of Nb(C,N) (Düfert et al. 1992; Gladman 1997). Since solute and precipitated niobium have different effects on recrystallization (Yamamoto et al. 1981), a diffusion-based
model (see Figure 46.5) has been developed to describe the dissolution of precipitation in the furnace or the rebuilding of precipitation at lower temperatures during roughing and finishing, respectively. To save computing time, a modified Aaron model (Aaron et al. 1970) is used, which corresponds well to the results of a finite volume model. In a third step, models for titanium microalloyed steel grades were developed, which deal with Ti(C,N) precipitation.

The transformation of austenite to ferrite, pearlite or bainite is modeled by algorithms described in detail by Herman et al. 1999. An overview of the various dependencies as used in our algorithm is given in Figure 46.6.

46.3.2 Physical Model plus Neural Network

The next step is the determination of mechanical properties from microstructural and process parameters. For plain (low) carbon steels a simple relationship between the ferrite grain size and the yield strength has been known since the 1950’s (Hall 1951; Petch 1953). For microalloyed steel grades precipitation hardening and softening effects due to recovery and grain growth take place. To close the gap between microstructure and mechanical properties, conventional statistical methods like multilinear regression or modern nonlinear methods like neural networks can be used. Due to the advantages described in Section 46.2, we use a correction network as schematically shown in Figure 46.7. A so-called ‘mixture of experts’, i.e. an average over the output of different types of neural networks (in our case a multilayer perceptron and a so-called Local-Global Map), is employed. The combination of mainly smooth global network types like a multilayer perceptron and mainly local types like radial basis functions or Local-Global Maps turns out to be a very robust approach, because ‘experts’ for different tasks have been ‘consulted’. The training procedure is performed off-line, i.e. a number of tensile test results are collected together with the respective input parameters of the network, and after a while the network is retrained. The neural network is capable of
performing an automatic input selection and outlier detection, and takes constraints into account. Overfitting of the neural network is prevented by means of cross-validation (Jansen et al. 1998). A combination of a physical model and a neural network facilitates double-checking the results of the neural network, especially in the extrapolation domain: the deviation of the models should be within reasonable limits.

46.3.3 Off-line System, on-line System and in-line System

The MicroStructureMonitor system can be used both off-line and on-line. Of course, an interface to the process data has to be provided in either case.

In an off-line application, the data can be entered by hand, read from a file (which is automatically generated by the on-line system) or calculated by off-line process models. Figure 46.8 shows a sample screen of the off-line system, which uses the Siemens HYBReX® system HYBREX® (Peuker et al. 1999). The off-line system allows cheap computer experiments to be conducted and enables the investigation of questions like “How much niobium is necessary to achieve a certain tensile strength value?” or “How does the coiler temperature affect the tensile strength for a certain steel grade?” and many other “What happens if…”-type questions.

In an on-line application, the measured data are supplied by the process automation system. Information about the mechanical properties of the rolled strip is available immediately
Figure 46.7: Overview of the microstructure model consisting of a physical model and an artificial neural network. The physical model takes care of the principal physics whereas the neural network is used for the precise calculation of mechanical properties.

Figure 46.8: Example screen of the off-line system. This frame depicts the results of an off-line simulation of a hot strip mill consisting of furnace, roughing mill, finishing mill and cooling section. The results for the mechanical properties as well as for the alpha iron grain size and for the fractions of ferrite, pearlite and bainite are displayed.
46.3 Joint Microstructure Model and Neural Network System

The time needed until the coil has cooled down (two to three days), samples have been cut, sent to the laboratory and analyzed can be saved. If, for example, the strip does not meet the required specifications, another one can be scheduled at once to hold the demanded shipping time. This is especially important nowadays when ‘time to market’ and ‘just in time production’ are issues of increasing importance.

Computing mechanical properties of the strip over the entire strip length is another feature of the on-line system. This improves conventional destructive quality control where typically only values from head end and tail end are available.

Based on the microstructure model, even an in-line (during running production) modification of, say, the coiling temperature is possible: the coiling temperature may be optimized such that the strip meets the target mechanical properties more accurately. Of course, this is only feasible for steel grades that react sensitively to the coiling temperature like titanium-microalloyed steel grades.

46.3.4 Results from Hot Strip Mills

As the objective of the microstructure model is to be applied in industry, it has to be tested against data from real hot strip mills. The first tests were carried out at a German medium strip rolling mill in 1996. The small dots in Figure 46.9 indicate the accuracy of the model for the tensile strength as an example. The result for the yield strength is similar, though the accuracy is not quite as good due to a slightly higher scatter in measurement. Each dot refers to one strip for which on the one hand tensile tests were carried out and on the other hand the mechanical properties were calculated by the on-line MicroStructureMonitor. The figure shows results for several thousands strips: calculated values are plotted versus measured ones. Data of low-carbon manganese steels and high strength low alloy steel grades are displayed. The accuracy is close to the precision of measurement in the region between 300 and 850 MPa.

Since mid 2001 the system has been working successfully on-line at an other German wide strip mill in Beeckerwerth. The small dark triangles indicate the results of over 1000 samples. The good results using the basis network of our first mill could be improved even further by an adaptation to the Beeckerwerth mill.

For a practical application it is important that the model (including the neural network) is transferrable to other plants – otherwise repeated model development would be necessary if there is a change in production or the mill is modernized. Process data were collected from several wide strip hot rolling mills. The author would like to point out that the model was applied without further modification of parameters in the algorithm or retraining of the neural network. The corresponding results are also displayed in Figure 46.9. They indicate that the microstructure model is transferrable from one hot strip mill to another.

When the accuracy of the predicted mechanical properties is at such high level it makes sense to actively influence process parameters in-line. With the in-line system, the coiling temperature $T_c$ can be automatically optimized to achieve a set point tensile strength $R_m$ (or at least drive $R_m$ closer to the target value).

After some initial offline tests the MicroStructureOptimizer works successfully on-line for microalloyed steel grades at a German steel mill since the beginning of 2001. An optimized coiling temperature is computed as new set point value. This new value is transferred to the cooling section’s automation system. The maximum variation has been restricted to
Figure 46.9: Results of the MicroStructureMonitor. Calculated values for the tensile strength are plotted versus measured ones. Small dots mark strips from our first hot strip mill, small dark triangles come from our second mill in Beeckerwerth, bigger symbols refer to test strips from 7 other hot strip mills. As the deviation between measurement and calculation usually is small, most points lie on the diagonal in the left picture and are therefore less pronounced than the few off-diagonal ones. The right picture shows the distribution of the relative error between calculation and measurement, which is very much Gaussian shaped.

±40°C around the original set point value such that to avoid possible problems not related to microstructure like the formation of scale at elevated temperatures. Correspondingly to that permitted variation, the tensile strength may change by several percent. Applying this procedure to selected Nb-microalloyed steel grades improved the correspondence of measured and target tensile strength by about 10–20%. Currently tests with Ti-microalloyed steel grades are under way, where the potential improvement of the tensile strength is even greater due to the higher sensitivity to the coiling temperature. The potential of this method for reducing the spread of mechanical properties around the respective target values has been estimated to be about 20% for titanium-microalloyed steels.

46.4 Conclusions

Artificial neural networks are widely used in modern process automation systems. Their application covers both modeling and adaptation tasks. The combination of physical models and neural networks (either as parameter or correction networks) was discussed. It turned out that this combination is valuable in many respects: the accuracy of the results can be improved by using neural networks for adaptation of the physical models, the physical model can be
used for double checking network results, and in contrast to a purely empirical approach all knowledge about the underlying physics can fully enter the model.

Modeling the microstructure and the mechanical properties of steel while being processed in a hot strip mill is a demanding task and can serve as an example where the described combination of physical and empirical modeling has been applied successfully. In the hybrid microstructure model presented in this paper, the neural network is employed to establish the link between chemical composition and process parameters on the one hand and mechanical properties on the other hand. The physical model is responsible for calculating recrystallization, grain growth, precipitation and phase transformation.

The prediction of the mechanical properties has become more and more important over recent years. The microstructure model has been demonstrated to produce results reliably and with an accuracy close to the precision of measurements for a wide range of steel grades (low-C Mn, Nb-microalloyed, Ti-microalloyed). Practical application of the model is beneficial in terms of a reduced number of necessary measurement samples, time savings in quality control, optimization of the production process and cost savings through reductions of alloying elements.

In conclusion, microstructure-monitoring utilizing neural networks can be regarded as a valuable tool for quality control in hot rolled steel strip production.

References

References


Index

40\textdegree[111] 350, 352
3IVM model 662, 663, 665–667

ABAQUS 520, 523
Abbruzzese–Lücke equations 312
abnormal grain growth 91, 98, 99, 101
abnormal subgrain growth 343
accelerated damage 800, 802
activation energy 61, 62, 64, 404, 407, 414–416, 593
active slip system 115, 118
adaptation 182, 186, 194, 809–811, 814, 815, 821, 822
adaptive control 182
adaptive grid 636
Al-alloys 596
ALE see Abbruzzese–Lücke equations
aluminum 459, 660, 661, 665
cold rolling 672, 673, 680, 683–685, 688, 690, 691, 700, 702, 704, 705
hot rolling 671, 672, 678, 679, 682, 683, 690, 694–696, 698–700, 705
sheet 688–690, 700, 706, 708
anelasticity 766
ANN see artificial neural network
annealing model 300
antiphase boundary 379
APB see antiphase boundary
applied stress 37, 269, 276, 280, 281
Arrhenius relation 404, 413
artificial intelligence 182
artificial neural networks 182–185, 809, 812, 822
ASY see axially symmetric model
austenite 816, 818
automotive crash 775
automotive part 752
average orientation 123, 126
axially symmetric model 583
back-diffusion 252, 260
back-scattered diffraction 120
Barlat
YLD89 764
YLD96 764
Bauschinger effect 760, 761, 765, 766, 768
Bauschinger type behavior 398
benchmark tests 759
bending 757, 758, 760
binary alloy 729
birdstrike 781–783
body-centered cubic 122
Boltzmann automaton 56, 71, 72
boundary condition 88, 89, 427
at cell boundaries 14, 16, 25, 26
at moving interfaces 4, 9
at surfaces 23
boundary migration 601
brass 531, 532, 542, 543
Brillouin zone 429
brittle-to-ductile transition 404, 405, 414, 415
bulk modulus 579
Burgers vector 406, 415, 421–423, 425, 426, 430, 464, 475
Butler-Volmer relation 568
CAFE method 247, 250, 257, 260–263
Cahn-Allen field relation 489, 493, 495
Cahn-Hilliard equation 38, 40, 46, 269, 276, 277
calculated stress-strain curve 579
calculated yield stress 580
CALPHAD 3, 4
carburization 26, 28, 29
casting 238, 661, 662
Cauchy stress 119, 123, 124
CDF  see coordination distribution function
cell diameter 576
cell size 9, 14, 15, 17, 23
cell switch 63
cellular structure 427
cementite 21, 24, 25, 30–32
CET  see columnar-to-equiaxed transition
CFD  see computational fluid dynamics
channels 759, 761
ClaNGrid 662
climb of dislocations 594, 597
cluster parameter 580, 582, 586, 590
clustering 189–191
clusters 623–626
coalescence 798, 800–802
coarse graining 66, 355, 419, 427
coarsening 37, 39, 44, 47, 269, 276, 277, 279–281, 283, 284
coherence stress 383, 384
cohesive energy 615
cohesive model 614, 616–619
cohesive strength 615, 617, 619
coil annealing 682
cold rolling 661, 666
collapsed element 608
collision rule 66–69
columnar grains 522
columnar-to-equiaxed transition 247, 250, 260
compaction 641–646, 648, 651, 655
competitive growth 23
composite
Fe/Ag 583
W/Cu 586
ZrO2/NiCr 80 20 574
composite model 437, 443, 600–602
compression test 694
computational cells 803
computational fluid dynamics 728
computational fracture mechanics 738
conservation law 422, 431
constant crack tip opening angle 610
constitutive equation 470, 471, 798
constitutive laws 572
constitutive model 675, 676, 692, 694, 706
constitutive sintering models 649
constitutive softening 785
constrained plasticity 500, 510
contact friction 758
continuity equation 304, 306
continuous cooling 22
continuous time 80, 110
continuum 57, 60, 67, 70
continuum theory 423, 424
continuum thermodynamics 489, 492
contour integral 612
convected coordinate system 747
coordination 306, 307, 309, 310, 313, 318, 319
coordination distribution function 308
cooper 202, 207, 208, 214, 216, 219
correlation function 427
Cosserat theory 502
Coulomb friction condition 750
coupling 791
CPFEM  see crystal-plasticity finite element method
crack 216–218, 510, 511, 608–613, 615, 617
crack propagation 803, 805
crack tip opening displacement 607
crack-tip plasticity 404, 405, 408–410, 414–416
cracking 642, 644, 645
creep 435, 436, 439–442
creep test 592
critical damage 776, 777
critical events 355, 358, 360–362, 364
critical phenomena 431
critical resolved shear stress 117, 368
CRSS  see critical resolved shear stress
crystal plasticity 61, 65, 114, 125, 127, 133, 138, 143, 153, 154, 159, 161, 170, 175, 489–491, 495, 496, 500–502,
Index

504, 509–511, 513, 515, 518, 521–523, 532, 534, 542, 543, 660
finite element method 448, 460, 547, 548, 550, 554, 557
theory 340
crystal structure 122
crystal-plasticity finite element method 660
crystalline strip 489, 490, 495
crystallographic
lattice 732
orientation 121, 307, 337, 548, 551
plane 114
slip 515, 516, 518, 519, 525, 529
texture 114, 120, 127, 133–135, 143, 152, 159, 170, 175, 529, 547, 555, 556, 710, 719, 722, 723
CTOA see crack tip opening angle
CTOD see crack tip opening displacement
cup drawing 173, 174, 552, 557
curvature driven grain growth 314, 315, 318, 320
curvature driven growth 355, 364
curvature driven motion 362, 363
curvature flow 321
damage 797–806
damage development 785, 786
damage percolation 785, 786, 790, 791, 794
DDD see discrete dislocation dynamics
deep drawing 547, 552, 556, 557, 703, 704, 706
deformation 299, 300
band 525
bands 523
gradient 115, 121
incompatibility 489, 490, 493, 495, 496
model 299
rate 115–119, 122–125, 128–131
spin 115, 116
texture 337
winning 529–532, 534, 536, 540–543
zone 337–341, 344, 346, 349–352
degradation 797
dendrite 247, 248, 255–257, 259, 727, 730, 732, 736, 738, 739, 742
dendrite arm 733, 736, 741
dendrite tip 727, 738
dendritic growth 634, 635, 639
design 757, 758, 760
design sensitivity 637
deviatoric crystal stress 117
DICTRA 4, 5, 7, 8, 17, 19, 28–34
diffraction technique 120
diffuse interface 37–40, 43
model 266, 269, 272, 281, 321, 323, 325, 327, 329, 334
diffuse interface model 321
diffusion 3–14, 16, 17, 19, 21, 26, 28–30, 32, 33, 60, 61, 66, 69, 321, 322, 325, 327, 332
dimples 798
direct design method 746
discrete dislocation dynamics 389, 408
discrete dislocation dynamics (2D) 404, 416
discrete dislocations 419–421, 424
dislocation 197–200, 202–217, 220–222, 662, 663, 667
annihilation 598
character 374, 377, 378
core 369, 373, 374
tensor 421–426, 501, 502
dipoles 594, 597, 602
interaction 421, 428, 430
nucleation 405, 407–409, 411, 415
pattern 431
reaction 420, 421
segmentation 371
sources 395–401
structure 120, 157
self-interaction 369, 372, 378, 385
velocity 407, 408, 411, 412, 414–416
dispersion strengthening 368, 370, 377, 378, 382, 386
dispersion field 420–422, 429
dissipation 420, 430, 431
potential 489, 492–494, 496
principle 492
dissipative 355, 358, 360
Index

domain integral  612, 613
domain structure  39, 43, 273, 280, 281, 283, 284, 286
dome height test  134–136
drag force  298, 300
draw bending  759
draw-bend test  758, 762
driving force  3–5, 7, 33, 56, 60–62, 72
Drucker-Prager-Cap model  642–644
ductile damage  798
ductile fracture  777, 778, 786, 790
ductile tearing  610, 617, 619
dynamic recovery  596, 598
dynamic recrystallization  85, 108, 109
dynamics of dislocations  199, 200
earing  698–700, 703–706
electron back-scattered diffraction  93, 120, 262
effective strain-rate  748
elastic anisotropy  547, 556
elastic energy  40, 45, 68, 269–272, 274–276, 278, 280, 281, 283
elastic modulus  116
elastic stretch  115, 117, 124
elastic-plastic fracture mechanics  607, 618
elastic-plastic model  515
elastic-plastic springback  763
elasto-plastic formulation  747
field variable  673
finite difference method  727–729, 733, 734
finite element method  61, 65, 158, 173, 203, 209,
finite element analysis  746, 753
finite element method  114, 130, 132, 133,
138, 140, 435, 448, 515, 520–523,
547, 549, 552, 561, 562, 571, 572,
634, 638, 642, 649, 653, 660–668,
693, 694, 706, 728, 746
finite element modeling  758, 759
fix-end torsion  478–480
flanged channel  759, 761
flow rule  143, 147, 148, 153, 154, 159, 168,
173
flow stress  120, 127–129, 133–135, 139, 436,
440, 441, 671, 674–677, 679, 680,
683, 691, 692, 694, 700
flow surface  123, 124, 126–131
fluid dynamics  56, 66, 67, 69, 72, 73
fluid flow  623, 625, 634, 636, 727, 728, 737–
739, 742
foreign object damage  401, 411
forging  709, 710, 712, 713, 715, 716, 718,
719
formability  124, 133–136, 143, 144, 167, 688,
693, 700, 702, 704, 706, 785, 791,
792
forming  448, 452
Fourier transform  466, 468
fracture  606, 607, 609–611, 614, 615, 617–
619
mechanics  606–608, 610, 611, 613, 617,
618
propagation  776
Index

- toughness 404–407, 409, 412, 413, 415, 416
- Frank–Read source 393, 396
- free energy 323, 324, 326, 327, 489, 492–496
- free standing film 392, 393, 396, 398, 399
- free surface 392
- free-end torsion 477, 478, 480
- freezing 731, 733, 734
- frequency distribution 583
- gas-tungsten-arc 737
- generalized continua 500, 502, 509, 510
- generalized internal variables 492
- geometric hardening 120
- geometrically necessary dislocations 489, 501, 504, 505, 510
- GIA see grain interaction
- GIA model 454, 664, 665
- Ginzburg-Landau equation 430
- GND see geometrically necessary dislocations
- governing equation 321, 323, 324, 327, 334
- gradient of internal variable 500, 505
- grain 119, 138
  - boundary 60–63, 65, 337, 341, 344, 347, 390, 392, 396, 398, 401
  - boundary curvature 294, 296, 297, 302, 304
  - boundary diffusion 650, 651
  - boundary energy 304, 306, 314
  - boundary mobility 61, 63–65, 294, 296, 300, 302
  - boundary motion 304
- coarsening 648, 651
- fragmentation 525
- interaction 458, 683
- selection 250, 257, 262
- shape 463, 471, 474
- size effects 511, 564
- structure 37, 248–250, 257, 262, 263, 327, 331, 334
- Green function method 466
- green strength 645
- grid structure 72, 73
- growth 797–800, 802, 805, 806
- metastable phases 25
- stable phases 25
- growth kinetics law 250, 254, 255
- GTA see gas-tungsten-arc
- Gurson–Tvergaard–Needleman model 800
- HAGB see high angle boundaries
- Hall–Petch effect 490, 498, 500, 511
- Hall–Petch model 401
- Hamiltonian 422, 428
- hardening 115, 118, 120, 132, 133, 464, 472, 474, 475, 477–479, 481, 484
- hardening behaviour 490, 493, 495
- HEM see homogeneous equivalent medium or homogeneous effective medium
- Herring Condition 355, 358, 360, 361, 364
- hexagonal close-packed 122, 280, 481
- hexagonal metals 530
- high angle boundaries 343
- high resolution transmission electron microscopy 249
- high strain rate 202
- high strength low alloy 812
- Hill quadratic yield 764
- Hillert distribution 306
- homogeneous effective medium 463
- homogeneous equivalent medium 709
- hot deformation 340, 342, 344
- hot rolling 660, 662–665
- hot strip rolling 809
- HRR-field 611, 612, 615
- HRTEM see high resolution transmission electron microscopy
- HSLA see high strength low alloy
- hydroforming deep drawing 744
- hydrostatic pressure 144, 146, 159, 162, 165, 170
- ideal forming theory 746, 753
  - bulk forming 746
  - design theory 747
  - sheet forming 746
- image force 200
impact 780, 781
implicit schemes 760
inclusion in a homogeneous effective medium 122
inclusion problem 509
incoherent particles 377, 378, 380, 384
inelastic deformation 114, 123
input pattern 183, 184, 189, 190
integration scheme 759
inter-particle dilational spacing 785, 788
interaction rules 72, 73
interface 321–323, 326–328, 331
dislocations 390–395
element 614, 615
velocity 730, 734
internal degrees of freedom 492
internal stress 757, 767
inverse problems 634, 636
inverse update scheme 297
IPDS see inter-particle dilational spacing
IPE see isotropic plasticity equivalent
irregular cellular automata 292, 293
isotropic hardening 438, 763–766
isotropic plasticity equivalent 750
jerky glide 599
J-integral 607, 609, 611–614
kinematic framework 437, 442–445
kinematics 472, 485
kinetic equation 436–439, 441, 443
kinetics 4, 13, 14, 19, 21, 24, 29, 30, 32, 60, 61, 63, 77, 80, 85, 90, 94–96, 99–101
Kröner model 424
Lagrangian scheme 516–518
Lamé model 455–458, 460
Lankford value 547
large strain 709, 718, 723
lattice Boltzmann automaton 66, 69, 71
lattice gas 56, 66–71
lattice mismatch 368–370, 381–383, 385, 386
lattice orientation 114, 115, 119, 132, 138, 139
lattice rotation 115
Laves phase 33
LE see local equilibrium
LEFM see linear elastic-fracture mechanics
LENP see local equilibrium with negligible partitioning
line energy 368, 369
line tension 368, 369, 379, 385, 392, 393
model 420
linear elastic-fracture mechanics 607, 776
linear elasticity 420, 423
linearization 463–465, 471
liquid phase 727, 734, 735
local approach 797
local equilibrium 4
with negligible partitioning 3, 4, 13
with partitioning 12
localisation effects 803
localization 172–175, 463, 469–471
M–C2 24, 26, 29, 30
M23C6 24–26
macroscopic behavior 122
magnetic energy 40
matricity 574–579
model 576
matrix erosion tessellation 788
MCCG see Monte Carlo coarse graining
mechanical properties 809, 811, 812, 818–823
mechanical test 144, 167
mechanical threshold strength 133
melting 727, 730, 731, 733, 734
membrane element 747
mesh dependence 803
mesoscale 355, 364, 365, 419, 421, 422, 431, 433, 798
metal fracture 782
metallurgy 709, 723
micro-tensile tests 389–391, 401
micro-texture 521
microchemistry 668
micromechanical models 526
microsegregation 256, 727
evolution 37, 44, 46, 49, 60, 266, 267, 269, 270, 272, 277, 278, 285, 287
formation 321, 334
modeling 809, 816
simulation 688, 696
microstructure-based model 435, 436, 445
MicroStructureMonitor 811, 816, 817, 819, 821, 822
microvoids 785
minimum creep rate 593, 601
minimum plastic work path 746, 747
proportional true strain path 748
Mises equivalent stress and strain 579
misorientation 292, 300, 302
mixture of experts 194
MLP see multilayer perceptron
mobility 6, 11, 61–63, 304, 314
modulus mismatch 386
molybdenum 214
Monte Carlo coarse graining 428
Monte Carlo method 76, 79, 80, 100
Monte Carlo technique 341, 343
morphology 728, 730, 732, 736
moving interfaces 4, 9
MTS see mechanical threshold strength
Mullins Equation 355, 357, 360
multi-cell calculations 23
multi-component systems 3–5
multi-grain 454, 458
multi-parameter model 437, 441
multi-step 710, 723
multicomponent alloys 325, 329
multifield model 489
multifunctional materials 560
multilayer perceptron 182, 185–187, 191, 192, 812
multiple solutions 451
multiplicative decomposition 115
multipole expansion 425
multiscale approach 489
multiscale asymptotic developments 506
mushy zone 733, 737, 738, 740

nanscale 114
Navier–Stokes equation 66, 68, 69
ND see neutron diffraction
negligible partitioning 3, 4, 13
network training 186, 191
neural networks 231
neutron diffraction 120
Nix–Freund model 390–394, 396
non-equilibrium 72
non-faceted growth 736
non-local behaviour 500
non-local effects 424
nucleation 91, 92, 96–100, 108, 798, 800–803
law 250, 255
mechanism 341
nucleus 337, 339, 340, 347, 350
numerical simulation 327
obstacle arrangement 369, 378
obstacle potential 326, 327
obstacle stress 370, 379, 382
ODF see orientation distribution function
olivine 484, 485
OM see optical microscopy
on-line model 819, 821
one-parameter model 436, 437
OOF 559–564, 567, 568, 570, 571, 573
OOF2 559, 562, 564, 571–573
optical microscopy 249
optimal design 636, 638, 639
optimization 746, 747, 750, 757
optimum initial blank 746, 747, 749, 750
optimum preform 751–753
procedure 748–750
order parameter 421
order strengthening 368, 370, 379, 382, 386
ordered intermetallic phase 45, 277
orientation 56, 57, 60–62, 64, 65
distribution 119–121, 123
distribution function 114, 119, 120, 125,
299, 448, 458, 547, 548, 551
gradient 338, 339
selection mechanism 350
oriented growth 347
Orowan process 368, 377–379, 382
overaged crystal 385
overfitting 182, 192, 194, 195
pair correlation function 427
para-equilibrium 3, 4, 14, 17, 34
parallel computing 83
partial dislocations 384, 385
particle 661, 662, 666
arrangement 369, 375–377
cluster 786
<table>
<thead>
<tr>
<th>Term</th>
<th>Page Numbers</th>
</tr>
</thead>
<tbody>
<tr>
<td>distribution</td>
<td>383</td>
</tr>
<tr>
<td>pinning</td>
<td>86, 100, 102</td>
</tr>
<tr>
<td>radius distribution</td>
<td>368, 376, 379</td>
</tr>
<tr>
<td>shearing</td>
<td>375</td>
</tr>
<tr>
<td>size distribution</td>
<td>376, 386</td>
</tr>
<tr>
<td>strengthening</td>
<td>368</td>
</tr>
<tr>
<td>velocity</td>
<td>67, 68</td>
</tr>
<tr>
<td>particle-stimulated nucleation</td>
<td>337, 345–347, 672</td>
</tr>
<tr>
<td>partitioning</td>
<td>12, 13</td>
</tr>
<tr>
<td>path dependence</td>
<td>613, 614, 619</td>
</tr>
<tr>
<td>pattern</td>
<td>213, 215</td>
</tr>
<tr>
<td>pattern recognition</td>
<td>182, 183, 187, 188</td>
</tr>
<tr>
<td>PCA</td>
<td>see principal component analysis</td>
</tr>
<tr>
<td>PDF</td>
<td>see probability density function</td>
</tr>
<tr>
<td>PE</td>
<td>see para-equilibrium</td>
</tr>
<tr>
<td>peak-aged crystal</td>
<td>382, 385</td>
</tr>
<tr>
<td>penalty method</td>
<td>748, 753</td>
</tr>
<tr>
<td>penalty term</td>
<td>424</td>
</tr>
<tr>
<td>percolation</td>
<td>625</td>
</tr>
<tr>
<td>perforation simulation</td>
<td>782, 783</td>
</tr>
<tr>
<td>periodic boundary condition</td>
<td>427</td>
</tr>
<tr>
<td>periodic homogenization</td>
<td>508</td>
</tr>
<tr>
<td>PFT</td>
<td>see pseudo-front tracking</td>
</tr>
<tr>
<td>phase change</td>
<td>709</td>
</tr>
<tr>
<td>phase diagram computations</td>
<td>250, 262</td>
</tr>
<tr>
<td>phase field</td>
<td>489, 493</td>
</tr>
<tr>
<td>phase separation</td>
<td>38, 266, 275, 278</td>
</tr>
<tr>
<td>phase transformations</td>
<td>37, 39, 42, 47, 60, 72, 266, 267, 269, 274, 275, 281, 283, 285, 286</td>
</tr>
<tr>
<td>phase transition</td>
<td>267, 283, 321, 323, 327, 329, 334, 727, 729</td>
</tr>
<tr>
<td>phase-field method</td>
<td>37, 38, 47–49, 247, 248, 250, 259, 266, 267, 269, 280, 281, 284–286, 634, 638</td>
</tr>
<tr>
<td>phase-field model</td>
<td>321–323, 326, 327, 329, 334, 420, 421, 424, 430</td>
</tr>
<tr>
<td>physical models</td>
<td>809, 813, 815, 822</td>
</tr>
<tr>
<td>piecewise linear interface calculation</td>
<td>251</td>
</tr>
<tr>
<td>piezoelectrics</td>
<td>559, 561, 563, 564</td>
</tr>
<tr>
<td>plane strain compression</td>
<td>522, 542, 543</td>
</tr>
<tr>
<td>plastic</td>
<td></td>
</tr>
<tr>
<td>deformation</td>
<td>209, 221, 435, 436, 440, 441, 443, 445</td>
</tr>
<tr>
<td>deformation gradient</td>
<td>115</td>
</tr>
<tr>
<td>flow</td>
<td>114, 124, 129</td>
</tr>
<tr>
<td>potential</td>
<td>123, 125, 127, 132, 140</td>
</tr>
<tr>
<td>stiffness</td>
<td>122, 123</td>
</tr>
<tr>
<td>work</td>
<td>147, 148, 170, 171</td>
</tr>
<tr>
<td>zone</td>
<td>406, 409</td>
</tr>
<tr>
<td>pluggable physics</td>
<td>571</td>
</tr>
<tr>
<td>point defects</td>
<td>422</td>
</tr>
<tr>
<td>Poisson’s ratio</td>
<td>579</td>
</tr>
<tr>
<td>pole figure</td>
<td>137</td>
</tr>
<tr>
<td>polycrystal</td>
<td>515, 520, 521, 524, 530, 531, 533, 540–543, 751</td>
</tr>
<tr>
<td>plasticity</td>
<td>120, 121, 140</td>
</tr>
<tr>
<td>yield surface</td>
<td>479, 480</td>
</tr>
<tr>
<td>polycrystalline aggregates</td>
<td>507, 511</td>
</tr>
<tr>
<td>polycrystalline material</td>
<td>114</td>
</tr>
<tr>
<td>polygonization</td>
<td>428</td>
</tr>
<tr>
<td>polyphase material</td>
<td>114</td>
</tr>
<tr>
<td>porosity</td>
<td>158, 159, 798–802</td>
</tr>
<tr>
<td>porous materials</td>
<td>799</td>
</tr>
<tr>
<td>Potts model</td>
<td>77, 83</td>
</tr>
<tr>
<td>powder</td>
<td>641–647, 652, 655</td>
</tr>
<tr>
<td>power law</td>
<td>117, 129</td>
</tr>
<tr>
<td>PPM2OOF</td>
<td>561, 562</td>
</tr>
<tr>
<td>precipitates</td>
<td>158, 274, 276, 278–280</td>
</tr>
<tr>
<td>predictive fracture</td>
<td>775</td>
</tr>
<tr>
<td>preform design</td>
<td>747, 751</td>
</tr>
<tr>
<td>primary creep</td>
<td>593, 600</td>
</tr>
<tr>
<td>primary phase formation</td>
<td>250, 253</td>
</tr>
<tr>
<td>primary static recrystallization</td>
<td>63</td>
</tr>
<tr>
<td>principal component analysis</td>
<td>191, 193</td>
</tr>
<tr>
<td>probability density function</td>
<td>76</td>
</tr>
<tr>
<td>process automation</td>
<td>809, 810, 812, 813, 815, 816, 822</td>
</tr>
<tr>
<td>process chain</td>
<td>670, 672, 673, 678, 682</td>
</tr>
<tr>
<td>process simulation</td>
<td>812</td>
</tr>
<tr>
<td>pseudo-front tracking method</td>
<td>247, 250, 251, 259, 262</td>
</tr>
<tr>
<td>PSN</td>
<td>see particle-stimulated nucleation</td>
</tr>
<tr>
<td>quasistatic loading</td>
<td>798</td>
</tr>
<tr>
<td>r-value</td>
<td>547, 552</td>
</tr>
<tr>
<td>radial basis function networks</td>
<td>182, 188–190, 192</td>
</tr>
<tr>
<td>radial basis functions</td>
<td>188–190, 812</td>
</tr>
<tr>
<td>radius of curvature</td>
<td>758</td>
</tr>
<tr>
<td>rate dependence</td>
<td>404</td>
</tr>
<tr>
<td>rate sensitivity</td>
<td>464, 516, 537</td>
</tr>
<tr>
<td>RBF</td>
<td>see radial basis functions</td>
</tr>
</tbody>
</table>
Index

RBFN see radial basis function networks
reaction–diffusion 67
reaction-diffusion model 431
rechargeable batteries 559, 561
recrystallization texture 338–340, 344
reference frame 5–8
relaxed constraints 452
remeshing 709, 715, 723
renormalization 431
representative volume element 798
residual stresses 575, 580, 582, 589, 767
resolved shear stress 116
resonance method (ASTM C1259-94) 586
reverse loading 150, 156, 170
rheology 623
rigid-plastic formulation 747
Rodrigues space 120
roll force 685, 686
rolled products 671–673
rolled stainless steel 119, 120
rolling mill 809, 810, 812, 816, 817, 821
room-temperature creep 767
rotation 115, 117, 132–134, 474
Rousselier model 800, 803
R/t ratios 763, 766
r-value 703
scaling relation 404, 408, 414–416
scanning electron microscopy 249
Schmid tensor 115, 464, 473
SDD see stochastic dislocation dynamics
SDF see size distribution functions
secant approximation 465, 471
second gradient 500–502, 504, 505, 510
second phase 158
second phase particle 785–787, 789, 791
secondary creep 593, 594
secondary phases formation 252, 254
segregation 239, 661
self organizing maps 191
self piercing rivets 779
self-consistent 422, 428, 431
embedded cell model 576, 578
model 454, 458, 460, 472, 484
theories 422
SEM see scanning electron microscopy
semi-brittle fracture 404, 406
semisolids 623
sensitivity analysis 637
SFE see stacking fault energy
shape factor 308, 314, 316, 317
shapeless cellular automaton 293
shear stress 116
sheet forming 143, 152, 159, 171, 173, 175
sheet forming processes 746
sheet metal 791
sheet processing 689
sheet tension 757, 760–762, 764, 766, 768
shell elements 763
shock wave 214, 219
SIBM see strain induced boundary migration
sidewall curl 758, 761
simple shear 489, 490, 495
simulation 355, 360, 363, 364, 820
single dislocations 597, 602
sinter distortions 654, 655
sinter forging tests 651, 652
sintering 641, 642, 648–655
size distribution functions 304, 308
skeleton lines 575, 576, 578, 581, 584
skew symmetric tensor 115
slanted fracture 802
slip 462, 464, 472–475, 477, 478, 481–485
direction 115, 117, 118
hardening 540, 541
plane 115, 117, 118
systems 427, 448–451, 453, 458, 459
SLR see special linear relationship
snap back 806
softening 606, 619
solid elements 763
solid state transformation 727

solute diffusion 727, 729

SOM see self organizing maps

source model 391, 397

SPC see strong pair coupling

special linear relationship 312

spin 121, 122

SPR see self piercing rivets

springback 757–762, 764–768

SST see standard stereographic triangle

stacking fault energy 529, 530, 533

stainless steel 133, 138

standard stereographic triangle 107

StaRT model 664–666

state variable 670, 673–676, 686

statistical mechanics 427

steel 456, 458, 459, 661, 667, 710, 715, 716, 718, 721

steel industry 811

stiffness 123, 124, 130, 131, 133, 138

stochastic dislocation dynamics 206

stochastic force 206

strain gradient 504–506, 510

strain hardening 154, 156, 157, 172, 530–532, 534, 540, 541, 543, 688, 691, 692, 700, 701, 703, 760

strain induced boundary migration 347

strain rate 114, 115, 119, 121, 123–125

strain-rate potential 159, 167, 747

Srp93 747

Srp98 750

strengthening mechanisms 369, 375, 377, 386

stress equilibrium 371, 375

stress exponent 598, 599

stress field 199, 205, 208, 211–214, 216, 221

stress space 123, 125–129

stress-strain curve 398, 399, 593

stretch bending 759

stretch flange 785, 787, 791–794

strong pair coupling 380

structure evolution equation 438, 441, 443, 445

subgrain boundaries 596, 600, 602, 603

subgrain strengthening 601

superalloy 380, 386

superdislocation 379, 380

superposition principle 211

supervised training 191

surface diffusion 650

suspensions 623, 624

switching probability 61–63

switching rule 61, 63

symmetric tensor 115

symmetry 422

tangent approximation 471

tangent modulus 579, 580

target pattern 183

Taylor factor 120, 126, 127, 451, 675, 677, 703

Taylor model 125, 520, 523

TEM see transmission electron microscopy

temperature field 728, 736, 739

temperature simulation 690, 692, 694–698, 705

tempered martensitic steels 600, 601

tensile strength 809, 811, 812, 819, 821, 822

tension-compression 761

ternary alloy 323, 332, 333

tertiary creep 593, 594


component 547–553, 556, 557

development 102

effects 570

evolution 120, 250, 262

simulation 684, 694, 695, 698, 703, 705

thermal activation 206, 210, 404, 405, 416

THERMO-CALC 4

thermodynamics 3, 4, 30, 431

thermomechanical

process 114

processing 337, 688, 699, 703, 706

simulation 674

Theta projection concept 594


plasticity 389, 390, 393, 395

thixotropy 623

threading dislocations 390, 395

through process model 670
through-process modeling 660, 661
through-thickness 661
through-thickness compression 481
Ti-alloys 598
time scale 419, 422, 430
time scaling 84
Titanium 532
tool 758, 761, 762
tooling 757–760, 762, 768
torsion 462, 477–480
TPM see through process model
training data 183, 194
training patterns 812, 814
transfer coefficient 28–32
transformation 3, 4, 10, 13, 16, 30, 32
transformation probability 63, 65
transformation rules 56–60
transition metals 404
transmission electron microscopy 393
triaxial tests 643, 645
triple junction 296, 297, 328, 329
triple point 309
trivalent mosaics 307, 314
TTC see through-thickness compression
tube hydroforming 744–747
tungsten 404, 406–408, 414, 415
turbulence 66
Turnbull equation 61
twin hardening 541, 542
twin volume fractions 533, 542
twin-matrix interfaces 540, 541
twinned fraction 475, 476
twinning 463, 464, 472–476, 481, 482
two-parameter-model 597, 602
UMAT subroutine 520
underaged crystal 382, 385
undercooling 247, 248, 254–256, 259, 262, 727, 730, 734
unsupervised training 191
upper bound 62
USER-materials 801
validation data 183, 194
variational method 486
velocity gradient 115, 116, 121
virtual crack extension 609, 612
virtual experiments 560, 562
viscoplastic inclusion 467
viscosity 623–625, 627
viscous glide 599
Voce 474, 481, 483
void coalescence 790–792
void volume fraction 798, 800–802, 804
VPSC model 707
wafer curvature 389–391
weighting factor 577, 578
weighting function 132
weld pool 727, 728, 730, 736–738, 740, 742
welding 727, 728, 730, 737–739, 742
wetting 321
work hardening 668, 672, 675, 678, 685, 692
X-ray diffraction 120
X-ray texture 547, 548
XRD see X-ray diffraction
yield behavior 801
yield function 143, 144, 147, 148, 153, 157, 159, 160, 162–166, 170, 171, 173
yield locus 706
yield strength 674, 680, 685, 699, 702, 703, 809, 811, 814, 818, 821
yield stress 227, 231, 232, 239, 240, 438, 440
yield stress potential 747
yield surface 125, 127, 132–135
analytical 557
zirconium 481